LNCS 5727

Tom Gross Jan Gulliksen
Paula Kotzé Lars Oestreicher
Philippe Palanque Raquel Oliveira Prates
Marco Winckler (Eds.)

Human-Computer Interaction – INTERACT 2009

12th IFIP TC 13 International Conference
Uppsala, Sweden, August 2009
Proceedings, Part II

2 Part II

ifip

Springer

Lecture Notes in Computer Science 5727

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

Tom Gross Jan Gulliksen
Paula Kotzé Lars Oestreicher
Philippe Palanque Raquel Oliveira Prates
Marco Winckler (Eds.)

Human-Computer Interaction – INTERACT 2009

12th IFIP TC 13 International Conference
Uppsala, Sweden, August 24-28, 2009
Proceedings, Part II

Springer

Volume Editors

Tom Gross
Bauhaus University Weimar, 99423 Weimar, Germany
E-mail: email@tomgross.net

Jan Gulliksen
Uppsala University, 75105 Uppsala, Sweden
E-mail: jan.gulliksen@it.uu.se

Paula Kotzé
Meraka Institute, 0001 Pretoria, South Africa
E-mail: paula.kotze@meraka.org.za

Lars Oestreicher
Uppsala University, 75120 Uppsala, Sweden
E-mail: larsoe@dis.uu.se

Philippe Palanque
Marco Winckler
University Paul Sabatier, 31062 Toulouse Cedex, France
E-mail: {palanque, winckler}@irit.fr

Raquel Oliveira Prates
Universidade Federal de Minas Gerais, 31270010 Belo Horizonte, MG, Brazil
E-mail: rprates@dcc.ufmg.br

Library of Congress Control Number: 2009932606

CR Subject Classification (1998): H.5.2, H.5.3, H.3-5, I.2.10, D.2, K.3-4, K.8

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

ISSN 0302-9743
ISBN-10 3-642-03657-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-03657-6 Springer Berlin Heidelberg New York

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12740727 06/3180 5 4 3 2 1 0

Foreword

INTERACT 2009 was the 12th of a series of INTERACT international conferences supported by the IFIP Technical Committee 13 on Human–Computer Interaction. This year, INTERACT was held in Uppsala (Sweden), organized by the Swedish Interdisciplinary Interest Group for Human–Computer Interaction (STIMDI) in cooperation with the Department of Information Technology at Uppsala University.

Like its predecessors, INTERACT 2009 highlighted, both to the academic and to the industrial world, the importance of the human–computer interaction (HCI) area and its most recent breakthroughs on current applications. Both experienced HCI researchers and professionals, as well as newcomers to the HCI field, interested in designing or evaluating interactive software, developing new interaction technologies, or investigating overarching theories of HCI, found in INTERACT 2009 a great forum for communication with people of similar interests, to encourage collaboration and to learn.

INTERACT 2009 had Research and Practice as its special theme. The reason we selected this theme is that the research within the field has drifted away from the practical applicability of its results and that the HCI practice has come to disregard the knowledge and development within the academic community. However, to continue to develop and prosper, the HCI community needs to pay more attention to the cooperation and mutual learning between research and practice. We witness accidents happening in health care, airplanes crashing due to problems in their operation, we see tremendous money losses due to inefficiencies in computer support. On a smaller scale we find many tools and applications that are, if not difficult, annoying to use for reasons that are well known. At the same time we see an amazing development of new IT designs, games, robots and applications designed for pleasure and fun. IT is an essential source for efficiency and productivity in work as well as a large source of problems. For the future we need to learn from the positive development and reach for an understanding of the mechanisms behind the successes in order for us to let such processes influence the more challenged application domains. We hope that INTERACT 2009 provided such an inspiring meeting between research and implementation.

INTERACT 2009 brought the conference to Sweden and the Nordic countries. Through conferences such as NordiCHI, the Nordic countries have come to be recognized as one of the major local communities within the field of HCI. Nordic participants are one of the biggest regional participation categories at HCI conferences in the world. The Nordic tradition of HCI focuses on utility, equality, participation and quality of our everyday life and work. We aim at continuing these traditions while being in the forefront of development of the HCI field.

We hope that INTERACT 2009 will be remembered as a conference that brought research and practice closer together, focusing on making the computerized world a better place for everyone, regardless of functional abilities or particular requirements.

August 2009

Jan Gulliksen
Lars Oestreicher

Note from the Technical Co-chairs

You are holding the proceedings of the 12th Human–Computer Interaction Conference INTERACT 2009. The creation of this document is really a testament to the enthusiasm and collegial spirit amongst the HCI research and practitioner community worldwide. Thank you to all the authors who chose INTERACT 2009 as the vehicle to publish their research: a total of 723 submissions were received for the conference this year. Of these submissions, 244 were accepted for presentation at the conference. These include:

- 104 Full Research Papers
- 79 Short Research Papers
- 7 Interactive Poster Papers
- 17 Workshops
- 12 Tutorials
- 3 Special Interest Groups
- 6 Demonstrations
- 3 Panels
- 13 Doctoral Consortium Papers

The acceptance rate for the research papers was 29% of full papers submitted for review and 20% of initial abstracts submitted. The acceptance rate for the short research papers and interactive posters was 39%. Each research paper was reviewed by between three and six members of the International Programme Committee.

Our sincere gratitude goes to the all the International Programme Committee members who willingly helped out and ensured the high quality of the INTERACT Conference papers was properly maintained. Although some people had to be bullied into reviewing (sorry about that), everyone submitted their reviews on time without a murmur of complaint. Thank you all for the effort that you so obviously put into this task.

Finally we wish to express a special thank you to the Proceedings Publication Chair, Marco Winckler, who painstakingly put this volume together.

August 2009

Paula Kotzé
Tom Gross

IFIP TC13

Established in 1989, the International Federation for Information Processing Technical Committee on Human–Computer Interaction (IFIP TC13) is an international committee of 30 member national societies and 7 Working Groups, representing specialists in human factors, ergonomics, cognitive science, computer science, design and related disciplines. INTERACT is its flagship conference, staged biennially in different countries in the world.

IFIP TC13 aims to develop the science and technology of human–computer interaction (HCI) by encouraging empirical research, promoting the use of knowledge and methods from the human sciences in design and evaluation of computer systems; promoting better understanding of the relation between formal design methods and system usability and acceptability; developing guidelines, models and methods by which designers may provide better human-oriented computer systems; and, cooperating with other groups, inside and outside IFIP, to promote user-orientation and humanization in system design. Thus, TC13 seeks to improve interactions between people and computers, encourage the growth of HCI research and disseminate these benefits worldwide.

The main orientation is toward users, especially the non-computer professional users, and how to improve human–computer relations. Areas of study include: the problems people have with computers; the impact on people in individual and organizational contexts; the determinants of utility, usability and acceptability; the appropriate allocation of tasks between computers and users; modelling the user to aid better system design; and harmonizing the computer to user characteristics and needs.

While the scope is thus set wide, with a tendency toward general principles rather than particular systems, it is recognized that progress will only be achieved through both general studies to advance theoretical understanding and specific studies on practical issues (e.g., interface design standards, software system consistency, documentation, appropriateness of alternative communication media, human factors guidelines for dialogue design, the problems of integrating multimedia systems to match system needs and organizational practices, etc.).

IFIP TC13 stimulates working events and activities through its Working Groups (WGs). WGs consist of HCI experts from many countries, who seek to expand knowledge and find solutions to HCI issues and concerns within their domains, as outlined below.

In 1999, TC13 initiated a special IFIP Award, the Brian Shackel Award, for the most outstanding contribution in the form of a refereed paper submitted to and delivered at each INTERACT. The award draws attention to the need for a comprehensive human-centered approach in the design and use of information technology in which the human and social implications have been taken into account. Since the process to decide the award takes place after papers are submitted for publication, the award is not identified in the proceedings.

WG13.1 (Education in HCI and HCI Curricula) aims to improve HCI education at all levels of higher education, coordinate and unite efforts to develop HCI curricula and promote HCI teaching.

WG13.2 (Methodology for User-Centred System Design) aims to foster research, dissemination of information and good practice in the methodical application of HCI to software engineering.

WG13.3 (HCI and Disability) aims to make HCI designers aware of the needs of people with disabilities and encourage development of information systems and tools permitting adaptation of interfaces to specific users.

WG13.4 (also WG2.7) (User Interface Engineering) investigates the nature, concepts and construction of user interfaces for software systems, using a framework for reasoning about interactive systems and an engineering model for developing user interfaces.

WG13.5 (Human Error, Safety and System Development) seeks a framework for studying human factors relating to systems failure, develops leading edge techniques in hazard analysis and safety engineering of computer-based systems, and guides international accreditation activities for safety-critical systems.

WG13.6 (Human-Work Interaction Design) aims at establishing relationships between extensive empirical work-domain studies and HCI design. It will promote the use of knowledge, concepts, methods and techniques that enable user studies to procure a better apprehension of the complex interplay between individual, social and organizational contexts and thereby a better understanding of how and why people work in the ways that they do.

WG13.7 (Human–Computer Interaction and Visualization) is the newest of the working groups under the TC.13. It aims to establish a study and research program that will combine both scientific work and practical applications in the fields of Human–Computer Interaction and Visualization. It will integrate several additional aspects of further research areas, such as scientific visualization, data mining, information design, computer graphics, cognition sciences, perception theory, or psychology, into this approach.

New Working Groups are formed as areas of significance to HCI arise. Further information is available at the IFIP TC13 website: http://www.ifip-hci.org/

IFIP TC13 Members

Australia
Judy Hammond
Australian Computer Society

Austria
Horst Hörtner
Austrian Computer Society

Belgium
Monique Noirhomme-Fraiture
Federation des Associations Informatiques de Belgique

Brazil
Simone Diniz Junqueira Barbosa
Brazilian Computer Society (SBC)

Canada
Gitte Lindgaard (Secretary)
Canadian Information Processing Society

China
Chengguing Zong
Chinese Institute of Electronics

Czech Republic
Vaclav Matousek
Czech Society for Cybernetics and Informatics

Denmark
Annelise Mark Pejtersen (Chair)
Danish Federation for Information Processing

Finland
Kari-Jouko Räihä
Finnish Information Processing Association

France
Philippe Palanque
Societe des Electriciens et des Electroniciens (SEE)

Germany
Horst Oberquelle
Gesellschaft fur Informatik

Greece
Nikos Avouris
Greek Computer Society

Iceland
Marta Kristin Larusdottir
The Icelandic Society for Information Processing (ISIP)

India
Anirudha Joshi
Computer Society of India

Italy
Fabio Paternò
Italian Computer Society

Ireland
Liam J. Bannon
Irish Computer Society

Japan
Masaaki Kurosu
Information Processing Society of Japan

New Zealand
Mark Apperley
New Zeeland Computer Society (NZCS)

Nigeria
Chris C. Nwannenna
Nigeria Computer Society

Norway
Norwegian Computer Society

Poland
J. L. Kulikowski
Poland Academy of Sciences

Portugal
Joaquim A. Jorge
Associacao Portuguesa de Informatica

Singapore
Kee Yong Lim
Singabore Computer Society

South Africa
Janet L. Wesson (Vice chair)
The Computer Society of South Africa

Spain
Julio Abascal
Federacion Espanola de Sociedades de Informatica

Sweden
Lars Oestreicher
Swedish Computer Society

Switzerland
Ute Klotz
Swiss Federation for Information Processing

The Netherlands
Gerrit C. van der Veer
Nederlands Genootschap voor Informatica

UK
Phil Gray
British Computer Society

USA-Based
John Karat
Association for Computing Machinery (ACM)
Nahum Gershon
The Computer Society, Institute of Electrical & Electronics Engineers (IEEE-CS)

Expert Members
Jan Gulliksen
Sweden
Tom Gross
Germany

Working Group Chairpersons

WG13.1 (Education in HCI and HCI Curricula)
Paula Kotzé, *South Africa*
WG13.2 (Methodology for User-Centered System Design)
Peter Forbrig, *Germany*
WG13.3 (HCI and Disability)
Gerard Weber, *Germany*
WG13.4 (also 2.7) (User Interface Engineering)
Nick Graham, *Canada*
WG13.5 (Human Error, Safety, and System Development)
Philippe Palanque, *France*
WG13.6 (Human-Work Interaction Design)
Torkil Clemmensen, *Denmark*
WG13.7 (Human Computer Interaction and Visualization)
Mark Apperley, *New Zealand*

INTERACT 2009 Technical Committee

Conference Committee

General Co-chairs

Jan Gulliksen	*Uppsala university and KTH Royal Institute of Technology in Stockholm, Sweden*
Lars Oestreicher	*Uppsala University, Sweden*

Technical Programme Co-chairs

Paula Kotzé	*Meraka Institute, South Africa*
Tom Gross	*Bauhaus University Weimar, Germany*

Technical Programme Committee

Demonstrations Co-chairs

Olav W. Bertelsen	*University of Aarhus, Denmark*
Janet C. Read	*University of Central Lancashire, UK*

Doctoral Consortium Co-chairs

Else Nygren	*Uppsala University, Sweden*
Kasper Hornbæk	*University of Copenhagen, Denmark*

Full Paper Co-chairs

Paula Kotzé	*Meraka Institute, South Africa*
Tom Gross	*Bauhaus University Weimar, Germany*

Keynote Speaker Chair

Mats Edenius	*Uppsala University, Sweden*

Panel Co-chairs

John Karat	*IBM Research, USA*
Masaaki Kurosu	*The Open University of Japan, Japan*

Short Papers and Interactive Posters Co-chairs

Philippe Palanque	*University Paul Sabatier, France*
Raquel Oliveira Prates	*Federal University of Minas Gerais, Brazil*

Special Interest Groups Co-chairs

Ute Klotz — *University of Applied Sciences and Arts of Lucerne, Switzerland*
Nikolaos Avouris — *University of Patras, Greece*

Tutorial Co-chairs

Fabio Paternò — *ISTI-CNR, Italy*
Mats Lind — *Uppasala University, Sweden*

Workshop Co-chairs

Anders Jansson — *Uppsala University, Sweden*
Gitte Lindegaard — *Carleton University, Canada*

Organization

Proceedings Publication Chair

Marco Winckler — *University Paul Sabatier, France*

Program Chair

Bengt Sandblad — *Uppsala University, Sweden*

Sponsoring Co-chairs

Tomas Berns — *Ergolab, Sweden*
Bengt Sandblad — *Uppsala University, Sweden*

Students Volunteers Co-chairs

Elina Eriksson — *Uppsala University, Sweden*
Rebecka Janols — *Uppsala University, Sweden*

Website Co-chairs

Lars Oestreicher — *Uppsala University, Sweden*
Nils-Erik Gustafsson — *Frontwalker/IT-arkitekterna, Sweden*

Programme Committee Members

Julio Abascal, *Spain*
Silvia Abrahão, *Spain*
Heikki Ailisto, *Finland*
Xavier Alamán, *Spain*
Mir Farooq Ali, *USA*
Dima Aliakseyeu, *The Netherlands*
Margarita Anastassova, *France*
Elisabeth Andre, *Denmark*
Paul André, *UK*
Daniela Andrei, *Romania*
Jean-Yves Antoine, *France*
Mark Apperley, *New Zealand*
Cecilia Aragon, *USA*
Liliana Liliana, *Italy*
Carmelo Ardito, *Italy*
Myriam Arrue Recondo, *Spain*
Richard Atterer, *Germany*
Simon Attfield, *UK*
Juan Carlos Augusto, *UK*
Nikolaos Avouris, *Greece*
Cedric Bach, *France*
Jonathan Back, *UK*
Nilufar Baghaei, *Australia*
Sebastiano Bagnara, *Italy*
Rafael Ballagas, *USA*
Cecilia Baranauskas, *Brazil*
Simone Barbosa, *Brazil*
Javier Bargas-Avila, *Switzerland*
Louise Barkhuus, *USA*
Len Bass, *USA*
Konrad Baumann, *Austria*
Gordon Baxter, *UK*
Russell Beale, *UK*
Elke Beck, *Austria*
David Benyon, *UK*
Francois Berard, *France*
Regina Bernhaupt, *Austria*
Karsten Berns, *Germany*
Nadia Berthouze, *UK*
Nigel Bevan, *UK*
Nicola J Bidwell, *Australia*
Jacob Biehl, *USA*
Thomas Binder, *Denmark*
Jeremy Birnholtz, *USA*
Staffan Björk, *Sweden*
Renaud Blanch, *France*
Ann Blandford, *UK*
Eli Blevis, *USA*
Pieter Blignaut, *South Africa*
Florian Block, *UK*
Susanne Bodker, *Denmark*
Mads Bodker, *Denmark*
Davide Bolchini, *USA*
Marcela Borge, *USA*
Ron Boring, *USA*
Giorgio Brajnik, *Italy*
Emily Brown, *USA*
Mike Brown, *Isle of Man*
Mihaela Brut, *Romania*
Martin Brynskov, *Denmark*
George Buchanan, *UK*
Steffen Budweg, *Germany*
Sabin-Corneliu Buraga, *Romania*
Jamika Burge, *USA*
Andreas Butz, *Germany*
Jacob Buur, *Denmark*
M. Claudia Buzzi, *Italy*
Marina Buzzi, *Italy*
Daragh Byrne, *Isle of Man*
Paul Cairns, *UK*
Eduardo H. Calvillo, *UK*
José Campos, *Portugal*
Pedro Campos, *Portugal*
Stefan Carmien, *Spain*
Luis Carriço, *Portugal*
John M. Carroll, *USA*
Tiziana Catarci, *Italy*
Augusto Celentano, *Italy*
Anxo Cereijo Roibás, *UK*
Teresa Chambel, *Portugal*
Olivier Chapuis, *France*
Kelvin Cheng, *Australia*
Keith Cheverst, *UK*
Luca Chittaro, *Italy*
Luigina Ciolfi, *Isle of Man*
Torkil Clemmensen, *Denmark*

Nathalie Colineau, *Australia*
Karin Coninx, *Belgium*
Stéphane Conversy, *France*
Michael Cooke, *Isle of Man*
Nuno Correia, *Portugal*
Maria Francesca Costabile, *Italy*
Adrien Coyette, *Belgium*
Stefan Cronholm, *Sweden*
João Cunha, *Portugal*
Mary Czerwinski, *USA*
Carla Maria Dal Sasso Freitas, *Brazil*
Peter Dalsgaard, *Denmark*
Catalina Danis, *USA*
Peter Dannenmann, *Germany*
Nigel Davies, *UK*
Janet Davis, *USA*
Antonella De Angeli, *UK*
Oscar de Bruijn, *UK*
Berardina De Carolis, *Italy*
Alexander De Luca, *Germany*
Maria De Marsico, *Italy*
Giorgio De Michelis, *Italy*
Boris de Ruyter, *The Netherlands*
Clarisse de Souza, *Brazil*
Ruth de Villiers, *South Africa*
Andy Dearden, *UK*
Yogesh Deshpande, *India*
Francoise Detienne, *France*
Nicoletta Di Blas, *Italy*
Lily Diaz, *Finland*
Anke Dittmar, *Germany*
Tanja Döring, *Germany*
Carlos Duarte, *Portugal*
Emannuel Dubois, *France*
Casey Dugan, *USA*
Sophie Dupuy-Chessa, *France*
Achim Ebert, *Germany*
Niklas Elmqvist, *USA*
David England, *UK*
Paul Engelfield, *UK*
Thomas Erickson, *USA*
Vanessa Evers, *The Netherlands Antilles*
Giorgio Faconti, *Italy*
Stefano Faralli, *Italy*
Carla Faria Leitão, *Brazil*
William Farr, *UK*
Loe Feijs, *The Netherlands*
Daniel Felix, *Switzerland*
Ylva Fernaeus, *Sweden*
Amyris Fernandez, *Brazil*
Danyel Fisher, *USA*
Morten Fjeld, *Sweden*
Daniela Fogli, *Italy*
Eelke Folmer, *USA*
Joan Fons, *Spain*
Manuel Joao Fonseca, *Portugal*
Amanda Fonville, *USA*
Peter Forbrig, *Germany*
David Fourney, *Canada*
Silvia Gabrielli, *Italy*
Luigi Gallo, *Italy*
Craig Ganoe, *USA*
Nestor Garay-Vitoria, *Spain*
Roberto García, *Spain*
Stavros Garzonis, *UK*
Franca Garzotto, *Italy*
Jean-Marc Gauthier, *Singapore*
Miguel Gea, *Spain*
Helene Gelderblom, *South Africa*
Cristina Gena, *Italy*
Gautam Ghosh, *Norway*
Carlo Giovannella, *Italy*
Audrey Girouard, *USA*
Kentaro Go, *Japan*
Paula A. Gomes da Silva, *Portugal*
Daniel Gonçalves, *Portugal*
Maria Paula Gonzalez, *Argentina*
Connor Graham, *Australia*
Nicholas Graham, *Australia*
Raphael Grasset, *New Zealand*
Marianne Graves Petersen, *Denmark*
Phil Gray, *UK*
Marde Greefff, *South Africa*
William Green, *The Netherlands*
Joshua Gross, *USA*
Tom Gross, *Germany*
John Grundy, *New Zealand*
Nuno Guimaraes, *Portugal*
Jan Gulliksen, *Sweden*

Adriana Mihaela Guran, *Romania*
Francisco Luis Gutierrez Vela, *Spain*
Jörg Haake, *Germany*
Michael Haller, *Austria*
Lars Hallnäs, *Sweden*
Kim Halskov, *Denmark*
Morten Borup Harning, *Denmark*
Richard Harper, *UK*
Michael Harrison, *UK*
Steve Harrison, *USA*
Kirstie Hawkey, *Canada*
Steve Haynes, *USA*
Matthias L. Hemmje, *Germany*
Michael Herczeg, *Germany*
Morten Hertzum, *Denmark*
Martin Hitz, *Austria*
Wolfgang Hochleitner, *Austria*
Paul Holleis, *Germany*
Lars Erik Holmquist, *Sweden*
Kristina Höök, *Sweden*
John Hosking, *New Zealand*
Aaron Houssian, *USA*
Nancy Houston, *USA*
Stefan Hrastinski, *Sweden*
William Hudson, *UK*
Jina Huh, *USA*
Stéphane Huot, *France*
Duke Hutchings, *USA*
Ebba Hvannberg, *Iceland*
Netta Iivari, *Finland*
Shamshi Iqbal, *USA*
Poika Isokoski, *Finland*
Minna Isomursu, *Finland*
Giulio Jacucci, *Finland*
Lars-Erik Janlert, *Sweden*
Eva Jettmar, *USA*
Sune Alstrup, *Denmark*
Chris Johnson, *UK*
Matt Jones, *UK*
Steve Jones, *New Zealand*
Joaquim Jorge, *Portugal*
Anirudha Joshi, *India*
Christophe Jouffrais, *France*
Oskar Juhlin, *Sweden*
Anne Kaikkonen, *Finland*
Matthew Kam, *USA*
Bridget Kane, *Isle of Man*
Anne Marie Kanstrup, *Denmark*
Maurits Kaptein, *The Netherlands*
Victor Kaptelinin, *Sweden*
John Karat, *USA*
Kostas Karpouzis, *Greece*
Dinesh S. Katre, *India*
Rick Kazman, *USA*
Christina Keller, *Sweden*
Babak Khazaei, *UK*
Olga Khroustaleva, *USA*
Stephen Kimani, *Australia*
Alex Klippel, *USA*
Anders Kluge, *Norway*
Michael Koch, *Germany*
Christina Köffel, *Austria*
Christophe Kolski, *France*
Andreas Komninos, *UK*
Joseph A. Konstan, *USA*
Ilpo Koskenin, *Finland*
Paula Kotzé, *South Africa*
Kostas Koukouletsos, *Greece*
Matthias Kranz, *Germany*
Subramanian Krishnamurthy, *India*
Steinar Kristoffersen, *Norway*
Antonio Krüger, *Germany*
Sari Kujala, *Finland*
Juliusz Kulikowski, *Poland*
Kari Kuutti, *Finland*
Jennifer Lai, *USA*
Michael Lankes, *Austria*
Ann Lantz, *Sweden*
Rosa Lanzilotti, *Italy*
Andreas C. Larsson, *Sweden*
Marta Larusdottir, *Iceland*
Effie Law, *UK*
Marco Lazzari, *Italy*
Eric Lecolinet, *France*
Minkyung Lee, *New Zealand*
Jair Leite, *Brazil*
Barbara Leporini, *Italy*
Stefano Levialdi, *Spain*
Henry Lieberman, *USA*
Ann Light, *UK*

Youn-Kyung Lim, *Republic of Korea*
Gitte Lindgaard, *Canada*
Silvia Lindtner, *USA*
Zhengjie Liu, *China*
Steffen Lohmann, *Germany*
Nicole Lompré, *France*
Vctor M. López Jaquero, *Spain*
Christopher Lueg, *Australia*
Artur Lugmayr, *Finland*
Paul Lukowicz, *Denmark*
Christof Lutteroth, *New Zealand*
Kris Luyten, *Belgium*
Saturnino Luz, *Isle of Man*
Baillie Lynne, *UK*
José Antonio Macías, *Spain*
Lachlan MacKinnon, *UK*
Allan MacLean, *UK*
Charlotte Magnusson, *Sweden*
Ana Gabriele Maguitman, *Argentina*
Lone Malmborg, *Denmark*
Jani Mäntyjärvi, *Finland*
Mari-Carmen Marcos, *Spain*
Annelise Mark Pejtersen, *Denmark*
Gary Marsden, *South Africa*
Jean-Bernard Martens, *The Netherlands*
Benoît Martin, *France*
Jean-Claude Martin, *France*
Masood Masoodian, *New Zealand*
Sara Mastro, *USA*
Vaclav Matousek, *Czech Republic*
Tara Matthews, *USA*
Emanuela Mazzone, *UK*
Graham McAllister, *UK*
Scott McCrickard, *USA*
Gregor McEwan, *Australia*
Tom McEwan, *UK*
Joseph McKnight, *Isle of Man*
Gerrit Meixner, *Germany*
Eduarda Mendes Rodrigues, *UK*
Alexander Meschtscherjakov, *Austria*
Jorn Messeter, *Sweden*
Thomas Mirlacher, *Austria*
Sonia Modeo, *Italy*
Tonja Molin-Juustila, *Finland*
Francisco Montero Simarro, *Spain*
Anders Mrch, *Norway*
Roberto Moriyon, *Spain*
Ann Morrison, *Finland*
Florian Mueller, *Australia*
Christian Mueller-Tomfelde, *Australia*
Jörg Müller, *Denmark*
Piero Mussio, *Italy*
Miguel A. Nacenta, *Canada*
Linda Napoletano, *Italy*
Domenico Natale, *Italy*
David Navarre, *France*
Karsten Nebe, *Germany*
Luciana Nedel, *Brazil*
Jeff Nichols, *USA*
Colette Nicolle, *UK*
Janni Nielsen, *Denmark*
Laurence Nigay, *France*
Erik G Nilsson, *Norway*
Sylvie Noël, *Canada*
Monique Noirhomme, *Belgium*
Mie Norgaard, *Denmark*
Jasminko Novak, *Switzerland*
Nuno Nunes, *Portugal*
Clemens N. Klokmose, *Denmark*
Horst Oberquelle, *Germany*
Marianna Obrist, *Austria*
Sergio Ochoa, *Chile*
Lars Oestreicher, *Sweden*
Michael OGrady, *Isle of Man*
Brian OKeefe, *UK*
Marta Oliva, *Spain*
Rikke Orngreen, *Denmark*
Alvaro Ortigosa, *Spain*
Dan Orwa Ochieng, *Kenya*
Lidia Oshlyansky, *UK*
Benoit Otjacques, *Luxembourg*
Antti Oulasvirta, *Finland*
Saila Ovaska, *Finland*
Philippe Palanque, *France*
Emanuele Panizzi, *Italy*
Luca Paolino, *Italy*
Eleftherios Papachristos, *Greece*
Tessy C. Pargman, *Sweden*

Cecile Paris, *Australia*
Timo Partala, *Finland*
Avi Parush, *Canada*
Oscar Pastor, *Spain*
Isabelle Pecci, *France*
Samuli Pekkola, *Finland*
Helen Petrie, *UK*
Emmanuel Pietriga, *France*
Marcelo Pimenta, *Brazil*
Ian Pitt, *Isle of Man*
Fabio Pittarello, *Italy*
Andreas Pleuss, *Germany*
Beryl Plimmer, *New Zealand*
Karolien Poels, *The Netherlands*
Klaus Pohl, *Germany*
Roberto Polillo, *Italy*
Vesna Popovic, *Australia*
Dorin Mircea Popovici, *Romania*
Chris Power, *UK*
Mathura Prasad Thapliyal, *India*
Marco Pretorius, *South Africa*
Costin Pribeanu, *Romania*
Angel Puerta, *USA*
Pernilla Qvarfordt, *USA*
Victor Ruiz Penichet, *Spain*
Kari-Jouko Raiha, *Finland*
Roope Raisamo, *Finland*
Nitendra Rajput, *USA*
Ismo Rakkolainen, *USA*
Roberto Ranon, *Italy*
Dimitrios Raptis, *Greece*
Matthias Rauterberg, *The Netherlands*
Madhu Reddy, *USA*
Johan Redström, *Sweden*
Karen Renaud, *UK*
Yann Riche, *USA*
Laurel D. Riek, *UK*
Thomas Rist, *Germany*
Chris Roast, *UK*
Markus Rohde, *Germany*
Teresa Roselli, *Italy*
Chiara Rossitto, *Sweden*
Ling Rothrock, *USA*
Amanda Rotondo, *USA*
Anne Roudaut, *France*
Jose Rouillard, *France*
Nicolas Roussel, *France*
Elisa Rubegni, *Switzerland*
Marco Sá, *Portugal*
Saqib Saeed, *Germany*
Antti Salovaara, *Finland*
Nithya Sambasivan, *USA*
Jaime Sanchez, *Chile*
Frode Eika Sandnes, *Norway*
Chris Sandor, *Australia*
Carmelina Santoro, *Italy*
Giuseppe Santucci, *Italy*
Corina Sas, *UK*
Anthony Savidis, *Greece*
Dominique Scapin, *France*
Chris Schmandt, *USA*
Albrecht Schmidt, *Germany*
Johannes Schöning, *Germany*
Gerhard Schwabe, *Switzerland*
Montserrat Sendin Veloso, *Spain*
Tacha Serif, *Turkey*
Cecília Sik Lányi, *Hungary*
Gavin Robert Sim, *UK*
Carla Simone, *Italy*
Mikael B. Skov, *Denmark*
Andy Smith, *UK*
Andrew Smith, *South Africa*
Marc Smith, *USA*
Eva Snee, *USA*
Hugo Solis, *USA*
Jan Stage, *Denmark*
Emil Stanescu, *Romania*
Christian Stary, *Austria*
Gunnar Stevens, *Germany*
Didier Stricker, *Denmark*
Sriram Subramanian, *UK*
Noi Sukaviriya, *USA*
S. Shyam Sundar, *USA*
Yngve E Sundblad, *Sweden*
John Tang, *USA*
Jean-Claude Tarby, *France*
Deborah Tatar, *USA*
Patrice Terrier, *France*
Nadia Thalmann, *Switzerland*
Jakob Tholander, *Sweden*

Michael Toomim, *USA*
Stefan Trausan-Matu, *Romania*
Daniela Trevisan, *Brazil*
Sanjay Tripathi, *India*
Manfred Tscheligi, *Austria*
Maurizio Tucci, *Italy*
Tuomo Tuikka, *Finland*
Phil Turner, *UK*
Susan Turner, *UK*
Jaime Urquiza-Fuentes, *Spain*
Kaisa Väänänen-Vainio-Mattila, *Finland*
Heli Väätäjä, *Finland*
Pedro Valderas, *Spain*
Judy van Biljon, *South Africa*
Elise van den Hoven, *The Netherlands*
Charles van der Mast, *The Netherlands*
Gerrit van der Veer, *The Netherlands*
Betsy van Dijk, *The Netherlands*
Darelle van Greunen, *South Africa*
Lex van Velsen, *The Netherlands*
Jean Vanderdonckt, *Belgium*
Alison Varey, *UK*
Radu-Daniel Vatavu, *Romania*
Angel Velázquez, *Spain*
Gina Venolia, *USA*
Markel Vigo, *Spain*
Giuliana Vitiello, *Italy*
Arnd Vitzthum, *Germany*
Jörg Voskamp, *Germany*
Ljiljana Vukelja, *Switzerland*
Dhaval Vyas, *The Netherlands*
Marion Walton, *South Africa*
Gerhard Weber, *Germany*
Carolyn Wei, *USA*
Alexandra Weilenmann, *Sweden*
Astrid Weiss, *Austria*
Janet Wesson, *South Africa*
Mikael Wiberg, *Sweden*
Daniel Wigdor, *USA*
David Wilfinger, *Austria*
Wolfgang Woerndl, *Germany*
William Wong, *UK*
Peta Wyeth, *Australia*
Heng Xu, *USA*
Diana Xu, *UK*
Pradeep G Yammiyavar, *India*
Alvin Yeo, *Malaysia*
Yeliz Yesilada, *UK*
Massimo Zancanaro, *Italy*
Luke Zhang, *USA*
Jürgen Ziegler, *Germany*
John Zimmerman, *USA*

Sponsors and Supporters

Gold

IT-arkitekterna

Swedish Foundation for Strategic Research

VINNOVA

Municipality of Uppsala

Silver

Microsoft Research

Bronze

IBM Research

ANTROP

Local Organizers

IFIP TC13

Uppsala University, Department of Information Technology

Swedish Interdisciplinary Organisation for Human-Computer Interaction (STIMDI)

Swedish Interdisciplinary Interest Group for Human–Computer Interaction

The Swedish Interdisciplinary Interest Group for Human–Computer Interaction (STIMDI) is a non-profit organization in Sweden that has the purpose of promoting cooperation between industry and academia within the field of HCI. STIMDI has been around since 1986 and also took the initiative of starting the NordiCHI conference series with NordiCHI 2000 in Stockholm. STIMDI is happy to be the co-organizer of INTERACT 2009 together with Uppsala University. More information can be found on www.stimdi.se

Uppsala University

Uppsala University is the oldest university in northern Europe. It was founded in 1477 and has about 40,000 students. Uppsala University is a complete university with all subjects that exist in its research area. It is the biggest center for HCI research in Sweden with three different departments of HCI at the Faculty of Science and Technology, at the Faculty of Social Science and at the Faculty of Medicine.

The research conducted by the HCI group at the department of Information Technology at Uppsala University concerns the analysis and design of computerized work environments. A goal is to ensure usability (satisfaction, efficiency and effectiveness) and reduce complexity for computer users. Most of the research is performed within governmental authorities, organizations and industrial settings, where the role is to guide the development of usable systems through user-centered design activities. A significant part of the research concerns complex computer systems intended for professional users in their everyday work. This also involves studying organizational aspects and work environment problems. Current research domains are:

- Human control of complex systems: high-speed ferry operation, train driver operation, train traffic control, simulators
- Administrative work environments: case-handling, call center operation
- Health care: decision support systems, tele-radiology, home health care
- Visualization: tactical control and command systems, virtual learning lab, 3D tools
- Work environment: Physical, psychosocial and cognitive work environment in computer supported work

Table of Contents – Part II

Multimodal Interfaces 1

Multimodal Interfaces 2

Multimodal Interfaces 3

Multi-user Interaction and Cooperation 1

Multi-user Interaction and Cooperation 2

Novel User Interfaces and Interaction Techniques 1

Novel User Interfaces and Interaction Techniques 2

Novel User Interfaces and Interaction Techniques 3

Novel User Interfaces and Interaction Techniques 4

Novel User Interfaces and Interaction Techniques 5

Novel User Interfaces and Interaction Techniques 6

Social Media/Social Networks

Tangible User Interfaces and Robotics

Tools for Design, Modelling and Evaluation 1

Tools for Design, Modelling and Evaluation 2

Usability Evaluation Methods

User Experience 1

User Experience 2

User Interfaces for Safety Critical Systems and Health Care 1

User Interfaces for Safety Critical Systems and Health Care 2

User Interfaces for Web Applications and E-commerce

Visualisation Techniques

Part Three: Demonstrations

Part Four: Doctoral Consortium

Part Five: Interactive Posters

Part Six: Panels

Part Seven: Special Interest Groups

Part Eight: Tutorials

Part Nine: Workshops

Table of Contents – Part I

Part One: Keynote Speakers

Part Two: Long and Short Papers

Accessibility 1

Accessibility 2

Affective HCI and Emotion

Child Computer Interfaces

Ethics and Privacy

Evaluation 1

Evaluation 2

Games, Fun and Aesthetic Design 1

Games, Fun and Aesthetic Design 2

HCI and Web Applications 1

HCI and Web Applications 2

Human Cognition and Mental Load 1

Human Cognition and Mental Load 2

Human Error and Safety

Human-Work Interaction Design

Interaction with Small and Large Displays 1

Interaction with Small and Large Displays 2

International and Cultural Aspects of HCI

Mobile Computing 1

Mobile Computing 2

Mobile Computing 3

Mobile Computing 4

Model-Based Design of Interactive Systems

Ambiguous Keyboards and Scanning: The Relevance of the Cell Selection Phase

Julio Miró-Borrás, Pablo Bernabeu-Soler, Raul Llinares, and Jorge Igual

Department of Communications
Universidad Politécnica de Valencia, Plaza Ferrandiz-Carbonell s/n,
03801 Alcoy, Spain
{jmirobo,pbernabe,rllinares,jigual}@dcom.upv.es

Abstract. This paper focuses on the relevance of the cell selection phase in the overall performance of a text entry system based on scanning and with an ambiguous keyboard. Most of the layouts are designed trying only to minimize the ambiguity of the keyboard, and taking into consideration only the disambiguation process when entering text. Nevertheless, the number of scan cycles necessary for selecting the cells has great importance in the overall performance. As we show, the performance depends on the number of cells and the linguistic model used in the cell selection phase.

Keywords: Text Entry, Scanning, Ambiguous Keyboards, Disambiguation.

1 Introduction

Keyboards with fewer keys[1] than characters (ambiguous keyboards) can be used instead of scan matrices in text entry systems based on scanning in order to increase the text entry rate. Nevertheless, most proposals try to create layouts that minimize the ambiguity of the keyboard based only on the disambiguation process. This is the standard way when considering physical keyboards such as the ones used in mobile phones or small devices. MacKenzie *et al.* [1] compared most of the ambiguous keyboards proposed by other researchers in terms of KSPC (keystrokes per character). Although some of these keyboards work with scanning, they were all treated the same way: direct access. An ambiguous keyboard with indirect access and character-level disambiguation is presented in [2], [3]. Lesher *et al.* [4] proposed several ambiguous keyboards that minimize the ambiguity using what is called the confusability matrix. All these proposals consider only the disambiguation process, without analyzing the effect of scanning the different cells on the overall performance.

This paper describes the research carried out with a single-switch text entry system using an ambiguous virtual keyboard and a word-level disambiguation process. Some of the current ambiguous layouts proposed by researchers are analyzed in our scanning system concluding that the best physical ambiguous layouts are not the best layouts when are used with scanning.

---

[1] We use cells instead of keys when talking about virtual keyboards accessed through scanning.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 1–4, 2009.

The remainder of this paper is organized as follows. In the next section we explain the methods employed in the research, specifically the mathematical model and the layouts to be analyzed. Next, the results are presented and analyzed and finally, some conclusions are drawn.

2 Methods

The research is done using the proposal in [3]. It consists of a scanning system with ambiguous keyboard and a word-level disambiguation method. Instead of using a separate key for space and next function, it uses a combination of automatic and inverse scanning, and short and long keypresses as follows. The operation of this proposal is composed of two phases. The first one is the cell selection phase. Using direct scanning, the user presses the switch when the desired cell is highlighted. This is done for all the characters in a word except the last one. For the last letter, the user presses and keeps the switch pressed, indicating to the system that the word is finished. Then, the disambiguation phase starts, presenting the suggested words one after another using inverse scanning. When the desired word is shown, the user releases the switch and the word followed by a space is entered into the system.

In this study, we used the same word-frequency list derived from the British National Corpus and described by Silverberg *et al.* [5], with the 9022 most common words in English. In the first phase, we obtained the probabilities of all characters for each position in a word, considering all the words in the list. Then, for each position, we obtained the key probabilities considering all letters in that key. In the second phase, we used the 9022 word list.

In a scanning environment, both the action of pressing the switch as well as the cursor movement are traditionally considered as keystrokes. A parameter that comprises all kinds of keystrokes is n_W or average weighted number of keystrokes per word (Equation 1), where $n_W(w)$ is the weighted number of keystrokes in word w, and $p(w)$ its probability. Additionally, w_S, w_C, and w_L are the weights for the scan cycles, the short and the long presses of the switch respectively, and n_S, n_C and n_L are the average values of the respective number of keystrokes.

$$n_W = \sum_w n_W(w) \cdot p(w) = w_S \cdot n_S + w_C \cdot n_C + w_L \cdot n_L . \tag{1}$$

When calculating the keystrokes values, n_L and n_S keep constant, so the best layouts are those that minimize the number of scan cycles. When considering physical layout performance, the main parameter is KSPC [1] or keystrokes per character (Equation 2), where K_W is the number of keystrokes needed for entering word w, F_W is the frequency of w, and C_W is the number of letters in w.

$$KSPC = \frac{\sum (K_w \times F_w)}{\sum (C_w \times F_w)} . \tag{2}$$

In our research we evaluate the layouts of other keyboards presenting the lowest value of KSPC for a fixed number of keys. These are: L2K, UKO-II, TMK4, L6K, Mobile Phone, and QP10213.

3 Results

We present in this section the number of scan cycles for the best physical layouts with 2, 3, 4, 6, 8 and 9 keys, i.e., those that present the smaller value of KSPC for the predictive method of disambiguation. The values for the number of short and long keystrokes for all layouts are: n_{C1}= 3.50, n_{L1}=1, n_{C2}=1 and n_{L2}=0 in both phases.

In Table 1 we present the number of scan cycles in phase 1 (n_{S1}), phase 2 (n_{S2}) and total (n_{ST}) for the known layouts. The last column shows the KSPC. In Table 2 we present the values obtained for other layouts with alphabetical arrangement of letters with better performance when used with our scanning system. The value of KSPC is also shown in the same table.

Table 1. Analysis of the best current keyboards

Layout	Keys	n_{S1}	n_{S2}	n_{ST}	KSPC
QP10213	9	10.82143	0.02337	10.84480	1.00431
Phone	8	10.41355	0.03479	10.44835	1.00641
L6K	6	6.11822	0.16165	6.27987	1.02979
TM4K	4	5.22093	0.27777	5.49869	1.05118
UKO II	3	3.74514	0.54090	4.28605	1.09967
L2K	2	2.06310	2.96924	5.03235	1.54712

4 Discussion

The value of KSPC for physical keyboards working in predictive disambiguation mode, decreases generally when increasing the number of keys. The reason is that the ambiguity decreases when less characters share the same cell. As we can see in Table 1, the best physical layout is QP10213 with a KSPC of 1.00431. Nevertheless the number of scan cycles is 10.84, much greater than UKO-II, which needs only 4.29 scan cycles with only 3 keys. KSPC for UKO-II is 1.09967, i.e., it requires 8.7% more keystrokes than QP10213 when considering the disambiguation phase, but saves 153.0% scan cycles.

The data in Table 1 shows that the keyboards with smaller KSPC are not the best keyboards when using scanning. In Table 2, we propose other keyboards with an alphabetical ordering of letters on the keys. They permit us to compare pairs of keyboards with the same number of keys. As we can see, all keyboards in Table 2 except AK2 present larger values of KSPC than their equivalents in Table 1. And what is more interesting, all keyboards present a smaller number of scan cycles. The highest differences are found in the 8-key devices: the AK8 needs 5.1% more keystrokes than the Phone keyboard, but it saves 51.8% scan cycles. Moreover, AK3 needs 15.1% more keystrokes than UKO-II, and saves 15.1% scan cycles. Notice that L2K presents a smaller number of scan cycles than AK2 but with a greater KSPC. In general, we can state that with higher number of keys, the savings in keystrokes are smaller than with keyboards with 2 or 3 keys, but the savings in scan cycles are just the opposite. This is because with more cells, the average number of scan cycles for selecting them increases, while the ambiguity is smaller.

Table 2. Proposal of several alternative keyboards (AK) with a smaller number of scan cycles

Layout	Keys	Layout	n_{S1}	n_{S2}	n_{ST}	KSPC
AK9	9	a-i, jk, lm, no, pq, rst, uv, wx, yz	5.35468	0.28484	5.63952	1.05249
AK8	8	a-i, jk, lm, no, pq, r-v, wx, yz	4.70206	0.32906	5.03112	1.06063
AK6	6	a-i, j-l, m-o, p-t, u-w, x-z	4.46755	0.34948	4.81703	1.06440
AK4	4	a-j, k-n, o-v, w-z	3.37419	0.63304	4.00723	1.11665
AK3	3	a-g, h-t, u-z	2.16769	1.59956	3.76724	1.29474
AK2	2	a-l, m-z	2.13221	2.93368	5.06589	1.54057

Traditional studies with ambiguous keyboards to be used with scanning try to get the best layouts for the disambiguation phase, and many times with a large number of keys [5]. Nevertheless, as we have shown here, it is more efficient to consider the cell selection phase in the design of the layout. The relevance of the cell selection phase in the total number of scan cycles depends on the linguistic model being used. The results presented in this work are obtained using a 1-gram model. With a fixed scanning pattern, the differences would be even greater, and with more elaborated prediction systems, the differences would be smaller. Only in the case of perfect prediction in phase 1, will the layouts with smaller KSPC have the best performance when using scanning.

Acknowledgments. This study has been partly funded by Cátedra Telefónica - UPV.

References

1. MacKenzie, I.S., Tanaka-Ishii, K.: Text entry using a small number of buttons. In: Text Entry Systems: Mobility, Accessibility, Universality, pp. 105–121. Morgan Kaufmann, San Francisco (2007)
2. Miró, J., Bernabeu, P.A.: Text entry system based on a minimal scan matrix for severely physically handicapped people. In: Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I. (eds.) ICCHP 2008. LNCS, vol. 5105, pp. 1216–1219. Springer, Heidelberg (2008)
3. Miró-Borrás, J., Bernabeu-Soler, P.: Text Entry in the E-Commerce Age: Two Proposals for the Severely Handicapped. J. theor. appl. electron. commer. res. 4, 101–112 (2009)
4. Lesher, G.W., Moulton, B.J., Higginbotham, D.J.: Optimal Character Arrangements for Ambiguous Keyboards. IEEE Trans. Rehab. Eng. 6, 415–423 (1998)
5. Silfverberg, M., MacKenzie, I.S., Korhonen, P.: Predicting text entry speed on mobile phones. In: SIGCHI Conference on Human Factors in Computing Systems, pp. 9–16 (2000)

Force Feedback Magnitude Effects on User's Performance during Target Acquisition: A Pilot Study

Lode Vanacken, Joan De Boeck, and Karin Coninx

Hasselt University – tUL – IBBT, Expertise Centre for Digital Media (EDM)
Wetenschapspark 2 B-3590 Diepenbeek (Belgium)
{lode.vanacken,joan.deboeck,karin.coninx}@uhasselt.be

Abstract. Only a few guidelines exist for defining the force properties in a haptic interface; as a consequence, they are mostly determined in an ad-hoc manner. We investigate how the user's performance changes during target acquisition when increasing force amplitudes are applied. Using a simple multidirectional point-select task, forces with variable amplitudes are applied to the user while traversing from one target to the other. We find that the user's performance suddenly degrades significantly, rather than decreasing progressively. This finding may be important for defining guidelines which forces may and may not be applied to a user in order to allow the haptic feedback to improve, rather than deteriorate the user's performance.

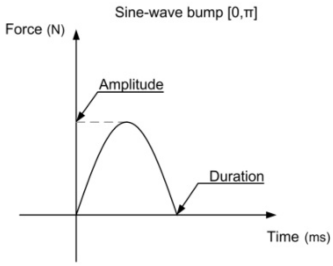
Keywords: Force Feedback, Target Acquisition.

1 Introduction and Related Work

Force feedback can be applied to support pointing tasks in a desktop application or in a virtual environment. This can be achieved either by assisting the user e.g. using 'gravity wells' (a 'snap-to' effect that pulls the user's pointer to the centre of the target) [1] or by giving haptic feedback in the form of a bump or a vibration as soon as an event occurs or a given zone is entered [2]. As not many guidelines exist, finding suitable values for the different parameters that define the forces is mostly performed using a 'trial-and-error' approach.

In our research, we want to answer the fundamental question what kind of forces may or may not be applied in order to support the user. Ultimately, this approach may lead to a set of guidelines avoiding the aforementioned trial-and-error approach. In this paper we want to investigate the user's behaviour with respect to the task completion time with changing forces. However, 'changing' a force may influence a number of degrees of freedom: force feedback device, force shape, duration, amplitude, force direction, …). Within the scope of this paper we define the force as a short force in a given direction, with given duration and amplitude. The amplitude of the force over time may follow a mathematical pattern such as a sine or a step function, which we define as the force shape. Fig. 1a illustrates such a sinusoidal 'force bump' which may occur lateral or longitudinal with respect to the user's movement.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 5–8, 2009.



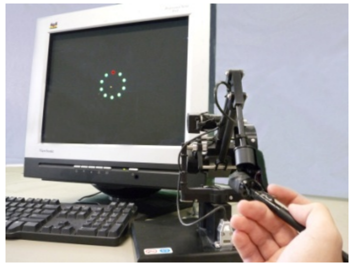


Fig. 1. a. (left) Force evolution of a sinusoidal haptic bump. b. (right) Setup of the experiment.

2 Experiment

We conducted a multi-directional tapping experiment in which we evaluate the influence of the force magnitude on the performance of the user.

2.1 Setup

As can be seen from Fig. 1b, our experimental setup consisted of a regular 19-inch CRT monitor and a Phantom premium 1.0 haptic device. The control display gain was 1 (one physical cm corresponds to one cm on the screen). Ten participants (one female and nine males) with an age between 22 and 30 (average 25), recruited among our co-workers, participated. All participants were right-handed and used their dominant hand during the experiment.

A simple multidirectional point-select task, as described in ISO 9241-9 [3], was used for this experiment. Ten targets are placed in a circle on the screen (see Fig. 1b). The diameter of the circle is determined at 6 cm and the size of a target is 0.7 cm. This task has a Fitts' index of difficulty of 3.26 bits, a measure typically used in Fitts' law experiments to indicate the difficulty of the task [4]. The value is chosen to be comparable to the task difficulty of a typical icon selection task [5].

During the test, the ten targets were highlighted one after the other and users were requested to select (by pointing and clicking) the highlighted target as fast and accurate as possible. Highlighting is altered between opposite sides of the circle so that user movements are equally distributed among all directions with a maximum distance between the targets. As the task to perform was a 2D selection task and the haptic device we used is a 3D input device, a vertical guiding plane restricted the task to two dimensions.

Finally, force feedback appearing in the form of a force bump with given shape, duration and varying amplitude was activated when exactly half-way in the path to the next target.

2.2 Design

A repeated measures within-participant design was adopted. The independent variables were force direction *D* (lateral to movement direction d_{lat}, longitudinal in the direction of the movement d^{+}_{long} and longitudinal opposed to the movement d^{-}_{long}) and force amplitude *A* (see Fig. 2). A fully crossed design resulted in 33 combinations of

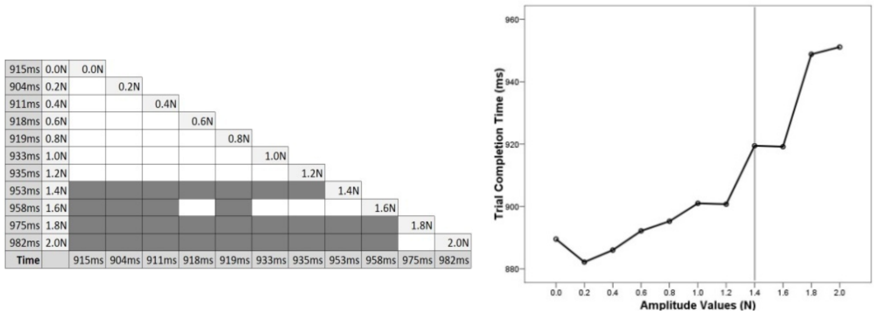


Fig. 2. a. (left) Pair wise post hoc comparisons of the amplitude conditions for trial completion time, all significant differences ($p < .05$) are shaded. b. (right) Trial completion times for the different amplitudes.

A and *D*. Each participant performed the experiment in one session lasting about 20 minutes. The session consisted of 5 blocks with each block containing the 33 combinations (11 *A*s and 3 *D*s) repeated 3 times in a random order. For a total of 99 trials per block, this resulted in 495 trials per participant. Between each block, users were obliged to take a break of at least 15 seconds to minimise fatigue during the test. Before the experiment, participants were given all 33 conditions in random order to familiarise them with the task. During the experiment the time it took to select a target was recorded, as well as the amount of errors made during selection.

2.3 Results and Discussion

We want to eliminate the results of any learning effects in our analysis; therefore we first investigated the influence of the block on the trial completion time. From this analysis, blocks 1 and 2 were significantly slower and were hence removed ($F_{4, 36} = 5.3$, $p < .003$).

Trial Completion Time. A repeated measures analysis of variance of the faultless trials, showed main effects for *D* ($F_{2, 18} = 5.42$, $p < .02$) and *A* ($F_{10, 90} = 12.6$, $p < .0001$). The average trial completion times for direction are 901.5ms for d^{+}_{long}, 916.2ms d^{-}_{long} and 905.5ms for d_{lat}. The differences between the *D*-conditions are small with respect to practical use. Post hoc comparison showed no significant differences except between d^{+}_{long} and d^{-}_{long}. This may not be surprising, as the first is assisting the user's movement while the latter is opposing it.

Post hoc comparisons for *A* showed many significant results. Fig. 2a represents these differences ($p < .05$) darkly shaded. Comparing the *no-force* condition with all other *A*-s (first column), a significant trial completion time deterioration, all $p < .002$, can be seen from *A*=1.4N. Fig. 2a shows that this deterioration is also confirmed by the comparison of the other *A* conditions. Fig. 2b also depicts all the average completion times for all *A* values: we see a small but non-significant deterioration for the smaller amplitudes and a stronger, significant degradation starting from 1.4N. This cut-off behaviour may be interesting for further research, when trying to formulate guidelines which forces may and may not support the user's task.

Error Rates. The same pattern can be found for the error-rate. An error was recorded when the user clicked the button without having the correct target underneath the cursor. The overall error rate for the experiment was 114 errors or 2.3%. The direction had no significant effect on error rate *D* ($F_{2, 18} = .668$, $p = .525$). However, the amplitude values did show a significant effect *A* ($F_{10, 90} = 2.62$, $p < .01$). Post hoc comparisons showed that the *no-force* condition (0.7%) has a significant difference with the last two *A* conditions (1.8N and 2.0N, respectively 5.2% and 5.6%). In the scope of this finding, we might refer to the remarks of several users reporting involuntary miss-clicks due to the large forces, and due to the forces diverting them from an anticipated target click.

3 Conclusion and Future Work

We investigated the user's performance when forces with different amplitudes are applied during a targeting task. We observed performance deterioration as forces become stronger. It is important to notice that the degradation occurs according to a cut-off behaviour, rather than decreasing progressively. This pilot study only took the force amplitude and direction into account. However, several other parameters such as duration and shape can be changed as well. From the physical relation between force, acceleration and velocity, we foresee that the calculation of the integral may be a good prediction of the user's behaviour, which we will investigate in our future work.

References

1. Hwang, F., Langdon, P., Keates, S., Clarkson, J.: The effect of multiple haptic distractors on the performance of motion-impaired users. In: Eurohaptics 2003, pp. 14–25 (2003)
2. Forlines, C., Balakrishnan, R.: Evaluating tactile feedback and direct vs. indirect stylus input in pointing and crossing selection tasks. In: CHI 2008, pp. 1563–1572 (2008)
3. ISO, International Organisation for Standardisation. ISO/TC 159/SC4/WG3 N147. Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9 - Requirements for nonkeyboard input devices, May 25 (1998)
4. Fitts, P.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47, 381–391 (1954)
5. Soukoreff, R.W., MacKenzie, I.S.: Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in HCI. Int. J. Hum.-Comput. Stud. 61(6), 751–789 (2004)

Gaze-Assisted Pointing for Wall-Sized Displays

Hans-Joachim Bieg[1], Lewis L. Chuang[2], and Harald Reiterer[1]

[1] University of Konstanz, HCI Group,
Box D 73, Universitätsstraße 10, 78457 Konstanz, Germany
[2] Max-Planck Institute for Biological Cybernetics,
Box 2169, Spemannstraße 38, 72012 Tübingen, Germany
{hans-joachim.bieg,harald.reiterer}@uni-konstanz.de,
Lewis.Chuang@tuebingen.mpg.de

Abstract. Previous studies have argued for the use of gaze-assisted pointing techniques (MAGIC) in improving human-computer interaction. Here, we present experimental findings that were drawn from human performance of two tasks on a wall-sized display. Our results show that a crude adoption of MAGIC across a range of complex tasks does not increase pointing performance. More importantly, a detailed analysis of user behavior revealed several issues that were previously ignored (such as, interference of corrective saccades, increased decision time due to variability of precision, errors due to eye-hand asynchrony, and interference with search behavior) which should influence the development of gaze-assisted technology.

Keywords: Eye-Tracking, Eye-Hand Coordination, Multimodal.

1 Introduction

Advances in display technology and the user's need for more screen real-estate are likely to lead to considerable display size increases in the coming years. Already, displays typically extend beyond the size of 20", up to that of large walls. With this comes a host of fresh challenges – such as the exploitation of large screen real-estate [1] and complex display configurations [2] – that could be addressed by utilizing knowledge of a user's current point of gaze on the screen.

A technique for gaze-assisted pointing was previously presented by Zhai, Morimoto, and Ihde [3] that was based on the simple assumption that users tend to look at display items that were potential targets for selection. Using information from a gaze-tracking system, they were able to translocate the pointer to the user's current point of fixation. In other words, the pointer followed the user's gaze and reduced the movement that was typically required of the user. To enable uninterrupted manipulations, the pointer was only translocated if a fixation occurred with considerable (120 px) distance to the current pointer position and if the device was not in motion at that time. In addition, the users were still able to use their pointing device to override this modification. Zhai and colleagues showed that gaze-assisted pointing techniques resulted in (marginally) faster response times on a simple pointing task.

We expected gaze-assisted pointing techniques to be especially useful on wall-sized displays, given the large movement amplitudes that are required on such

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 9–12, 2009.

interfaces, Furthermore, we decided to test these techniques on two tasks that reflected the complexities of a user's typical experience with display interfaces. The goals of this study were to evaluate whether the MAGIC technique would lead to substantial improvements of pointing speed and to see how the technique was affected by perceptual demand.

2 Method

The experiment was carried out by 12 participants with normal or corrected-to-normal vision. In this study, participants were equipped with a mouse and were asked to perform different search and drag-and-drop tasks of varying complexity (see Fig. 1). On half of these trials, participants were provided with gaze-assisted pointing (MAGIC) and this modification was disabled on the remaining half. This order was counterbalanced to control for practice effects.

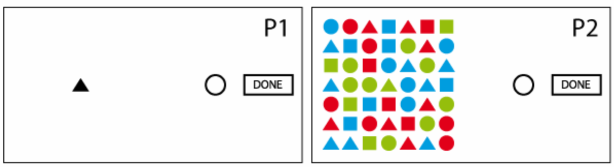


Fig. 1. Schematic representation of two experimental tasks. Left: in the first task, a single target had to be selected and dragged to a circular homing area, followed by a click on a button to confirm task completion. Right: in the second task participants had to visually search a 13 × 13 item grid for a single item, specified by shape (triangle, rectangle, circle) and color (red, green, blue) before performing the same drag-and-drop task. Such a combination of search and selection is frequently required during a range of human-computer interaction tasks [4]. The size of each item was approximately 20 px (3.3 cm, 1.92° visual angle).

Participants sat 100 cm in front of a back-projected display (resolution 1400 × 1050 px, size 94.1° × 89.1° visual angle). Subjects wore a head-mounted eye-tracker, with attached markers for head-tracking. These tracking data could thus be combined to compute the point of the user's gaze, on the screen [5]. This information was provided to the system in real-time and at 120 Hz. Experimental sessions lasted around 90 minutes and the system was recalibrated regularly to assure that gaze was tracked accurately.

3 Results and Discussion

Comparison of task completion time and selection times showed better task completion times for trials without MAGIC, relative to trials that enabled MAGIC for task p1 (mean difference: 1094 ms, $t(11)=6.6$, $p < 0.01$) and task p2 (mean difference: 1720 ms, $t(11)=6.7$, $p < 0.01$). This was the case despite latencies of approximately one second between fixations on the target and selections on the MAGIC trials. Although users fixated the target early, triggering a translocation of the pointer in the

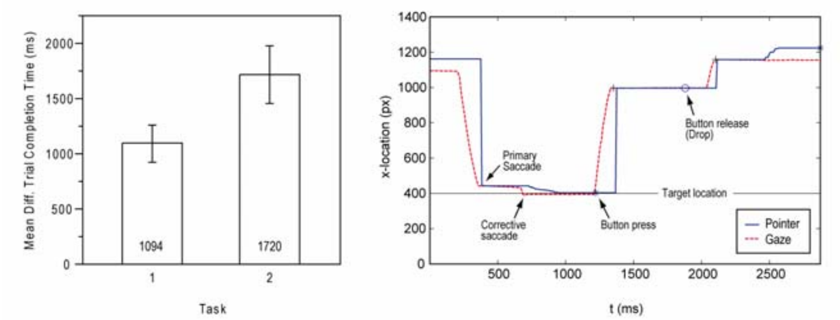


Fig. 2. Left: difference in mean trial completion time (ms) between trials with MAGIC enabled and disabled (whiskers show std error). Right: movement profile of a trial during task 1. The pointer is translocated to the first inaccurate fixation (primary fixation). This fixation is followed by a corrective saccade which is below the 120 px threshold.

MAGIC condition, they were not able to benefit from this translocation and required equal or even more time to carry out fine-adjustments for the selection. A detailed analysis of user behavior with MAGIC revealed several reasons for this and suggests potential improvements:

Large saccades tend to generate considerable error in their landing positions, relative to the intended targets. These are typically followed by small corrective saccades (see Fig. 2 right) which fall below our adopted threshold of 120 px [6]. In turn, these could have led to conflicts between the users' expectations and the displayed position of the pointer. A solution to this problem would be to make the threshold dependent on the amplitude of the preceding saccade or disabling it for a fixed time span after a saccade. Based on our current data, we believe that a better threshold would be one that adapts to the saccade amplitude. A separate investigation of the effect of saccade length on fixation errors could provide more information about the benefits of such an adaptation and whether the function of this adaptation could be calibrated to fit the individual's actual performance.

Precision variability could also have increased selection time by requiring an additional cognitive task of evaluating if the pointer was displayed "accurately". If the pointer was translocated precisely, the user was able to directly select the item underneath it. If it was off, a corrective action was required. This problem could be remedied by not translocating the pointer directly to a fixation but by introducing an artificial offset so that a more predictable correction can occur during the final manual acquisition movement.

Errors also occurred when gaze moved away from targeted locations before the mouse actions were completed. For instance, participants sometimes started searching for the next target before they had dropped the item. This idiosyncratic tendency could be calibrated on an individual level by modifying the latency of pointer translocation to suit the given individual's behavioral patterns.

A comparison of the second task against the first task showed an increase in task completion times because users required time to search for the target [7]. However, the increment was larger when MAGIC was enabled (see Fig. 2 left, $t(11) = 3.5$, $p < 0.01$). This result gives rise to the assumption that participants were distracted by translocations of the pointer while searching for the target. Comparison of fixation duration showed that fixations during search were much shorter (mean = 281 ms, SD = 108 ms) than fixations directly preceding target selections (mean = 988 ms, SD = 455 ms). An adaptive mechanism that prevents translocation during series of short fixations could increase the performance of gaze-assisted pointing by reducing distractions due to unintended pointer motion during search.

4 Conclusion

The applicability of MAGIC, a gaze-assisted pointing technique was tested with a large display. Contrary to our initial expectations, this technique did not induce any movement time savings during basic selection and search tasks. However, the experiment also revealed several issues that negatively affected performance. Specifically, we discuss how gaze-assisted pointing can benefit by being adapted to basic coordination patterns and individual differences in eye-movement behavior. The appropriate parameters for these adaptations can be approximated from a larger dataset.

References

1. Swaminathan, K., Sato, S.: Interaction design for large displays. Interactions 4(1), 15–24 (1997)
2. MacIntyre, B., Mynatt, E.D., Voida, S., Hansen, K.M., Tullio, J., Corso, G.M.: Support for multitasking and background awareness using interactive peripheral displays. In: Proceedings of the annual ACM symposium on User interface software and technology (UIST), pp. 41–50 (2001)
3. Zhai, S., Morimoto, C.H., Ihde, S.: Manual and gaze input cascaded (MAGIC) pointing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 246–253 (1999)
4. Fleetwood, M.D., Byrne, M.D.: Modeling the visual search of displays: a revised ACT-R model of icon search based on eye-tracking data. Human-Computer Interaction 21(2), 153–197 (2008)
5. Herholz, S., Chuang, L.L., Tanner, T.G., Bülthoff, H.H., Fleming, R.W.: LibGaze Real-time gaze-tracking of freely moving observers for wall-sized displays. In: Proceedings of the Vision, Modeling, and Visualization Workshop, VMV (2008)
6. Harris, C.M., Wolpert, D.M.: The Main Sequence of Saccades Optimizes Speed-accuracy Trade-off. Biological Cybernetics 95(1), 21–29 (2006)
7. Wolfe, J.M.: Visual Search. In: Pashler, E. (ed.) Attention, University College London Press, London (1998)

Hand Pointing Accuracy for Vision-Based Interactive Systems

Kelvin Cheng[1,2] and Masahiro Takatsuka[2]

[1] HxI Initiative, NICTA, Australian Technology Park,
13 Garden St, Eveleigh, NSW 2015, Australia
[2] ViSLAB, School of IT, The University of Sydney, NSW 2006 Australia
kelvin.cheng@nicta.com.au, masa@vislab.usyd.edu.au

Abstract. Vision-based hand pointing interactive systems always assume implicitly that users' physical pointing accuracy is perfect. However, this may not be the case. We investigated the accuracy provided by users in three pointing strategies. Result showed that pointing inaccuracy can be as high as 239mm at 3 metres away and suggest that the line-up method provides the best accuracy overall.

Keywords: Hand Pointing, Pointing Accuracy, Computer Vision.

1 Introduction

In most vision-based interactive systems the accuracy of the estimated pointing direction is an essential element to their success. The focus is usually on finding new ways of improving the detection, estimation and tracking the pointing hand or finger, and deduces the pointing direction [1, 2, 4]. However, these systems do not account for inaccuracies made by the user and assumed implicitly that the user is always pointing to the desired location. In this paper, we report on a study that was conducted to investigate the inherent accuracy provided by three common pointing strategies, which were widely used in previous interactive systems:

1) Forearm – only the forearm is used to point, with limited upper arm movement.
2) Straight-arm – the whole arm is used to point, without bending.
3) Line-up – the fingertip is positioned collinearly between the eye and the target. Several names have been used in the literatures for this method such as "eye-fingertip"[3] or "head-hand line"[2], we call this the "line-up" method.

Despite advances in computer vision, there is still no consistent method to segment and extract the pointing direction of the arm. To minimize errors introduced by computer vision systems, we made use of the laser pointing. A laser pointer can be attached to the arm and gives us a direct and simple resultant point, making it possible to quantitatively compare the different targeting strategies. However, it would be difficult to fix a laser pointer at a specific position on the arm. Consequently subjects were asked to hold the laser pointer in their hand in a way that best represent the extension of their arm direction, in a consistent manner across all three methods (Figure 1). Although physically pointing with hand or arm compared to holding the laser in the palm are slightly different, we believe that this difference would be consistent

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 13–16, 2009.

enough so that it would still be comparable relatively between the strategies. Therefore until a more suitable method is found, we felt that it was appropriate in this case to use the laser pointer. In addition, users were not provided feedback from the laser pointer to adjust their accuracy.

The effect of distance on the accuracy of pointing – whether pointing deteriorates as user moves away from the target – was also investigated in this experiment.

2 Experiment

Fifteen volunteers (3 female, 12 male) participated in this study. Their age ranged from 20 to 31 (average of 24.9). One is left handed while 8 are right eye dominant. A 5x5mm black target was constructed on a wall at a height of 1.5 metres. The laser pointer used was a pen (1x14.7cm) with a red laser pointer fitted at the back end. In order to detect the target and the red laser dot, we used a webcam (Logitech Quickcam Pro 4000) positioned 2 metres away from the target, together with a custom image processing software (Visual C++ and OpenCV). The lighting in the room was dimmed to allow improved detection of the laser dot.

The study was a within-subject study where each subject performed pointing tasks with all three pointing styles from three distances: 1, 2 and 3 metres from the target. Three blocks of trails were completed for each of the 9 combinations and the mean position for each combination was determined. The order of pointing styles and distances were counter-balanced. Without turning on the laser, subjects were asked to aim as accurately as possible, and hold the laser pointer in their dominant hand, in a way that best represents the direction of their pointing arm (for straight arm and forearm pointing, illustrated in Fig 1a). For the line-up method, users were asked to place the laser pointer so that both ends of the laser pointer are collinear with the target and the eye (Fig 1b). Subjects were free to use one or both eyes for this method.

To prevent subjects receiving feedback from the laser dot, the laser was only turned on after they have taken aim. Subjects were asked to keep the orientation of the pen consistent throughout the entire experiment, in order to minimize unwanted deviation from the button press. A snapshot was then taken by the webcam. Accuracy was measured in terms of the distance between the target and the laser dot produced on the wall. The positions of the laser dots were determined off-line. In summary, the experimental design was: 15 subjects x 3 pointing techniques x 3 distances x 3 blocks = 405 pointing trials.

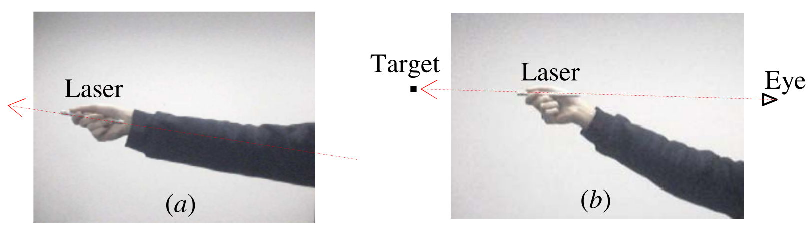


Fig. 1. (*a*) A laser pointer was used to represent the arm's direction (*b*) The laser pointer used with the line-up method, where the target, two ends of the laser pointer and eye are collinear

3 Results and Discussion

Figure 2 illustrates the mean distance from target for each pointing method at the three distances, and their interactions, for all trials.

A two-way analysis of variance (ANOVA) with repeated measures reveals a significant main effect for pointing technique on accuracy ($F[2,28]=37.97$, $p<0.001$), as well as for distance on accuracy ($F[2,28]=47.20$, $p<0.001$). We also observed a significant interaction between technique and distance ($F[4,56]=9.879$, $p<0.001$).

Multiple pairwise means comparisons were tested within each pointing technique. The p-values resulted with each comparison between the distances are shown in table 1 with Bonferroni correction. Trend analyses were also performed on each technique. Significance in the linear component for the particular technique signifies a linear increase in accuracy with increasing distance.

Multiple pairwise means comparisons were also performed within each distance to investigate possible differences between each technique, and the p-values are shown in the table 2.

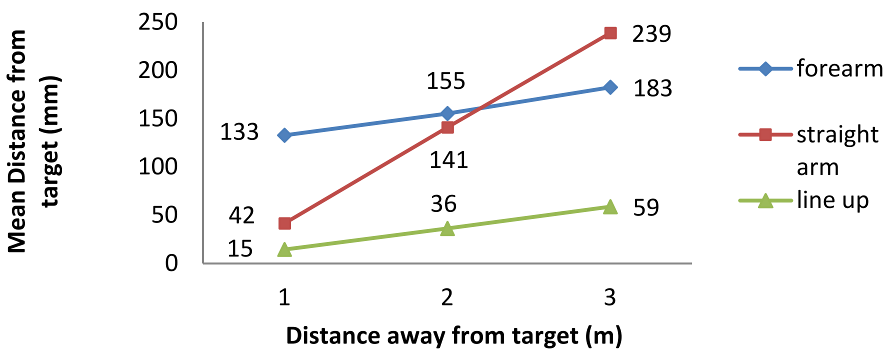


Fig. 2. Mean distance between target and the laser dot position

Table 1. The significance of multiple pairwise means comparisons within each pointing technique

Technique	1m vs 2m	2m vs 3m	1m vs 3m	Linear component
Forearm	1.000	0.378	0.055	0.018*
Straight arm	<0.001*	0.002*	<0.001*	<0.001*
Line up	0.003*	0.116	0.014*	0.005*

*Denotes significance at the 0.05 level

Table 2. The significance of multiple pairwise means comparisons within each distance

Distance	Forearm vs straight arm	Straight arm vs line up	Forearm vs line up
1m	0.001*	<0.001*	<0.001*
2m	1.000	<0.001*	<0.001*
3m	0.182	<0.001*	<0.001*

*Denotes significance at the 0.05 level

Results suggest that the line-up method is the most accurate across all distances. A linear increase throughout the 3 distances at a rate of 14.7mm per meter can also be observed. The forearm pointing technique is consistently less accurate than the line-up method. Even though the linear increase for this technique (16.7mm per metre) is similar to the line-up method, the difference in accuracy between the two methods is at least 115mm at all three distances. It is interesting to note the insignificance between the three distances within the forearm pointing method. This may suggest a high tolerance with increasing distance from target for this pointing method.

On the other hand, the straight arm pointing method is highly affected by the increase in distance from target. This is illustrated by the significant difference across all distances and the relatively high linear increase (65.7mm per meter). Compared to forearm pointing, the accuracy at 2m and 3m are not significant. While the only difference between forearm and straight arm pointing is at 1m. This may be due to the higher level of feedback given by the longer arm extension, and that the straight arm pointing resembles the line-up method at close proximity to the target.

4 Conclusion

From this experiment, we have identified inaccuracies in users pointing performance, which varies depending on the strategy used. We observed that the line-up method is the most accurate pointing method, and that the straight arm method is more accurate than the forearm method only at a distance of one metre. These results provide us a basic understanding on the accuracy of each method. The question of whether these methods are sufficiently accurate depends largely on the specific task, and the decision rest on the individual interaction designer.

To complement this study, a follow-on experiment may be conducted to study the ergonomics of these pointing methods. Understanding the trade off between accuracy and ease-of-use can assist future vision-based hand pointing interaction researchers and practitioners to decide the input strategy that best suit their users in completing the required tasks.

References

1. Cipolla, R., Hollinghurst, N.J.: A Human-Robot Interface using Pointing with Uncalibrated Stereo Vision. In: Cipolla, R., Pentland, A. (eds.) Computer Vision for Human-Machine Interaction, pp. 97–110. Cambridge University Press, Cambridge (1998)
2. Colombo, C., Bimbo, A.D., Valli, A.: Visual Capture and Understanding of Hand Pointing Actions in a 3-D Environment. IEEE Transactions on Systems, Man, and Cybernetics 33, 677–686 (2003)
3. Hung, Y.-P., Yang, Y.-S., Chen, Y.-S., Hsieh, I.-B., Fuh, C.-S.: Free-Hand Pointer by Use of an Active Stereo Vision System. In: Proceedings of 14th International Conference on Pattern Recognition (ICPR), Brisbane, pp. 1244–1246 (1998)
4. Nickel, K., Stiefelhagen, R.: Pointing Gesture Recognition based on 3D-Tracking of Face, Hands and Head-Orientation. In: Proceedings of ICMI 2003 International Conference on Multimodal Interfaces, Vancouver, pp. 140–146 (2003)

Pen-Based Video Annotations: A Proposal and a Prototype for Tablet PCs

Diogo Cabral and Nuno Correia

CITI, Departamento de Informática, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, Quinta da Torre, 2829-516 Caparica, Portugal
{diogocabral,nmc}@di.fct.unl.pt

Abstract. Pen computing or pen-based technology can be considered one of the most natural and intuitive computer input interface for humans. The combination of digital video and pen-based technology can support a new approach for *electronic paper* by associating handwritten notes and drawings with video segments. This paper presents a proposal for pen-based video annotations and a prototype for Tablet PCs, implementing this concept, particularly dynamic pen-based video annotations combined with real-time video tracking.

Keywords: Annotations, Video Annotations, Pen-based Annotations, Pen-based video Annotations, Tablet PCs.

1 Introduction

The pen is an old tool, which resulted from the need to communicate and memorize by writing down information [1]. The pen has changed through time as well as the substrate used to write [1]. The first attempts in pen computing occurred in the 1950s and 1960s [1] and the most recent and widely used development of this technology is the Tablet PC, a laptop computer equipped with pen-based technology and a touch screen. The size of the displays, almost the same as a regular sheet of paper, and its computational power, as a regular laptop, make the Tablet PCs an interesting approach to *electronic paper* and a precursor to interactions that will be possible in the future [2].

Annotations are a useful form of writing down information using pen and paper and can be associated to other media types like text and images, in a process that combines reading and critical thinking usually called active reading [3]. Nevertheless, annotations can also be applied to digital platforms and documents [3, 4]. More recently, digital video became a popular type of digital document and a medium of communication worldwide. Video can be considered one of the most complete and complex medium. Besides the pictorial and verbal (written text or speech) attributes, video includes motion and time characteristics. A process analogous to the active reading, called active watching, can associate annotations to video content [5].

The combination of digital video and pen-based technology allows association between handwritten annotations/drawings with video segments. Pen-based video annotations introduce new notions for the note taking act, such as temporal position and motion, when compared with regular pen and paper annotations.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 17–20, 2009.

2 Related Work

The topic of pen-based video annotations is not new and some systems [6,7,8,9,10] already explored this idea. Nevertheless, most of them [6,8,9,10] do not totally explore the dynamic dimension of video content, making the annotations spatially static, while the content is moving. Although, the Ambulant Annotator [7] explores this dimension, the annotation spatial path is defined by the user and not by tracking the video content. A more recent work [11] explores graphical video annotations combined with motion tracking, but do not consider the particular case of pen-based annotations and the tracking feature needs a long time of video pre-processing.

3 Pen-Based Video Annotations

The input modality of digital annotations can be divided in three types: ink, keystroke and multimedia [12]. Pen-based annotations generate ink-based annotations; nevertheless, if a keyboard annotation is a set of characters [12], then a pen-based annotation can be considered a set of ink strokes. Each ink stroke has its own attributes, like color or thickness, which can differ from one to another [3]. Therefore, a pen-based annotation can be considered a set of ink strokes with different attributes. In addition, a temporal dimension can be added to the ink strokes. Normally, a user makes these ink strokes sequentially [4]. This temporal order associated to each ink stroke can be crucial to understand the idea transmitted by the annotation.

Video annotations should include spatial and temporal dimensions also associated with video content [5]. Together, these dimensions generate the idea of motion, spatial position that varies with time, which should be also associated to video annotations. Video annotations should have the ability to follow specific motion features included in the video content.

The combination of both concepts, pen-based annotations and video annotations, generate a new kind of annotations, pen-based video annotations, which can be defined as annotations composed by a set of time dependent ink strokes, each with specific attributes, and associated with video changeable temporal and spatial positions.

The attributes of pen-based video annotations can include ink properties (e.g. color, thickness), private or public definitions, hypertext properties (e.g. spatial anchors on the video content and links to other media types) and tracking features (e.g. tracking objects, colors and textures).

4 Prototype

A prototype was developed with the aim to implement pen-based video annotations. The system was developed for Tablet PCs (Fig. 1a) and includes pen-based annotations displayed over video content, associated by time to video intervals, and with some changeable ink attributes, such as color and thickness. In addition, the system has a frame based timeline making possible to browse the video content, by selecting

a) b)

Fig. 1. a) The prototype running on a Tablet PC b) Annotation Motion Tracking

a particular frame, and to draw over a set of frames. A scrollbar was also developed, with the aim to navigate through the frame timeline, with the particular feature that the marker has fixed dimensions, sufficient to be selected using the pen, without missing the target, as can happen with regular scrollbars.

The prototype includes annotations motion tracking, which allows annotations to follow motion changes in the video content, as shown in Fig. 1b.

The motion tracking is based on the difference of frames on the area defined by the annotation, in real-time. The annotation area, defined by a rectangle embracing the ink strokes, is equally divided in four smaller rectangles and the rectangle (of these four) that presents a larger change of pixels, represents the direction that the annotation should follow. This method does not require a lot of processing power, enabling its real-time usage. The annotations are drawn for every frame in a transparent layer over the video window, without the need to change the original video content. The system also redraws the ink strokes in the same order that they were made. The annotation lifetime is by default equal to the video time length but the time interval can be changed using a dual slider. Annotation motion stops when the object stops or when it is not visible in the video window.

5 Conclusions and Future Work

This paper presents a proposal for pen-based video annotations and the different notions introduced by this type of annotations. The dynamic property of pen-based annotations was implemented in a prototype using Tablet PCs, allowing an annotation to track video content movements, among other features. Additional capabilities, such as pressure and gesture widgets and tracking other features and multiple objects will be developed as next steps. Finally, usability tests will be carried out to test the proposed system.

Acknowledgments. The presented work is partially funded by the UTAustin-Portugal Program and by the HP Technology for Teaching Grant Initiative 2006.

References

1. Meyer, A.: Pen Computing, A Technology Overview and a Vision. SIGCHI Bulletin 27, 46–90 (1995)
2. Hayes, R.A., Feenstra, B.J.: Video-speed electronic paper based on electrowetting. Nature 425, 383–385 (2003)
3. Marshall, C.C.: Annotation: from paper books to the digital library. In: 2nd ACM International Conference on Digital libraries, pp. 131–140. ACM Press, New York (1997)
4. Bargeron, D., Moscovich, T.: Reflowing digital ink annotations. In: SIGCHI Conference on Human factors in Computing Systems CHI 2003, pp. 385–393. ACM Press, New York (2003)
5. Correia, N., Chambel, T.: Active video watching using annotation. In: 7th ACM International Conference on Multimedia (Part 2), pp. 151–154. ACM Press, New York (1999)
6. Ramos, G., Balakrishnan, R.: Fluid interaction techniques for the control and annotation of digital video. In: 16th annual ACM Symposium on User interface Software and Technology, pp. 105–114. ACM Press, New York (2003)
7. Bulterman, D.C.A.: Animating Peer-Level Annotations Within Web-Based Multimedia. In: Eurographics Multimedia 2004, pp. 49–57. Eurographics Association (2004)
8. Diakopoulos, N., Essa, I.A.: Videotater: an approach for pen-based digital video segmentation and tagging. In: 19th annual ACM symposium on User interface software and technology, pp. 221–224. ACM Press, New York (2006)
9. Rahn, N.C., Lim, Y., Groth, D.P.: Redesigning video analysis: an interactive ink annotation tool. In: CHI 2008 Extended Abstracts on Human Factors in Computing Systems, pp. 3339–3344. ACM Press, New York (2008)
10. Graça, C., Pimentel, M., Goularte, R., Cattelan, R.G., Santos, F.S., Teixeira, C.: Ubiquitous Interactive Video Editing Via Multimodal Annotations. In: Tscheligi, M., Obrist, M., Lugmayr, A. (eds.) EuroITV 2008. LNCS, vol. 5066, pp. 72–81. Springer, Heidelberg (2008)
11. Goldman, D.B., Gonterman, C., Curless, B., Salesin, D., Seitz, S.M.: Video object annotation, navigation, and composition. In: 21th annual ACM Symposium on User Interface Software and Technology, pp. 3–12. ACM Press, New York (2008)
12. Cousins, S.B., Baldonado, M., Paepcke, A.: A Systems View of Annotations. Technical Report, Xerox PARC (2000)

Human Perception of Near-Duplicate Videos

Rodrigo de Oliveira, Mauro Cherubini, and Nuria Oliver

Telefonica Research, Via Augusta, 177
08021 Barcelona, Spain
{oliveira,mauro,nuriao}@tid.es

Abstract. Popular content in video sharing websites (*e.g.,* YouTube) contains many duplicates. Most scholars define near-duplicate video clips (NDVC) as identical videos with variations on non-semantic features (*e.g.,* image/audio quality), while a few others also include semantic features (different videos of similar content). However, it is unclear what exact features contribute to human perception of similar videos. In this paper, we present the results of a user study conducted with 217 users of video sharing websites. Findings confirm the relevance of both classes of features, but the exact role played by semantics on each instance of NDVC is still an open question. In most cases, participants had a preference for one video when compared to its NDVC and they were more tolerant to changes in the audio than in the video channel.

Keywords: NDVC, near-duplicate, similarity, user study, YouTube.

1 Introduction

In the last few years, different research groups have tried to understand how video-sharing web sites are used and in particular the impact that near-duplicate video clips (NDVC) have on video information retrieval tasks [4], spam creation [1] and identification of copyright infringements [2]. Most of the previous work has focused on identifying and removing NDVC. However, there is no agreement on the technical definition of what features identify almost identical copies of the same video. According to Wu and colleagues [4], near duplicate videos differ in file format or encoding parameters, photometric variations (*e.g.*, color) and editing operations (*e.g.*, overlay, captions, add/remove scenes). Other scholars have employed an extended definition of *similarity between features* including changes on capturing time [2], for instance a different camera viewpoint. Finally, Basharat *et al.* included *similarity at the semantic level* [1], where the same semantic concept can occur under different scene settings (*e.g.*, two videos of different deer in different forests grazing moss).

We believe these studies are important but could benefit from additional information gathered via user studies for at least three reasons: 1) little is known about how users are affected by the presence of NDVC; 2) we have limited knowledge of whether the definition of near duplicate videos we choose to adopt matches the users' understanding of what NDVC are; and 3) we need empirical proofs that removing NDVC from the results set of a video search task would satisfy the users' needs. Therefore, we state the hypotheses of our study as:

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 21–24, 2009.

H1 Users of video sharing websites search for videos more than browse.
H2 Users of video sharing websites perceive NDVC according to their definition [1, 3, 4].
H3 Users of video sharing websites have preferences over near-duplicate videos and usually don't want to have all of them listed and displayed after executing a search query.

To test these hypotheses, we deployed a large-scale qualitative questionnaire where respondents characterized their common use of video sharing websites, watched pairs of NDVC and stated their similarity degree (pairs differing by only one feature), and presented their preferences (if any) about which duplicate they would like to have in the search results. These measurements led us to a user-centered definition of NDVC.

2 Experiment Design

Procedure. A questionnaire was deployed on a popular news portal in Spain (www.terra.es) to test our hypotheses. In terms of H1, we investigated the users' behavior in a video search task from two perspectives: *purpose* and *proactivity*. With respect to *purpose*, subjects were asked if they usually use services like YouTube to: (1) search for specific videos, (2) browse without a specific video in mind, or (3) do something else. In terms of *proactivity*, participants answered if the videos they watch on these systems are typically: (1) found by themselves, (2) suggested by someone else, or (3) found by other means. All subjective answers were manually categorized.

In order to validate H2, we looked for the most viewed videos in YouTube from "last month" and "at all times", and created queries to retrieve these videos. From the results set, five NDVC pairs were identified to exemplify variations of non-semantic features [3, 4], and two pairs to exemplify variations of semantic features [1]. Videos were edited such that all pairs would have the same length ($\bar{x} = 37$ seconds), except in one condition (see Table 1). Video examples were presented on a Latin square basis to avoid bias, thus creating seven groups. Each participant was submitted randomly to only one group. For each of the seven pairs, participants were asked to fully watch both videos and rate how similar they thought these videos were.

H3 was addressed by asking participants if they had a preference between the videos belonging to the same condition and which one would it be if they were searching for videos using the same query (see Table 1 for conditions and queries).

Table 1. Descriptions of the NDVC pairs used in the study (http://tinyurl.com/youtubestudy)

Condition	Query	Video 1	Video 2
A: Photometric variation	crazy frog champions	**A1:** standard image	**A2:** higher quality (color and lighting)
B: Editing operation (add/remove scenes)	skate Rodney Mullen	**B1:** fewer scenes, more content per scene	**B2:** more scenes, fewer content per scene
C: Different length	how to search in Google Maps	**C1:** first 38 seconds of video C2	**C2:** C1 with 24 seconds of extra content
D: Editing operation (audio/image overlays)	plane airport Bilbao wind	**D1:** no overlays	**D2:** overlays (audio comments and logo)
E: Audio quality	More than Words	**E1:** stereo, 44Khz	**E2:** mono, 11Khz
F: Similar images and different audio	atmospheric pressure	**F1:** experiment with a soda can	**F2:** experiment with a beer can
G: Similar audio and different images	Beatles all you need is love	**G1:** original musical clip	**G2:** G1 song performed by another band

Participants. From an initial pool of 647 participants who self-volunteered to answer the questionnaire, 217 (115 male, 122 female) complied with all the study requirements: (1) fluent in Spanish, (2) experience with at least one video sharing website, (3) could listen to the videos using the computer speakers or headphones, and (4) no relevant audio or visual impairment. Their median age was 31 years (min: 16, max: 63), 98.2% were Spanish, and had a wide range of occupations. Subjects reported using computers everyday and video sharing websites from 4 to 6 days/week.

3 Results and Discussion

Validation of H1: *Search was confirmed to be the users' purpose when using video sharing websites.* Sixty-one percent (133 subjects) of participants declared using these systems to search for a *specific* video while 39% (84 subjects) use them to spend time without anything in mind. Moreover, participants were *active users of these systems:* the majority of participants declared watching videos found by themselves (53%).

Validation of H2: Table 2 summarizes the results obtained in terms of H2, where the figure in bold reflects the highest value for each video pair.

Table 2. Similarity levels attributed to each NDVC pair (see Table 1)

Similarity level (five point Likert scale)	Video examples (% of subjects)						
	A	B	C	D	E	F	G
1. Completely different	3.2	8.8	5.1	6.0	5.1	2.8	30.0
2. Essentially different	11.1	14.7	12.9	15.2	9.7	10.6	18.4
3. Related somehow	7.4	**33.2**	**34.6**	23.0	8.3	34.1	**41.9**
4. Essentially the same	**42.9**	**35.0**	**35.0**	**43.3**	31.3	**45.6**	9.7
5. Exactly the same	35.5	8.3	12.4	12.4	**45.6**	6.9	0.0

According to Table 2, identical videos with different image quality (condition A) were perceived as near-duplicates (42.9% stated that videos A1 and A2 are "essentially the same"). Interestingly, when identical videos differed only in audio quality (condition E), they were considered as "exactly the same" (45.6%). One could argue that differences in audio quality could be perceived easier with headphones than with speakers, which suggests that the participants' different audio sets affected decisions (speakers: 184; headphones: 33). However, this was not the case (p=0.11). Therefore, we conclude that in the context of video sharing websites, *users are more tolerant to changes in the audio than in the video modality.* Regarding the validation of H2, given that NDVC from conditions B (add/remove scenes), C (different lengths) and D (overlays) were also mostly rated as "essentially the same", we could corroborate that the human perception of NDVC matches the commonly adopted definitions associated to non-semantic features [3, 4]. However, it is important to note that participants were undecided whether videos from conditions B and C could also be considered as "related somehow" (33.2% *vs.* 35% and 34.6% *vs.* 35% respectively). We are investigating this tie effect on a further quantitative study. With respect to semantics [1], most subjects perceived videos in condition F as "essentially the same" and in condition G as "related somehow". Therefore, we conclude that the *human perception of NDVC also has a semantic component*. However, it is not clear from our study the exact role that semantics play on particular instances of videos.

Validation of H3. Our findings confirm that *given 2 NDVC, users usually prefer one of them in a video search task,* being it the one with best image quality (condition A) or additional information (conditions C and D). In the case of videos sharing audio semantics, they opted for the original musical clip (condition G). Subjects preferred both videos when they shared most scenes with additional information on each (condition B), or were semantically similar, but visually different (condition F). When videos differed in audio quality, subjects either had no preference or preferred the one with the best quality (condition E). Table 3 summarizes these results.

Table 3. Preferences over near-duplicates for each NDVC pair (see Table 1)

Preference (single choice)	Video examples (% of subjects)						
	A	B	C	D	E	F	G
Only video 1	1.8	6.0	5.1	6.0	**35.0**	6.0	**54.4**
Only video 2	**52.5**	14.7	**61.3**	**46.5**	3.2	13.4	6.5
Both videos 1 and 2	18.0	**53.5**	19.4	27.2	24.4	**44.7**	36.4
None of the videos	0.5	4.1	0.5	1.4	1.8	2.3	0.9
No preference	26.3	19.8	13.4	18.4	**35.0**	33.6	1.4
Didn't understand query	0.9	1.8	0.5	0.5	0.5	0.0	0.5

Implications. Near-duplicate detection algorithms could achieve better results by also comparing semantics between video clips (*e.g.*, similar scenes, same people or objects, *etc.*). Furthermore, near-duplicates in search results should be treated according to what feature(s) make clips alike, *e.g.,* when a NDVC has additional information (condition C), its relevance in the search results should be increased.

4 Conclusions

From our results, human perception of NDVC matches many of the features present in its technical definitions with respect to manipulations of non-semantic features [2,4]. However, it is yet not clear whether similar clips differing in overlaid or added visual content with additional information can be considered as near-duplicates. Furthermore, the definition should be extended to videos with similar semantics but different visual and audio information [1]. However, there is still the need to identify low-level features that influence semantic similarity. Results of a follow-up questionnaire are being analyzed to clarify these findings. We plan to extend research on the feature set and include psychophysical experiments of feature interaction.

References

1. Basharat, A., Zhai, Y., Shan, M.: Content based video matching using spatiotemporal volumes. Journal of Computer Vision and Image Understanding 110(3), 360–377 (2008)
2. Benevenuto, F., Duarte, F., Rodrigues, T., Almeida, V.A., Almeida, J.M., Ross, K.W.: Understanding video interactions in YouTube. In: MM 2008, pp. 761–764. ACM Press, New York (2008)
3. Shen, H.T., Zhou, X., Huang, Z., Shao, J., Zhou, X.: UQLIPS: a real-time near-duplicate video clip detection system. In: VLDB 2007, VLDB Endowment, pp. 1374–1377 (2007)
4. Wu, X., Hauptmann, A.G., Ngo, C.-W.: Practical elimination of near-duplicates from web video search. In: MULTIMEDIA 2007, pp. 218–227. ACM Press, New York (2007)

PressureMove: Pressure Input with Mouse Movement

Kang Shi[1], Sriram Subramanian[2], and Pourang Irani[1]

[1] Computer Science Department, University of Manitoba
Winnipeg, R3T 2N2, Canada
{kangshi,irani}@cs.umanitoba.ca
[2] Computer Science Department, University of Bristol
Bristol, BS8 1UB, UK
sriram@cs.bris.ac.uk

Abstract. We present *PressureMove* a pressure based interaction technique that enables simultaneous control of pressure input and mouse movement. Simultaneous control of pressure and mouse movement can support tasks that require control of multiple parameters, like rotation and translation of an object, or pan-and-zoom. We implemented four variations of PressureMove techniques for a 2D position and orientation matching task where pressure manipulations mapped to object orientation and mouse movement to object translation. The *Naive* technique mapped raw pressure-sensor values to the object rotation; the *Rate-based* technique mapped discrete pressure values to speed of rotation and *Hierarchical* and *Hybrid* techniques that use a two-step approach to control orientation using pressure. In user study that compared the four techniques with the default mouse-only technique we found that Rate-based PressureMove was the fastest technique with the least number of crossings and as preferred as the default mouse in terms of user-preference. We discuss the implications of our user study and present several design guidelines.

Keywords: Pressure-input, integrality of input dimensions, pressure and movement alternative interaction techniques.

1 Introduction

In line with recent incarnations of the mouse, Cechanowicz et al [2] have augmented the mouse with additional pressure input channels, and called this augmentation the PressureMouse. The PressureMouse builds upon the recently published set of guidelines for pressure based interaction [10,14,16]. However, recent studies on pressure interactions primarily provide insight on the strengths and limitations of pressure-based input and offer guidelines for creating pressure augmented interactions. Very little is known on how to fluidly integrate pressure input channels with the basic operations of the input device to which it is being augmented.

Pressure based interaction techniques proposed for the mouse are largely based on users manipulating the pressure channel independently of the movement degrees-of-freedom [10,14,16]. Pressure augmentation could potentially be designed such that the user can manipulate pressure input and cursor movement, enabling users to synchronously perform actions that can otherwise only be accomplished sequentially. For

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 25–39, 2009.

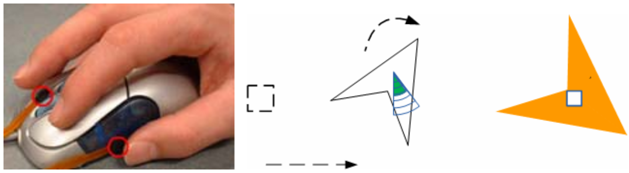


Fig. 1. (a) Mouse with two pressure sensors; (b) Rotating an object with pressure input and displacing it using mouse movement to achieve a common task in several applications. (see http://www.youtube.com/watch?v=YqyGaOSZhKY for a video)

example, a mouse could potentially enable users to rotate and translate an object synchronously, a task that is routinely carried out in drawing applications (Figure 1).

Based on results of an early pilot study and prior work (Zliding[14] and PressureMarks[15]), we observed that users can simultaneously control pressure and movement, but not all users utilize the simultaneous control in a fluid fashion. In this paper we investigate the design space and the resulting interaction techniques that allow simultaneous control of pressure and movement, referred to as *PressureMove*. To demonstrate the effectiveness of PressureMove, we concentrated on the task of simultaneous rotation and object translation. We designed four PressureMove techniques that provide users the flexibility of using the input dimensions of pressure and movement simultaneously or sequentially. Pressure manipulations controlled object orientation and mouse movement controlled movement. In a 2D rotate and translate task, similar to the tetrahedral docking task in 3D [8,21], we examined the proposed designs for integrating mouse movement and pressure rotation. Our results show that one of our PressureMove designs, the rate-based integration offered best control and performance and was significantly faster than all other techniques including the traditional mouse.

The main contributions of this paper are to: 1) extend the design space of a pressure augmented device (the mouse) to include simultaneous control of pressure and movement; 2) design integral interaction techniques; 3) identify strengths of various strategies for controlling non-competing degrees-of-freedom; and 4) outline design implications that emerge from our systems.

2 Related Work

We review the related research on pressure input and integral input channels.

2.1 Pressure Based Interaction

Ramos et al. [16] explored the design space of pressure-based interaction with styluses. They proposed a set of pressure widgets that operate based on the users' ability to effectively control a discrete set of pressure values. Ramos et al. [16] identified that adequate control of pressure values is tightly coupled to a fixed number of discrete pressure levels (six maximum levels), the type of selection mechanism and a high degree of visual feedback. However, their investigation does not explore the benefits of simultaneously integrating pressure control with stylus movement.

Mizobuchi et al. [10] conducted a study to investigate how accurately people control pressure exerted on a pen-based device. Their results show that continuous visual feedback is better than discrete visual feedback, users can better control forces that are smaller than 3N, and 5 to 7 levels of pressure are appropriate for accurate discrimination and control of input values. Since controlling pressure input is challenging, Shi et al [17] recently proposed PressureFish, a technique to discretize the pressure space using fisheye functions. With PressureFish, users are capable of manipulating pressure input with a higher level of control and more efficiently than common discretization functions.

Researchers studied pressure input in the context of multi-level interaction. Zeleznik et al. [19] proposed an additional “pop-through” state to the mechanical operation of the mouse button. Forlines et al. [3] proposed an intermediary “glimpse” state to facilitate various editing tasks. Multi-level input can facilitate navigation, editing or selection tasks but utilize pressure input in a limited way. Such techniques make it challenging to fluidly control another input channel such as mouse movement.

Cechanowicz et al [2] investigated the possibility of facilitating pressure-based input by augmenting a mouse with either one or two pressure sensors. Such an augmentation allows users to control a large number of input modes with minimal displacements of the mouse. Cechanowicz et al [2] developed several pressure mode selection mechanisms and showed that with two pressure sensors users can control over 64 discrete pressure modes. However, Cechanowicz et al [2] did not investigate the possibility of fluidly integrating pressure input with other mouse based operations.

Few results suggest how we can fully integrate pressure with the underlying input mechanisms of the device to which it is augmented. Ramos et al [14] proposed Zliding to control a scaling factor with pressure at the stylus’ tip and manipulating a parameter with the stylus’ x-y position. Similarly, with PressureMarks [15] the user can invoke several states by steering the stylus and simultaneously applying various degrees of pressure. While both these studies highlight the possibility of integrating pressure input with the movement of the device, they have not explored the large design space that results when integrating both input channels.

In general, very few of the reported results have explored the design space of fluidly integrating pressure input with the functional features of the device. Furthermore, little is known about how pressure integrates with the very common task of moving a pointer. Based on this limited knowledge it is challenging to propose applications that can benefit from integrating pressure with multiple input channels.

2.2 Fluidly Controlling Multiple Input Channels

There has been a long standing interest in identifying how to integrate and facilitate control of simultaneous input channels. Jacob et al [13] proposed a framework that can facilitate the understanding and categorization of integrality and separability of input devices and interactions afforded by these. Two input dimensions are considered integral if they are perceived as a single dimension or seperable if the dimensions seem unrelated [13]. In their study, performance was better when the device matched the tasks in integrality/separability dimensions. In light of their findings, coordinating multiple channels may suggest whether the input device is operating in the same dimension space as the task, i.e. good coordination and performance suggests that the device and perceptual structure of the task are in the same space. Integrality can be considered to some extent as a coordination measure.

Balakrishnan et al [1] used integrality to demonstrate that subjects could control three degrees of freedom simultaneously with the Rockin'Mouse, a X-Y translational and one Z-rotational DOF. Similarly, MacKenzie et al. [7] investigated the possibility of integrating rotation on the mouse, a device designed primarily for translation and selecting objects. The TwoBall mouse facilitates a number of common tasks, and makes certain application features, such as the rotate tool, redundant.

Studies have also investigated the benefits and possibility of integrating several tasks into one coherent and fluid action. Kruger et al. [5] designed a technique, RNT (Rotate'N Translate), to fluidly integrate rotation and translation. The motivation behind RNT was to provide in one seamless action the ability to rotate and translate an object in a collaborative environment. Results of their study show that RNT is more efficient than separately controlling translation and rotation. RNT further enhances a number of collaborative tasks, including coordination and communication with respect to user orientation.

Fluid integration of multiple input channels was examined in the context controlling an input device with the fingers instead of using the entire arm. In an empirical study, Zhai et al [21] investigated the effectiveness of finger muscle groups in controlling multiple degrees-of-input. Zhai et al [21] gave users two alternative 6DOF input devices, one that controlled a cursor with the movement of the entire arm (glove) and the other with the fingers of a hand (FingerBall [21]). The objective of the study was to compare finger control to arm control in finely rotating and positioning an object in 3D. The task consisted of docking a cursor with the target, both of which were equal size tetrahedral. They found that the finger-based device afforded simultaneous translation and rotation actions with better control.

In developing a metric for measuring the allocation of control in a 6 degree-of-freedom rotation and translation task, Masliah and Milgram [8] studied the interdependence and overlapping actions of the two tasks. They used a 3D virtual docking task, similar to that of Zhai [20] in which subjects were asked to align a tetrahedral shaped cursor onto an identically shaped target. Interestingly, their results showed that users would rarely control all 6 DOFs simultaneously. Instead, users would allocate their control to the rotational and translational DOFs separately. Wang et al [18] carried out a study to investigate the relationship between object transportation and object orientation by the human hand. In their experiment, subjects were asked to align a small wooden block with a graphical target cube. Manipulation tasks were designed that required both object translation and orientation. Their results demonstrate the existence of a parallel and independent structure for object translation and orientation. Their results suggest that object translation and orientation share characteristics of an integral structure according to the notion by Jacob et al [13].

3 PressureMove

We propose PressureMove, a pressure based technique that facilitates simultaneous control of mouse movement and pressure input. We considered two dimensions: controlling pressure input, and visual feedback.

3.1 Controlling Pressure

Sensors typically report pressure values between 0 to 1024 levels. Previous studies have suggested that users are not capable of distinguishing the granularity and controlling this range of pressure values [2, 10, 16]. This has led most investigations to discretizing the pressure space into controllable and haptically perceivable units. Ramos et al. [14, 16] revealed that adequate control of pressure values is tightly coupled to a fixed number of discrete pressure levels (maximum of six levels). Cechanowicz et al. [2] suggested that pressure discretization can include 8 to 10 discrete levels, when controlled by the thumb or index finger, on a mouse.

An alternative to discretizing pressure input is to map the raw pressure space (non-discretized referring to the fact that the discrete pressure values reported by the sensor are not further discretized) onto the task parameters. Each unit of pressure in the raw pressure space controls an input parameter, whether it be angular rotation, scalar, or other factor. Raw pressure input is not easily controlled, however facilitates a larger number of mappings.

We can also define a hybrid pressure space that uses continuous and discrete pressure values. With hybrid control, continuous pressure input provides the user with rapid access to a region of interest within the pressure space while switching to discrete control allows finer granularity and control over parameter values. PressureMove includes discrete, raw, and hybrid pressure control techniques.

3.2 Visual Feedback

Kinesthetic feedback alone is insufficient for adequately controlling pressure. Visual feedback is a dominant characteristic of most closed-loop pressure based interactions [2,10,14,16]. Different forms of Visual feedback for pressure based input have been explored in PressureWidgets [16]. However, the Visual feedback in PressureMove is inspired by the visual feedback mechanism used by Kittenakare et al [4] and Ramos et al [16]. Since the design of the visual feedback is intricately tied to the task, we describe the feedback designed for the task of simultaneously positioning and orienting an object. We expect that a similar form of visual feedback can be easily adapted for other simultaneous control tasks.

A pressure cursor is used to provide appropriate visual feedback. The default cursor is a solid triangular shaped object (see Figure 2(a)). When the user applies pressure a proportion of this cursor gets highlighted relative to the amount of pressure

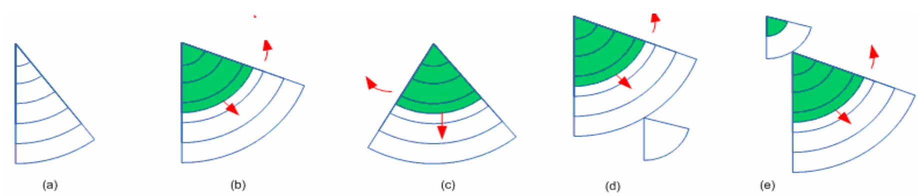


Fig. 2. Cursor state in the PressureMove techniques; a) standard PressureMove cursor without pressure; b) cursor fills up when pressure is being applied, movement is clockwise for one sensor and counter-clockwise for the other; c) in a hierarchical manner, for first pressure level, The red arrows are not part of the cursor and only used to explain how the cursor moves

being applied as in Figure 2(b) and 2(c). Visual feedback is always continuous, as this form of feedback has shown to enhance performance over non-continuous visual feedback. Additionally, we redundantly encode pressure amount to the aperture of the pressure cursor, i.e. the higher the pressure value, the larger the aperture of the cursor (as is seen in the difference in size of the cursor in Figure 2(a) and 2(b)).

In the case where we used a hybrid pressure space we used a two-step cursor as shown in Figure 2(d) and 2(e). The head-triangle (the triangle that represents the head of the cursor) represents the first pressure space the user can use while the second triangle corresponds to the second pressure space. In Figure 2(d) the user is currently controlling the first pressure space while in Figure 2(e) the user is operating with the second pressure space. In cases where multiple pressure spaces are composed to form the technique, multiple triangles can be concatenated. However, in our design we only used up to two pressure spaces composed to form a single technique.

4 PressureMove Techniques

We describe four variations of PressureMove techniques to manipulate mouse movement and pressure input simultaneously. All pressure interaction techniques used the thumb sensor to manipulate the parameter in one direction and the middle finger sensor to manipulate the parameter in the reverse direction.

4.1 PressureMove – Naïve

As the name suggests this is a naïve implementation of simultaneous control. In this technique the raw pressure values reported by the pressure sensor are mapped to the object parameter controlled by pressure. Figure 3(a) shows the mapping function - the pressure range is mapped to the complete range of the rotation parameter, i.e. 360° angle. When the user increases pressure the object orientation increases and when they release pressure the orientation reverses i.e., if the initial direction of rotation is clockwise then on releasing pressure the object change orientation in the counter-clockwise direction. When the user releases the pressure sensor the parameter value returns to the starting position. To fix the value the user can left-click before releasing pressure. When the user presses the thumb sensor the object rotates clock-wise and the visual feedback is as shown in Figure 2(b). When the user switches to the sensor located on the middle finger the object rotates counter-clockwise.

4.2 PressureMove – Rate-Based

In this technique each level of the discrete pressure space maps to the speed of rotation of the object as shown in Figure 3(b). When the user maintains pressure at discrete level 1 the object rotates by 1° at each timer event. To move the object faster the user moves higher up within the pressure levels. At level n the object rotates at n degrees per timer event. This mechanism provides the additional benefit of maintaining a given orientation when the user releases the pressure sensor, thus incorporating a clutching mechanism that is not available with the naïve technique. At discrete level 0 the user can tap the pressure sensor to nudge the object by 1° per tap. This gives the user additional fine control when honing in on the target. This tapping was inspired from the Tap-and-Refine technique in [2]. The visual feedback used was the same as for the Naïve implementation.

4.3 PressureMove – Hierarchical

PressureMove-Hierarchical allows users to control rotation in two steps – a coarse-step and a fine-step. The coarse and fine movement is controlled by a discrete pressure mapping. In the coarse-step moving to a pressure level 1 results in rotating the object by 24° (one step is 360°/15levels = 24°) and moving up successive levels rotated the object by 24° per level n (n Є [0,15], n is the coarse-step pressure level). Thus at any pressure level the object is rotated by $n \times 24°$; while in the fine-step moving up each pressure level rotates the object by 1° starting from n. The object rotates from n to $n \times 24-15$ using one sensor and from n up to $n \times 24+15$ using the other sensor where n is the point in the coarse-control when the user switches to fine-control. The user can toggle between coarse- and fine-step by using the left click button. Figure 3(c) shows the pressure vs angle profile for this technique. The dotted line at about 150° indicates the moment at which the user moved from coarse to fine control using left-click. Figure 2(d) and 2(e) show the visual feedback that was provided to the user when using the thumb sensor (so object rotates clockwise). The top triangle of the cursor changes with pressure when the user is performing a coarse-level action (as in Figure 2(d)) and the bottom triangle changes with pressure when the user if performing a fine-level action (as in Figure 2(e)).

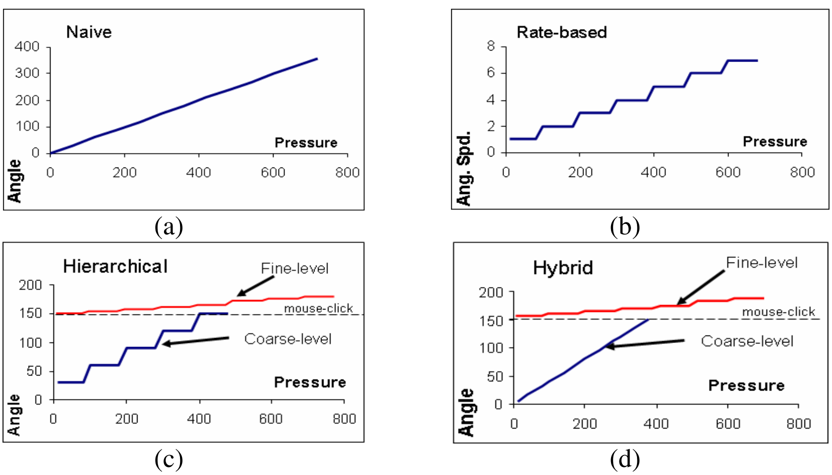


Fig. 3. The pressure mapping functions for each of the PressureMove techniques. a) Naïve implementation b) Rate-based technique c) Hierarchical technique; d) Hybrid technique. The dotted horizontal line in (c) and (d) at Angle = 150 indicates a left-click action. The red line is the fine-level control and the blue line is for the coarse-level.

4.4 PressureMove - Hybrid

Hybrid combines the simplicity available with Naïve with the fine control provided by Hierarchical. The coarse-step of Hierarchical is replaced by the continuous rotation control used in Naïve (see the bottom left part of Figure 3(d)). This enables the user to quickly rotate the object to approximately the desired orientation and then use finer step control to perform a more precise orientation. The fine-control step and the visual feedback mechanism worked exactly as in Hierarchical.

5 Experiment

The goal of this experiment was to evaluate PressureMove as a viable concept for simultaneous control of pressure input and mouse movement.

5.1 Task and Stimuli

The task, shown in Figure 4, required the user to reposition and reorient to a target location and orientation a small object (100×100 pixels) which initially appeared upright and in the left end of the screen. The target, of a slightly larger size than the object appeared to the right of the object. The size, the distance to the object and the orientation of the target were changed as part of the experimental design.

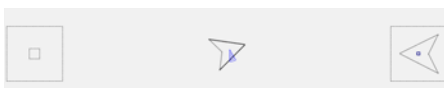


Fig. 4. The experimental task consisted of docking a triangular shaped object over a target. Rotation is controlled using pressure, and displacement controlled with mouse movement.

Users see the object and the target before the beginning of each trial. The trial begins when the user moved the cursor onto the object and pressed the left mouse-click. They reposition and reorient the object to the target location using the different interaction techniques. When the object position and orientation match the target position and orientation, the target bounding rectangle changes to a green color. The user then has to maintain the matching position and orientation for 1 second before the trial is completed. We did this to prevent users from accidentally matching the position and orientation. If the user moves the object away from the matched position, the 1 second timer is reset. The object position and orientation were considered to match those of the target if the difference in pixels and orientation was within the target-fit parameter controlled as factor of the experiment. When the trial is completed the target bounding rectangle briefly turns red and the next trial loads.

5.2 Hardware Configuration and Techniques

Our study used an optical mouse with pressure sensors mounted on its rim (Figure 1). The sensors (model #IESF-R-5L from CUI Inc.) could measure a maximum pressure value of 1.5Ns. Each sensor provided 1024 pressure levels. The application was developed in C# and the sensor was controlled using the Phidgets library [11]. The experiments were conducted in full-screen mode at 1280×800 pixels on an Intel T5600 1.83GHz, Windows Vista OS. Two sensors were mounted on the mouse such that they could be easily accessed by the thumb or the middle finger (as shown in Figure 1). All pressure interaction techniques used the thumb sensor to rotate the object clockwise and the middle finger sensor to rotate the object counter clockwise.

For all the discrete pressure based techniques we used the PressureFish discretiztion function [17] with 15 pressure levels. For the continuous pressure cases we only used pressure values between 0 and 720 as previous research has shown that users find it difficult to maintain pressures at higher values.

5.3 Procedure and Design

The study used a 5×2×3×2 within-participants factorial design. The factors were: Technique (Naive, Rate-based, Hierarchical, Hybrid, Mouse-only), Distance (500 pixels, 1100 pixels), Orientation (60, 135, 270) and Target Fit (tight, loose). A tight target-fit meant that the users had to position the center of the object within ± 4 pixels of the target center and the object orientation has to be within ± 5° of the target orientation. For loose target-fit these figures were ±12 pixels and ±8° respectively.

The order of presentation first controlled for technique and then for distance followed by orientation and target-fit. We explained the techniques and participants were given ample time to practice the techniques at the beginning of the experiment. The experiment consisted of three blocks with each block comprising of two repetitions for each condition. With 5 techniques, 2 distances, 3 orientations, 2 target-fits, 3 blocks, and 2 trials, the system recorded a total of (5×2×3×2×3×2) 360 trials per participant. The experiment took approximately 60 minutes per participant.

5.4 Performance Measure and Participants

The experimental software recorded trial completion time, and number of crossings as dependent variables. Trial completion time (MT) is defined as the total time taken for the user to position and orient the object within the target. The number of crossings (NC) is defined as the number of times the object enters and leaves the target position or orientation for a particular trial. Users were not able to proceed to the next trial without successfully completing the task and so there were no errors for the software to record. Participants were also asked in an exit questionnaire to rank the different pressure control techniques in terms of mental demand, physical demand, effort, overall performance and frustration. Thirteen participants (11 males and 2 females) between the ages of 19 and 40 were recruited from a local university. All participants had previous experience with graphical interfaces and used the mouse in their right hand. None of the participants had worked with a pressure based input device before.

5.5 Results

We used the univariate ANOVA test with participant number as a random factor and Tamhane post-hoc pair-wise tests (unequal variances) for all our analyses.

Completion Time. The average trial completion time was 6.1s (standard deviation 4.9s). Out of a total of 4680 trials 73 outliers (more than 3.5 standard deviations from the group mean) were excluded from further analysis. There was a significant effect of interaction technique ($F_{(4,48)} = 11.15$, $p<0.001$), target-fit ($F_{(1,12)} = 102.9$, $p<0.001$), distance ($F_{(1,12)} = 7.5$, $p<0.02$), Orientation ($F_{(2,24)} = 15.9$, $p<0.001$) and block-number ($F_{(2,24)} = 43.4$, $p<0.001$) on MT. Figure 5 shows the mean trial completion time for each technique and target-fit. Overall, Rate-based was the fastest technique followed by Naive, Mouse, Hierarchical, and Hybrid.

Post-hoc analysis showed that all pairs were significantly different except (Naive, Mouse), and (Mouse, Hierarchical). Block 3 was significantly faster than Block 2 which was significantly faster than Block 1. Users were significantly slower in completing the trials when the target-fit was tight (as opposed to loose); when targets were farther (1100 pixels followed by 500) and when the orientation of target was greater (all combinations significantly different with 270°> 135°> 60°).

Crossings. The average number of crossings per trial across all conditions was 2.5 (standard error = 0.048). There was a significant effect of interaction technique ($F_{(4,48)} = 55.15$, $p<0.001$), target-fit ($F_{(1,12)} = 68.1$, $p<0.001$), distance ($F_{(1,12)} = 7.5$, $p<0.02$), Orientation ($F_{(2,24)} = 19.8$, $p<0.001$) and block-number ($F_{(2,24)} = 13.7$, $p<0.001$) on MT. We found no effect of target distance on number of crossings. Figure 5(c) shows the mean crossings for each technique. Overall Rate-based had the least number of crossings, followed by Mouse, Hierarchical, Naive and Hybrid.

Post-hoc comparisons showed that there was a significant difference between all pairs except the (Rate-based, Mouse), (Mouse, Hierarchical) and (Naive, Hybrid). Block 3 had significantly fewer crossings (mean 2.1) than Block2 (mean 2.4) which in turn had significantly fewer crossings than Block1 (mean 2.9). Users had significantly fewer crossings in the loose target-fit condition (mean 2.1) than in the tight target-fit condition (mean 2.9). Users had significantly fewer crossings when the target orientation was 270° (mean = 2) when compared to 60° (mean = 2.7) or 135° (mean = 2.9). We found no statistical difference in number of crossings between 60° and 135°.

Subjective Ranking. Users ranked the Mouse as the best technique followed by Rate-based, Hybrid, Hierarchical and Naive. Anova test on the overall performance revealed a significant difference in terms of user ranking between the different

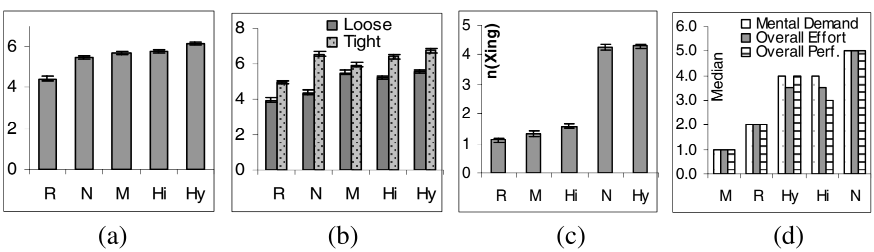


Fig. 5. Mean trial completion time (along the Y-axis in seconds) with standard error bars (a) for each technique and (b) for each technique and target-fit. (c) Mean scores for different techniques. (d) Median user-ranking of different techniques in terms of Mental Demand, Overall Effort and Performance.

techniques ($F(4,64) = 16.6$, $p<0.001$). Post-hoc analysis did not reveal any significant differences between (Mouse, Rate-based) and (Hierarchical, Hybrid) pairs. But all other pairs were significantly different. We found similar rank-ordering of the techniques in terms of Overall Effort, Mental Demand, Physical Demand, and Frustration (see Figure 5(d)).

6 Discussion

Users were constantly improving their performance over the three blocks for both trial completion times and number of crossings. The average MT in Block 3 was 5.0s compared to 5.3s for Block 2 and 6.1s for Block 1. However, the univeriate analysis we used in the previous section did not reveal any significant interaction between technique and block number for both MT and Crossings. Thus while users improved their performance over each block the overall order of the different techniques did not change. Observing improvement over blocks is in line with prior work suggesting that practice allows users to allocate better control to the simultaneous operation of different input dimensions [8].

6.1 Simultaneous Control

As part of the experimental log we collected continuous data of mouse movement and pressure values for each trial. Figure 6 shows typical movement and pressure profiles for the four pressure-based techniques. Each left-right pair is distance and pressure profile for the same trial of a user. However, each technique is from a different user, selected randomly to highlight that the movement profiles shown in the figures are stereotypical. In the left images when the distance is not changing the user has positioned the object near the target whereas the same interpretation is not true for all pressure based techniques. The Rate-based technique being a relative input technique, users don't need to maintain constant pressure to complete the task.

We can see from the figure (Figure 6(a) and 6(b)) that the Naïve implementation does not really encourage simultaneous control of pressure and movement. Users use the first second to complete positioning the object before applying pressure to change orientation. We observe a similar trend with the Hybrid technique.

In the case of the Hierarchical technique, users start applying pressure about the same time that they start moving (see Figure 6(e) and 6(f)). But in the first part of their motion (between 0 and 2s) they mostly focus on moving the object to the right location and then switch attention to orienting (between 2 and 4s) the object.

But in the case of Rate-based technique, users start applying pressure to change orientation at the same time as they are moving the object to position it. In Figures 6(g) and 6(h) we see that in the first 2s the user is both positioning the object while at the same time as applying pressure. However, unlike the Hierarchical technique, they have completed most of the positioning and orienting within the first 2 s and between 2 and 4s they are merely fine-tuning the object. We believe that the open-loop motion for both positioning and orienting coincide making rate-based

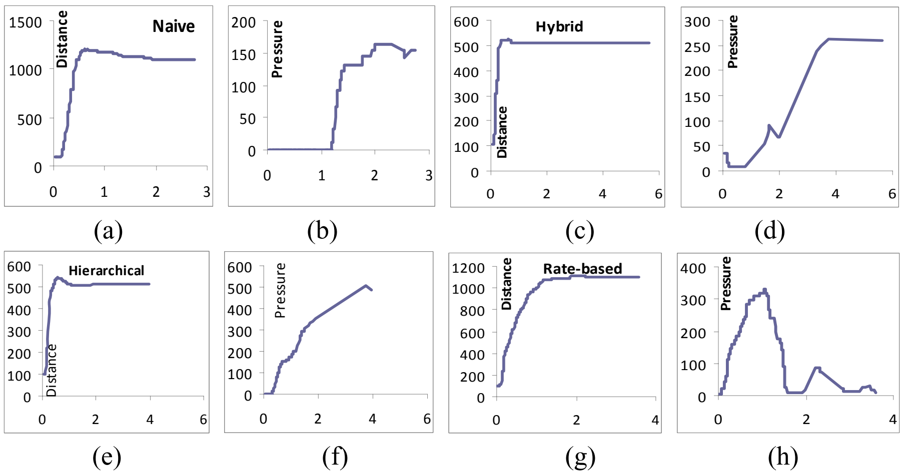


Fig. 6. Traces of a typical user control when using the four PressureMove techniques. The patterns reveal the degree of simultaneity employed in each of the techniques, ranging from low simultaneity with the Naïve technique to high simultaneity with Rate-based.

technique a powerful PressureMove technique. However, we did not test this hypothesis with our data.

We observed similar profiles across all users and believe that PresssureMove technique that's based on a rate-based mapping encourages users to simultaneously control both movement and pressure.

6.2 PressureMove – Rated-Based

PressureMove – Rate-based outperforms all other PressureMove designs; several unique properties of the technique contribute to its superior performance. This technique is based on discrete pressure control. A high degree of control is required to hold and maintain the pressure at given discrete levels; this is facilitated by a small number of discrete pressure levels and the use of discrete fisheye function.

Additionally, since each pressure level is assigned an angular velocity, pressure level 0 brings the rotating object to a halt, at the last applied orientation. The implicit clutching mechanism in the rate-based technique allows smaller close-loop movements than the other techniques. Finally, the technique allows fine adjustments at level 0, by nudging the object by 1° every tap. The fine grain control over angular displacement and the fluidity of this technique facilitates a higher degree of simultaneous control than any of the other systems.

Results similar to ours which show that rate-based technique improves performance in certain types of input devices, have been observed in 3D positioning tasks. Zhai [20] points out that using isometric devices (such as a joystick that self-centers) to operate in a position control mode (or zero-order) results in poorer performance than when operating the device in a rate-based mode (or a first-order).

6.3 Applications

PressureMove can enhance the interactive performance in a number of different applications. In all of the following applications, the simultaneous control of more than one input parameter would ease the task of the operator.

Zoomable User Interfaces. Zoomable user interfaces can largely benefit from the simultaneous control of several parameters. PressureMove can control various parameters by applying pressure to a scalar value and movement to direction of the zooming operation. For instance moving the mouse left or right could zoom in or out respectively, while pressure controls the resolution factor of the zooming operation. The rate-based technique would change the resolution of the zoom operation by one step at each level of angular velocity. Similarly, on a map the mouse movement would pan the document while pressure input zooms in or out. Unlike, most ZUI implementations where the center point of reference is defined by the position of the cursor or cross-hair before transitioning into the zoom, with PressureMove, the position of the cursor can be updated dynamically during zoom transitions, thereby facilitating a larger degree of freedom in moving around a workspace while zooming.

Drawing Applications. Drawing applications facilitate a large number of object positioning tasks with operations that involve rotating elements, scaling and/or skewing. Here, operations requiring coarse movements (such as scaling or skewing an object) could be relegated to the pressure input and precise positioning could be assigned to the mouse movement.

Dynamic Control-Gain. PressureMove could be utilized to dynamically manipulate control-gain ratios. Such manipulation is particularly useful on high resolution, large display interactions on which users operate with fine and coarse resolution.

6.4 Design Recommendations

There are several lessons that designers can take from our investigation:

- Pressure input can be appropriately integrated with mouse movement, such that both dimensions are operated simultaneously. This should result in higher performance gains than operating with either channel separately.
- PressureMove – Rate-based should be the first and preferred implementation of any PressureMove application. The discrete pressure control, fine grain pressure mapping and inherent clutching mechanisms in this technique are favorable properties that could be borrowed to implement other variations.
- Allowing users to gain experience with PressureMove is important and may be necessary in some cases, i.e. new implementation of a PressureMove technique should not be discarded without first giving consideration to proper training.

7 Conclusion

PressureMove is a novel technique that facilitates the simultaneous control of various input parameters. We designed PressureMove to specifically facilitate object rotation with pressure input and object movement with the mouse displacement. We designed and implemented four PressureMove techniques, based on existing pressure-based interactions [2, 14]. Our PressureMove techniques cover the wide spectrum of possibilities with pressure control and mouse displacement mappings. In a study, the Rate-based PressureMove technique, which maps pressure input to angular velocity allowed the maximum amount of simultaneous control of pressure with mouse movement. Users were able to perform a docking task more efficiently and with fewer crossing with the rate-based implementation. We have demonstrated the possibility of simultaneous control of pressure input and mouse movement. We believe other similar interactions involving simultaneous pressure and movement are possible and will enhance the interactive performance on tasks with multiple input dimensions.

We are considering several lines of investigation for future work. We want to apply PressureMove to a number of other applications and tasks. For example, we believe new PressureMove techniques can be developed for navigation interfaces where the user is controlling several parameters simultaneously. We will also investigate the possibility on integrating PressureMove on other devices such as the stylus. While TabletPC interactions already provide pressure as a form of input, the pressure applied can interfere with movement. Therefore additional investigation is required to evaluate the possibility of PressureMove on the stylus.

References

1. Balakrishnan, R., Baudel, T., Kurtenbach, G., Fitzmaurice, G.: The Rockin'Mouse: integral 3D manipulation on a plane. In: Proc. CHI 1997, pp. 311–318 (1997)
2. Cechanowicz, J., Irani, P., Subramanian, S.: Augmenting the mouse with pressure sensitive input. In: Proc. CHI 2007, pp. 1385–1394 (2007)
3. Forlines, C., Shen, C., Buxton, B.: Glimpse: a novel input model for multi-level devices. In: Proc. CHI 2005 Ext. Abstracts, pp. 1375–1378 (2005)
4. Kattinakere, R.S., Grossman, T., Subramanian, S.: Modeling steering within above-the-surface interaction layers. In: Proc. CHI 2007, pp. 317–326 (2007)
5. Kruger, R., Carpendale, S., Scott, S.D., Tang, A.: Fluid integration of rotation and translation. In: Proc. CHI 2005, pp. 601–610 (2005)
6. Kurtenbach, G., Buxton, W.: User learning and performance with marking menus. In: Proc. CHI 1994, pp. 258–264 (1994)
7. MacKenzie, I.S., Soukoreff, R.W., Pal, C.: A two-ball mouse affords three degrees of freedom. In: Proc. CHI 1997 Ext. Abstracts, pp. 303–304 (1997)
8. Masliah, M.R., Milgram, P.: Measuring the allocation of control in a 6 degree-of-freedom docking experiment. In: Proc. CHI 2000, pp. 25–32 (2000)
9. MightyMouse™, `http://www.apple.com/mightymouse/`
10. Mizobuchi, S., Terasaki, S., Keski-Jaskari, T., Nousiainen, J., Ryynanen, M., Silfverberg, M.: Making an impression: force-controlled pen input for handheld devices. In: Proc. CHI 2005 Ext. Abstracts, pp. 1661–1664 (2005)
11. Greenberg, S., Fitchett, C.: Phidgets: easy development of physical interfaces through physical widgets. In: Proc. UIST, pp. 209–218 (2001)
12. Jacob, R.J., Sibert, L.E.: The perceptual structure of multidimensional input device selection. In: Proc. CHI 1992, pp. 211–218 (1992)
13. Jacob, R.J., Sibert, L.E., McFarlane, D.C., Mullen, M.P.: Integrality and separability of input devices. ACM Trans. Comput.-Hum. Interact. 1(1), 3–26 (1994)

14. Ramos, G., Balakrishnan, R.: Zliding: fluid zooming and sliding for high precision parameter manipulation. In: Proc. UIST 2005, pp. 143–152 (2005)
15. Ramos, G.A., Balakrishnan, R.: Pressure marks. In: Proc. CHI 2007, pp. 1375–1384 (2007)
16. Ramos, G., Boulos, M., Balakrishnan, R.: Pressure widgets. In: Proc. CHI 2004, pp. 487–494 (2004)
17. Shi, K., Irani, P., Gustafson, S., Subramanian, S.: PressureFish: A Method to Improve Control of Discrete Pressure-based Input. In: Proc. CHI 2008, pp. 1295–1298 (2008)
18. Wang, Y., MacKenzie, C.L., Summers, V.A., Booth, K.S.: The structure of object transportation and orientation in human-computer interaction. In: Proc. CHI 1998, pp. 312–319 (1998)
19. Zeleznik, R., Miller, T., Forsberg, A.: Pop through mouse button interactions. In: Proc. UIST 2001, pp. 195–196 (2001)
20. Zhai, S.: User Performance in Relation to 3D Input Device Design. Computer Graphics 32(4), 50–54 (1998)
21. Zhai, S., Milgram, P., Buxton, W.: The influence of muscle groups on performance of multiple degree-of-freedom input. In: Proc. CHI 1996, pp. 308–315 (1996)

Bimanual Interaction with Interscopic Multi-Touch Surfaces

Johannes Schöning[1], Frank Steinicke[2], Antonio Krüger[1], Klaus Hinrichs[2], and Dimitar Valkov[2]

[1] DFKI (German Research Center for Artifical Intelligence), Intelligent User Interfaces Department, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany
{schoening,krueger}@dfki.de
[2] Visualization and Computer Graphics Group, Department of Computer Science University of Münster, Einsteinstr. 62, 48149 Münster, Germany
{frank.steinicke,khh,dimitar.valkov}@uni-muenster.de

Abstract. Multi-touch interaction has received considerable attention in the last few years, in particular for natural two-dimensional (2D) interaction. However, many application areas deal with three-dimensional (3D) data and require intuitive 3D interaction techniques therefore. Indeed, virtual reality (VR) systems provide sophisticated 3D user interface, but then lack efficient 2D interaction, and are therefore rarely adopted by ordinary users or even by experts. Since multi-touch interfaces represent a good trade-off between intuitive, constrained interaction on a touch surface providing tangible feedback, and unrestricted natural interaction without any instrumentation, they have the potential to form the foundation of the next generation user interface for 2D as well as 3D interaction. In particular, stereoscopic display of 3D data provides an additional depth cue, but until now the challenges and limitations for multi-touch interaction in this context have not been considered. In this paper we present new multi-touch paradigms and interactions that combine both traditional 2D interaction and novel 3D interaction on a touch surface to form a new class of multi-touch systems, which we refer to as *interscopic multi-touch surfaces* (iMUTS). We discuss iMUTS-based user interfaces that support interaction with 2D content displayed in monoscopic mode and 3D content usually displayed stereoscopically. In order to underline the potential of the proposed iMUTS setup, we have developed and evaluated two example interaction metaphors for different domains. First, we present intuitive navigation techniques for virtual 3D city models, and then we describe a natural metaphor for deforming volumetric datasets in a medical context.

Keywords: Multi-touch Interaction, Interscopic Interaction, 3D User Interfaces.

1 Introduction and Related Work

In recent years, the visualization of and the interaction with three-dimensional (3D) data has become more and more popular and widespread due to the requirements of numerous application areas. Two-dimensional (2D) desktop systems are often limited in cases in which natural and intuitive interfaces are desired. Sophisticated 3D user

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 40–53, 2009.

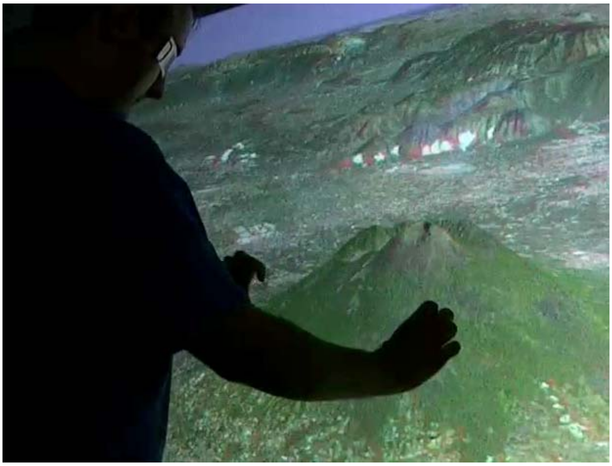


Fig. 1. A user interacting with an interscopic multi-touch surface (iMUTS) in a landscape planning scenario in which depth cues are important, e. g., in a landslide risk management scenario. The geospatial data is displayed in anaglyph stereoscopic mode.

interfaces, as they are provided by virtual reality (VR) systems consisting of stereoscopic projection and tracked input devices, are rarely adopted by ordinary users or even by experts often due to the overall complexity of the user interface. However, a major benefit of stereoscopy is binocular disparity that provides a better depth awareness. When a stereoscopic display is used, each eye of the user perceives a different perspective of the same scene. This can be achieved by either having the user wear special glasses or by using special 3D displays. For this reason, VR systems using tracking technologies and stereoscopic projections of 3D synthetic worlds have a great potential to support a better exploration of complex data sets. However, interaction with stereoscopic display may become difficult. For instance, in a stereoscopically displayed 3D scene, it may be hard to access distant objects [1]. This applies in particular if the interaction is restricted to a 2D touch surface. Objects might be displayed with different parallax paradigms, i.e., negative, zero, and positive parallax, resulting in different stereoscopic effects. Interaction with objects that are displayed with different parallaxes is still a challenging task even in VR-based environments. In addition, while the costs as well as the effort to acquire and maintain VR systems have decreased to a moderate level, these setups are only used in highly specific application scenarios within some VR laboratories. In most human-computer interaction processes, VR systems are only rarely applied by ordinary users or by experts – even when 3D tasks have to be accomplished [2]. One reason for this is the inconvenient instrumentation required to allow immersive interactions in such VR systems, i.e., the user is forced to wear stereo glasses, tracked devices, gloves etc. Furthermore, the most effective ways for humans to interact with synthetic 3D environments have not finally been determined [2, 1]. Even the WIMP (window, icon, menu, pointing device) desktop metaphor [3] has its limitations when it comes to direct manipulation of 3D data sets [4], e.g., via 3D widgets [5]. And as a matter of fact, 2D interactions are performed best with 2D devices usually supporting only two degrees of freedom

(DoF) [6, 7]. Hence, 3D user interfaces are often the wrong choice in order to accomplish tasks requiring exclusively or mainly two-dimensional control [2, 6]. Most 3D applications also include 2D user interface elements, such as menus, texts and images, in combination with 3D content. While 3D content usually benefits from stereoscopic visualization, 2D graphical user interfaces (GUIs) items often do not have associated depth information. Therefore, interactions between monoscopic and stereoscopic elements, so-called *interscopic interactions* [8], have not been fully examined with special consideration of the interrelations between the elements. Due to recent developments in the entertainment market, multi-touch has received considerable attention for 2D user interfaces. Multi-touch refers to a set of interaction paradigms, which allow users to control applications with several simultaneously performed touches. While multi-touch has shown its usefulness for 2D interaction by providing more natural and intuitive techniques such as 2D translation, scaling and rotation, it has not been considered if and how these concepts can be extended to 3D multi-touch interfaces. The challenge of 3D interaction with stereoscopically displayed objects using a multi-touch user interface has only been considered rudimentarily. Multi-touch surfaces can be realized by using different technologies, ranging from capacitive sensing to video analysis of infrared or full colour video images [9–13]. Recently, the FTIR (frustrated total internal reflection) technology and its cheap footprint [10] (being a camera based multi-touch solution) has accelerated the usage of multi-touch. With today's technology it is now possible to apply the basic advantages of bi-manual interaction [14–19] to any suitable domain. Multi-touch surfaces can be easily integrated into systems supporting interaction of multiple users with two- dimensional data sets. Bill Buxton gives a comprehensive overview of the history of development of multi-touch surface1. Another benefit of multi-touch technology is that the user does not have to wear inconvenient devices in order to interact in an intuitive way. The DoF are restricted by the physical constraints of the touch screen. In combination with autostereoscopic displays, such a system can avoid any instrumentation of the user, while providing an advanced user experience. However, the benefits and limitations of using multi-touch in combination with stereoscopic display have not been examined in-depth and are not well understood [20].

Only some researchers have addressed the problem of 3D interaction on a 2D surface. Grossman et al. [21] presented a suite of gestural interaction techniques for use with a spherical 3D volumetric display. However, in this setup the user can already use 3D inputs on a touch sphere in comparison to available inputs on a two-dimensional plane. To allow interactions in the 3D space Benko et al. [22] introduced the Balloon Selection, a 3D interaction technique that is modelled after the real-world metaphor of manipulating a helium balloon attached to a string. Balloon Selection allows for precise 3D selection in the volume above a tabletop surface by using multiple fingers on a multi-touch sensitive surface in conjunction with a 3D tracked glove. Although, this technique enables two-and-a-half-dimensional (2.5D is an informal term used to describe visual phenomena which is actually 2D with 3D looking graphics) selection, real 3D interaction has not been considered in their work.

In this paper we discuss the challenges for iMUTS-based interaction paradigms, and present two new interaction techniques, which underline their benefits (see figure 1). The remainder of this paper is structured as follows: First we discuss the general interaction problem with interscopic data in Section 2. In Section 3, we describe two new multi-touch interactions metaphors that are best suited for multi-touch input and that can be interfaced by the described iMUTS setup. Section 4 describes the iMUTS technical setup in detail. Section 5 presents an initial evaluation of the presented concepts. Finally, Section 6 concludes the paper and gives an overview about future work.

2 Interaction Challenges

While it appears quite obvious to use multi-touch for 2D interaction, one might argue that it is not useful to limit interaction with 3D data to a 2D touch surface. As mentioned in the introduction, the usage of complex VR systems still requires much user instrumentation, and it is still quite complex to interact in the 3D space [1]. And as a matter of fact, 2D interaction devices usually supporting only two DoFs [6, 7] are still in widespread use, although they are not the optimal devices for 3D interaction. These both facts motivated us to consider multi-touch interfaces with respect to the requirements of current 3D graphical applications.

2.1 Taxonomy of User interfaces

In Figure 2, a taxonomy is illustrated which classifies iMUTS paradigms within the broad field of graphical user interfaces. This classification includes conventional desktop systems as well as VR and current multi-touch environments. The taxonomy model is based on a coordinate system involving three axes representing generality, presence and nativeness. Generality captures the variety of interaction tasks that can be performed with the corresponding user interface, presence measures the degree to which a user believes the virtual environment (VE) is part of the physical surrounding or vice versa, and nativeness denotes how inartificial the user interface appears with respect to the required instrumentation. Current graphical user interfaces can be classified according to these characteristics. For example, as mentioned in Section 1,

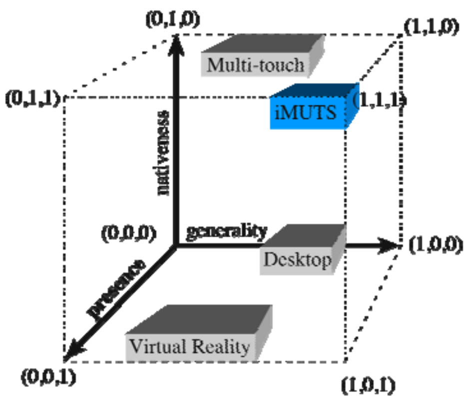


Fig. 2. Taxonomy of different user interface paradigms and the integration of GUIs based on iMUTS. The three axes representing *generality*, *presence* and *nativeness*.

desktop-based environments are most appropriate for 2D interaction tasks, but are not optimal for immersive 3D interaction. Hence they are specialized rather than general in this case. Furthermore, virtual scenes are usually displayed monoscopically in desktop-based environments, with only a small field of view and no-head tracking is supported; therefore, the user's sense of presence is often lower in this case. Although interaction with traditional input devices is quite natural, the instrumentation with keyboard and/or mouse is less natural than using the hands directly as it is possible with multi-touch user interfaces (see figure 2). For these reasons, desktop systems are mapped to the area close to the origin (0, 0, 0). VR systems increase presence and are suitable for 3D interaction, but lack support for 2D interaction (cf. Section 1). Furthermore, instrumentation by haptic devices such as data gloves and/or immersive display systems such as head-mounted-displays leads to non-native interaction. Hence we arrange VR systems similar to desktop-based environments, but closer to the front bottom edge. Current multi-touch user interfaces are suitable for ordinary desktop tasks, they provide the same sense of presence, but they also allow native interaction by the hands. Hence they are arranged above desktop systems showing increased nativeness. We believe that the iMUTS has the potential to increase presence (e.g., due to stereoscopic projection) in comparison to existing multi-touch user interfaces, and that they will support intuitive 2D as well as 3D interactions while not requiring additional instrumentation of the user. Hence, iMUTS user interfaces may meet most requirements for current GUIs. We admit that this taxonomy is neither perfect nor universal, but it points out some of the benefits that we believe iMUTS paradigms may provide.

To summarize: with iMUTS, humans may interact spontaneously while no heavy instrumentation is required. Further advantages of iMUTS are:

- Multiple Users: iMUTS easily supports the interaction between different users.
- Intuitivity: People have developed sophisticated skills for sensing and manipulating their physical environments [23]. iMUTS supports these skills.
- Spontaneity: User can switch between different tasks spontaneously.
- Costs: The presented iMUTS setup can be realized with low costs.

As mentioned in Section 1, interaction with stereoscopically displayed objects is still a challenging task [1], in particular when the interaction is restricted to a 2D touch surface. This is the main limitation of an iMUTS. In the following section, we explain this issue and discuss solutions, which have proven their usability in other domains.

2.2 Parallax Problems

In order to display graphical content stereographically, two half-images have to be generated, i.e., one for each eye. When using stereoscopic projection, a 3D impression occurs due to binocular disparity, which means that objects in space are projected to different positions on the screen. The corresponding horizontal displacement results in essentially three different stereoscopic display paradigms: negative, zero and positive parallax. Objects having zero parallax are displayed monoscopically and therefore are ideally suited for multi-touch interaction (see green-colored box in Figure 3). Both eyes perceive the same image, which causes a two-dimensional impression. As mentioned in the introduction, for such a situation multi-touch interfaces have

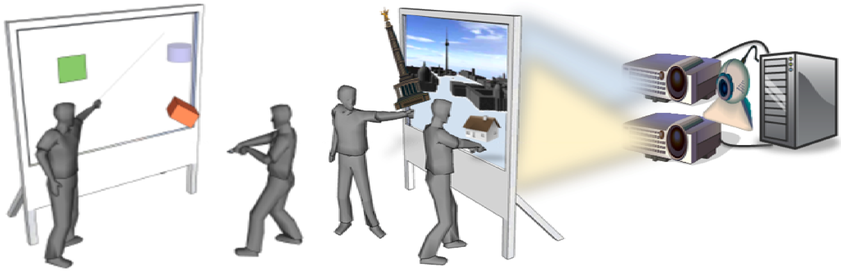


Fig. 3. (left) Illustration of two users interacting with stereo content, as well as monoscopic content (green rectangle: zero parallax, orange-colored box: negative parallax, purple-colored cylinder: positive parallax), and (right) during an interaction with a city planning application on an interscopic multi-touch surface (iMUTS))

considerable potential to enhance the interaction process, in particular when 2D manipulations are intended. Objects with positive parallax appear behind the touch screen and therefore cannot be accessed directly due to the screen limiting the reach of the user (see purple-coloured cylinder in Figure 3). This is a problem for any kind of direct interaction in stereoscopic environments, and several approaches address this issue [1, 24], e.g., distant objects behind the screen can be selected by casting a virtual ray [25]. Objects displayed with negative parallax appear in front of the projection screen (see orange-coloured box in Figure 3) When the user wants to interact with such objects by touching, s/he is limited to touching the area behind the objects since multi-touch screens capture only direct contacts. Therefore, the user virtually has to move fingers or her/himself through virtual objects, and the stereoscopic projection is disturbed. Consequently, immersion may get lost. This problem is a common issue known from two-dimensional representation of the mouse cursor within a stereoscopic image. While the mouse cursor can be displayed stereoscopically on top of stereoscopic objects [8], movements of real objects in the physical space, e.g., the user's hands, cannot be constrained such that they appear only on top of virtual objects. Therefore direct grabbing of objects in front of the touch screen is not possible, and moreover, the hands may interfere with the stereoscopic effect. When objects are displayed with negative parallax, any input devices tracked with six DoF support direct interaction. However, we admit that this requires an instrumentation of the user again. Hence, multi-touch devices in combination with image-based gesture recognition may be more appropriate, but have not been used in our setup until now.

3 Interaction Metaphors

As mentioned above, 3D visualization applications combine two-dimensional with three-dimensional content. While 3D data has the potential to benefit from stereoscopic display, visualization and interaction with 2D content should be restricted to two dimensions [2]. In this section, we discuss aspects that have to be taken into account when designing a multi-touch user interface for interscopic interaction. Therefore, we present interaction metaphors with interscopic data for a city-planning scenario as well as for medical volume deformation. The described metaphors are not limited to the city or medical domain, but can be applied also to other scenarios. Both interaction concepts are motivated by the work of Wilson [26], which combines multi-touch sensing with a physics engine. In the first scenario, users can manipulate a virtual window analog to a plate on a ball and socket joint. This enables an intuitive way of travelling through a 3D city model. Although such 3D city models are usually represented by polygonal descriptions, the iMUTS interaction paradigms are not limited to polygonal data. To underline that also other formats, such as raster-based and volumetric data representations can be interfaced with iMUTS-based interaction paradigms, we present a second interaction metaphor with which users can directly manipulate a volume dataset by multi-touch deformation.

3.1 Windows on the World Interaction with Interscopic Data

During our long-term cooperation with the urban development, city planning and transport planning office as well as the land surveying and land registry office of the city of Münster, Germany, the needs for intuitive interfaces for the urban planning process (see Figure 4 (left)) became obvious. City planning tasks are highly cooperative and dynamic problems that involve several users with different expertise according to Broll [27], e. g., architects, planners, designers, politicians, potential home buyers, etc. Due to their naturalness, multi-touch user interfaces have considerable potential to increase performance and creativity for such a setup. All co-operators can be involved from the beginning and can design simultaneously. While many research groups address visualization techniques, for example for seismic exploration data, less effort has been undertaken in order to provide intuitive 3D user interfaces for such data sets.

For iMUTS, we designed a new interaction metaphor to easily navigate through a stereoscopically displayed 3D city model. The *Windows on the World* navigation metaphor allows users to easily navigate through a virtual city model displayed in an application with 2D graphical user interface elements and a 3D virtual world. This metaphor should not be confused with the work of Feiner et al. [28], a prototype heads-up window system intended for use in a 3D environment. Our metaphor is

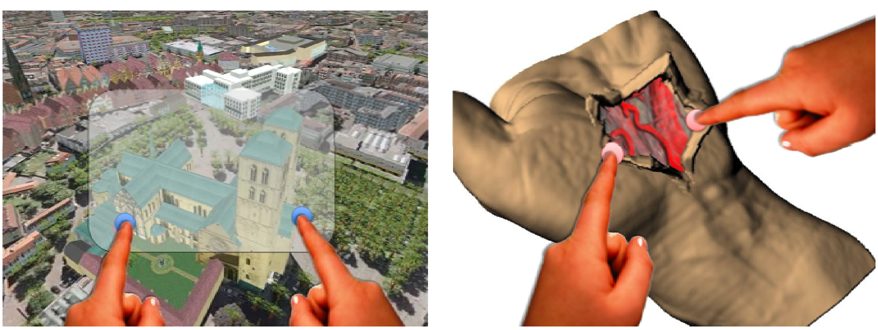


Fig. 4. Two example applications which can be interfaced by novel user interfaces based on the iMUTS paradigms: (left) navigating through a virtual city model with the Window of the World metaphor, and (right) multi-touch deformation of volume data in a medical scenario

based on a virtual analogy of a physical plate that is attached to a ball and socket joint. Hence, in analogy to such a physical setup in our approach users can manipulate a virtual window overlaid on the virtual city or landscape model. The window is displayed with zero parallax so the user has the feeling that she can manipulate this virtual window. By touching the window the user can control the point of view, e.g., pressing on the bottom of the virtual window will cause that the user will look up (see Figure 4 (left)). By touching the top window edge, the user will look down. Turning the window to the left will cause a right rotation and vice versa. By pushing the middle of the window with both hands, the user navigates forward through the virtual model. The user can also rotate the virtual window to rotate the view. Zooming the window will enlarge the virtual display; thereby widening the opening angle of the virtual camera. The metaphor is independent from the device via which the user gives touch input. In addition to FTIR-based multi-touch surfaces, the user can use mobile devices that are capable to track multiple touches simultaneously. We have tested the metaphor with the described FTIR-based setup as well as with Apple's iPhone. The landscape planning application is based on NASA's World Wind2 using the Java-based SDK. The NASA World Wind visualization platform is open source and comes with a rich SDK for data set and interface customization.

3.2 Volume Deformation with Interscopic Data

Volume deformation is one example technique that allows users to modify volume data by means of warping. We have developed techniques for deforming volumetric data sets based on physically inspired approaches. In our system, user interaction mainly consists of two operations: specifying and applying a cut, as well as deforming the newly cut object. These operations are usually applied sequentially and can therefore be handled independently during the interaction process. The common problem is that a three-dimensional object is transformed, while the user only sees a two-dimensional projection onto the screen and is using two-dimensional input devices. We implemented this deformation originally for mouse-based interaction. Due to the reduced degrees of freedom, the user interface is simplified by providing a small set of predefined cuts specified by cutting templates, which may be resized and positioned freely within the scene. These templates share some resemblance with a cookie cutter, with the main difference that they do not start cutting as soon as they touch an object, but only when the user explicitly initiates the cut. This is a typical example, in which mouse-based as well as 3D interaction is limited and does not support the mental model of a deformation.

For this particular kind of interaction, multi-touch has considerable potential to provide a natural as well as effective user interface. Figures 5 (right) and 4 (right) show an illustration of this approach. The user can intuitively define the deformation by means of warping the volume data set at certain points simply by touching and moving with multiple inputs simultaneously, for example by using the edge or part of the palm of one or both hands. We have implemented this volume deformation with stereoscopic projection, which we restricted to almost zero parallax, i.e., parts of the volume were slightly above and below the touch surface. This kind of projection provides a sufficient three-dimensional visualization of the volume data set, while the

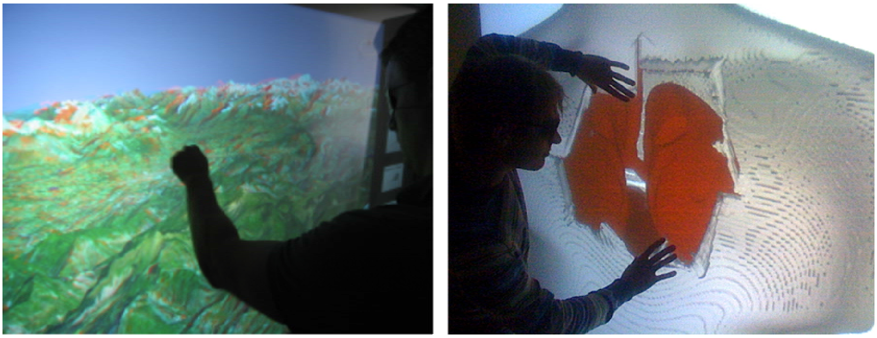


Fig. 5. Two examples of applications interfaced by the iMUTS interaction paradigms. (Left) a user interacts in a landscape-planning scenario on an FTIR multi-touch wall with an anaglyph-based stereoscopic projection. (Right) a user performs volume deformation in a medical scenario using a large FTIR-based passive back projection iMUTS.

interaction via the two-dimensional surface with the 3D data can still be performed without requiring additional tracked input devices.

Our physical-inspired implementation of the deformation of volume data sets exploits the 3D ChainMail algorithm in combination with a GPU-based ray-casting renderer in order to ensure interactive frame rates. We have experienced that this kind of physics-based interaction supports users because their knowledge from the real-world can be transferred easily to virtual interaction tasks [8, 26].

4 iMUTS - Technical Setup

In this section, we describe the system components of our iMUTS setup and discuss challenges for user interfaces based on such an interscopic multi-touch environment. There are several ways of combining multi-touch screens with a stereoscopic display in order to improve depth perception while providing intuitive interaction paradigms without further instrumentation of the user. Our multi-touch interscopic wall prototype is based on the FTIR principle introduced by Han [10]. We use a 180cm×220cm and 1.1cm thick sheet of acrylic, whose edges have been polished clear, as an optical waveguide. This sheet is edge-lit by high-power infrared LEDs, which are placed directly against the polished edges so as to maximize coupling into total internal reflection, while a digital video camera equipped with a matching band-pass filter is mounted orthogonally. Total internal reflection keeps the light trapped within the sheet, except at points where it is frustrated by some object (e. g. finger) in optical contact, causing light to scatter out through the sheet towards the camera, a PointGrey Dragonfly23 . The acrylic plate was mounted onto a wall and a wide-angle lens attached to the PointGrey Dragonfly 2 equipped with a matching infrared band-pass filter was mounted orthogonally at a two-meter distance. For display, an HKS rear projection screen is used and attached behind acrylic plate. Only simple image processing operations (rectification, background subtraction, noise removal, and connected components analysis) are required for each frame, and for that a Java multi-touch library4, developed at the Deutsche Telekom Laboratories, has been used and extended. The library is released under the GNU Public License. It contains a set of common algorithms designed to work with any multi-touch system, such as routines to label connected components and track features. By using an application layer, it is easy to manipulate objects and transform (position, rotate, scale) them. The library also comes with a module for accessing cameras (currently just the Point-Grey Dragonfly2 is supported), and the library is able to stream the data into the TUIO protocol [29], the de facto standard for broadcasting multi-touch events over the network. Video is captured as 8-bit monochrome at 30 fps at a resolution of 1024 × 768 (corresponding to 2, 1mm 2 precision on the surface). For stereoscopic projection we have tested two different setups landscape planning and volume deformation in a medical scenario:

- a passive stereoscopic back projection with an FTIR based interactive wall (see Figure 5 (right)),
- and a simple anaglyph-based stereoscopic projection again based on an FTIR multi-touch wall (see Figure 5 (left)).

The passive stereoscopic projection screen is illustrated in Figure 3. Two DLP projectors with a resolution of 1248 × 1024 provide half images for the left and right eye of the user. The half images are linearly filtered such that users have to wear corresponding polarized glasses. In the case of the anaglyph mode, the half images can be displayed by one projector with a resolution of 1920 × 1080. Colour masks applied to the rendering processes mask both half images such that the user can separate the half images with anaglyph glasses. For both setups the images are rendered by a computer with Intel dual-core processors, 4 GB of main memory and an nVidia GeForce 8800 GTX for rendering purposes.

5 Initial Evaluation

We conducted an initial user study to compare the subjects' self-reported experiences when interacting with interscopic multi-touch surfaces in comparison to traditional multi-touch surface. 2 female and 6 male subjects participated in the study (ages 21-33, mean 23.9 years). All subjects were members of the Department of Geoscience at university of Münster, Germany and familiar with landscape planning tasks. The study was set up as a within-subject design, where all participants had to perform the following task in two different conditions (interscopic multi-touch display and multi-touch surface): subjects had to judge the landslide risk in a certain region in the Alps. The data was displayed in anaglyph stereoscopic mode in the interscopic condition and the Windows on the World Interaction metaphor was used. In the "pure" multi-touch condition the Windows on the World Interaction metaphor was also used but the data was displayed directly on the interactive surface. The subjects had to explore a certain region by navigating to different viewports, to make a judgment about the region with the highest landslide risk. After the actual test, subjects were asked to rate both techniques by filling out a modified version of the "user interface evaluation questionnaire" of ISO 9241-9 with only a single Fatigue (seven-point rating; higher scores denote a better rating). The total time of the experiment was about 15 minutes for each participant.

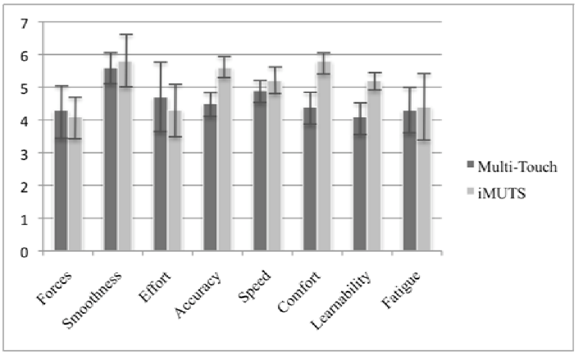


Fig. 6. Results of the user interface evaluation questionnaire ISO 9241-9 (seven-point rating; higher scores denote a better rating)

The answers varied strongly between subjects, which is reflected in the large confidence interval. Figure 6 shows that subjects rated accuracy with 5.6 on average when using the iMUTS setup. This is an increase of 20 % in comparison to the evaluation of the traditional multi-touch map interaction. Only the differences in the categories accuracy, comfort and learnability are significant at the 5% level. In general, from the user comments they preferred the interscopic, because they felt more embedded in the scenario and it was easier to get depth information, which is important for evaluating landslide risk. They tended to perform the task faster, because they could perform actions (e.g. panning and zooming) more accurately and navigate more easily to the spots they wanted to inspect than with a pure multi-touch system. In general, the users had no problem performing the Windows on the World Interaction, after basically getting the idea behind the interaction metaphor. Overall, the users preferred the iMUTS system and they gave us comments like: "It feels like I am there." or "Hopefully the fieldwork will be as easy as manipulating the whole world from my window board". Both statements underline the argument that the user believes the iMUTS is part of the physical surrounding or vice versa and supporting presence.

6 Conclusion

In this paper, we have introduced a new interscopic multi-touch paradigm that combines traditional 2D interaction performed in monoscopic mode with 3D interaction and stereoscopic projection. We discussed challenges and potentials for the use of multi-touch interfaces for the interaction with interscopic data. In addition, we have introduced two different systems of iMUTS:

- a passive back projection with an FTIR-based interactive wall (see Figure 5),
- and a simple anaglyph-based stereoscopic projection again based on an FTIR multi-touch wall.

For both applications, we have highlighted two new multi-touch interaction metaphors, which benefit from multi-touch in combination with monoscopic or stereoscopic projection: the Window on the World navigation method and direct volume deformation. Furthermore, we have presented two different application scenarios, i.e., city & landscape planning and medical exploration that can be interfaced with user interfaces based on iMUTS. We believe that iMUTS has great potential to fill the gap between WIMP and VR systems to form the basis of the next generation of 2D and 3D user interfaces. They provide intuitive, fast and spontaneous access to 3D information for multiple users at low costs without requiring user instrumentation. Moreover, we have outlined some challenges and limitations that might occur in such scenarios. Currently, multi-touch walls are horizontally or vertically mounted. VR-based display devices such as the responsive workbench allow users to turn the display from horizontal to vertical. In contrast to vertical multi-touch surfaces, horizontal ones provide the possibility to place physical objects on the surface 5. For some application domains, it might be beneficial to present the stereoscopic content on a non-planar surface like Microsoft's sphere [30]. With respect to the surface geometry, it might be possible that in some areas objects may not be placed due to instabilities caused by gravitation. The special problem of interaction with stereoscopic data displayed with negative parallax has not been addressed in detail within the scope of this paper and will be considered in future iMUTS. However, in such a case, the user can use arbitrary input devices that can be tracked with six degrees of freedom. Another solution might be to allow a user to interactively change the parallax of objects by using a mobile device attached to the user's body as a "soft slider". If the touch surface is portable, the screen can be moved through the VE (analog to the 3D Window on the World metaphor) until desired objects are displayed with zero or negative parallax and interaction can be performed as described above. In order to provide stereoscopic images without the need for glasses, in future work we will set up a transparent multi-touch wall in front of an autostereoscopic display. Hence, natural and intuitive interfaces can be provided which do not require any instrumentation, but support 3D interaction. In addition we will carry out further user studies with this next prototype.

References

1. Bowman, D., Kruijff, E., LaViola, J., Poupyrev, I.: 3D User Interfaces: Theory and Practice. Addison-Wesley, Reading (2004)
2. Balakrishnan, R.: A Grant of 3D. Why aren't we using 3D user interfaces, and will we ever? (3D User Interfaces Keynote) (2006)
3. Myers, B.A.: A taxonomy of window manager user interfaces. IEEE Comput. Graph. Appl. 8(5), 65–84 (1988)
4. Chun, W.S., Napoli, J., Cossairt, O.S., Dorval, R.K., Hall, D.M., Purtell II, T.J., Schooler, J.F., Banker, Y., Favalora, G.E.: Spatial 3-D Infrastructure: Display-Independent Software Framework, High-Speed Rendering Electronics, and Several New Displays. In: Stereoscopic Displays and Virtual Reality Systems XII, vol. 5664, pp. 302–312 (2005)
5. Conner, D.B., Snibbe, S.C., Herndon, K.P., Robbins, D.C., Zeleznik, R.C., van Dam, A.: Three-Dimensional Widgets. In: Symposium on Interactive 3D Graphics (1992)
6. Hanson, A.J., Wernert, E.: Constrained 3D Navigation with 2D Controllers. In: Proceedings of Visualization 1997, pp. 175–182. IEEE Computer Society Press, Los Alamitos (1997)

7. Salzman, T., Stachniak, S., Stürzlinger, W.: Unconstrained vs. Constrained 3D scene manipulation. In: Nigay, L., Little, M.R. (eds.) EHCI 2001. LNCS, vol. 2254, p. 207. Springer, Heidelberg (2001)
8. Steinicke, F., Ropinski, T., Bruder, G., Hinrichs, K.: Interscopic User Interface Concepts for Fish Tank Virtual Reality Systems. In: Proceedings of the Virtual Reality, pp. 27–34 (2007)
9. Dietz, P., Leigh, D.: DiamondTouch: a multi-user touch technology. In: Proceedings of the 14th annual ACM symposium on User interface software and technology, pp. 219–226 (2001)
10. Han, J.Y.: Low-cost multi-touch sensing through frustrated total internal reflection. In: UIST 2005: Proceedings of the 18th annual ACM symposium on User interface software and technology, pp. 115–118. ACM, New York (2005)
11. Malik, S., Laszlo, J.: Visual touchpad: a two-handed gestural input device. In: Proceedings of the 6th international conference on Multimodal interfaces, pp. 289–296 (2004)
12. Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., Oulasvirta, A., Saarikko, P.: It's mine, don't touch!: interactions at a large multi-touch display in a city centre. In: CHI 2008: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, pp. 1285–1294. ACM, New York (2008)
13. Schöning, J., Brandl, P., Daiber, F., Echtler, F., Hilliges, O., Hook, J., Löchtefeld, M., Motamedi, N., Muller, L., Olivier, P., Roth, T., von Zadow, U.: Multi-touch surfaces: A technical guide. Technical report, Technical University of Munich (2008)
14. Buxton, W., Myers, B.: A study in two-handed input. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 321–326 (1986)
15. Lee, S., Buxton, W., Smith, K.: A multi-touch three dimensional touch-sensitive tablet. ACM SIGCHI Bulletin 16(4), 21–25 (1985)
16. Matsushita, N., Rekimoto, J.: HoloWall: designing a finger, hand, body, and object sensitive wall. In: Proceedings of the 10th annual ACM symposium on User interface software and technology, pp. 209–210 (1997)
17. Epps, J., Lichman, S., Wu, M.: A study of hand shape use in tabletop gesture interaction. In: CHI 2006: CHI 2006 extended abstracts on Human factors in computing systems, pp. 748–753. ACM, New York (2006)
18. Wu, M., Balakrishnan, R.: Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In: Proceedings of the 16th annual ACM Symposium on User Interface Software and Technology, pp. 193–202 (2003)
19. Schöning, J., Hecht, B., Raubal, M., Krüger, A., Marsh, M., Rohs, M.: Improving Interaction with Virtual Globes through Spatial Thinking: Helping users Ask "Why?". In: IUI 2008: Proceedings of the 13th annual ACM conference on Intelligent User Interfaces. ACM Press, New York (2008)
20. Steinicke, F., Schöning, J., Krüger, A., Hinrichs, K.: Multi-touching cross-dimensional data: towards direct interaction in stereoscopic display environments coupled with mobile devices. In: AVI 2008: Workshop on designing multi-touch interaction techniques for coupled private and public displays (2008)
21. Grossman, T., Wigdor, D., Balakrishnan, R.: Multi-finger gestural interaction with 3d volumetric displays. In: SIGGRAPH 2005: ACM SIGGRAPH 2005 Papers, pp. 931–931. ACM, New York (2005)
22. Benko, H., Feiner, S.: Balloon selection: A multi-finger technique for accurate low-fatigue 3d selection. 3D User Interfaces (2007)
23. Ullmer, B., Ishii, H.: Emerging frameworks for tangible user interfaces. IBM Systems Journal 39(3), 915–931 (2000)

24. Pierce, J., Forsberg, A., Conway, M., Hong, S., Zeleznik, R., Mine, M.: Image Plane Interaction Techniques in 3D Immersive Environments. In: ACM Symposium on Interactive 3D Graphics, pp. 39–44 (1997)
25. Mine, M.: Virtual Environments Interaction Technqiues. Technical Report TR95-018, UNC Chapel Hill Computer Science (1995)
26. Wilson, A.D., Izadi, S., Hilliges, O., Garcia-Mendoza, A., Kirk, D.: Bringing physics to the surface. In: UIST 2008: Proceedings of the 21st annual ACM symposium on User interface software and technology, pp. 67–76. ACM, New York (2008)
27. Broll, W., Lindt, I., Ohlenburg, J., Wittkämper, M., Yuan, C., Novotny, T., Fatah, A., Mottram, C., Strothmann, A.: Arthur: A collaborative augmented environment for architectural design and urban planning. Journal of Virtual Reality and Broadcasting (JVRB) 1(1), 1–10 (2004)
28. Feiner, S., MacIntyre, B., Haupt, M., Solomon, E.: Windows on the world: 2D windows for 3D augmented reality. In: Proceedings of the 6th annual ACM symposium on User interface software and technology, pp. 145–155. ACM, New York (1993)
29. Kaltenbrunner, M., Bovermann, T., Bencina, R., Costanza, E.: TUIO: A protocol for table-top tangible user interfaces. In: Proc. of the The 6th Intl. Workshop on Gesture in Human-Computer Interaction and Simulation (2005)
30. Benko, H., Wilson, A.D., Balakrishnan, R.: Sphere: multi-touch interactions on a spherical display. In: UIST 2008: Proceedings of the 21st annual ACM symposium on User interface software and technology, pp. 77–86. ACM, New York (2008)

Multimodal Media Center Interface Based on Speech, Gestures and Haptic Feedback

Markku Turunen[1], Jaakko Hakulinen[1], Juho Hella[1], Juha-Pekka Rajaniemi[1], Aleksi Melto[1], Erno Mäkinen[1], Jussi Rantala[1], Tomi Heimonen[1], Tuuli Laivo[1], Hannu Soronen[2], Mervi Hansen[2], Pellervo Valkama[1], Toni Miettinen[1], and Roope Raisamo[1]

[1] University of Tampere, Tampere Unit for Computer-Human Interaction, Finland
firstname.surname@cs.uta.fi
[2] Tampere University of Technology, The Unit of Human-Centered Technology , Finland
firstname.surname@tut.fi

Abstract. We present a multimodal media center interface based on speech input, gestures, and haptic feedback (hapticons). In addition, the application includes a zoomable context + focus GUI in tight combination with speech output. The resulting interface is designed for and evaluated with different user groups, including visually and physically impaired users. Finally, we present the key results from its user evaluation and public pilot studies.

Keywords: Speech, haptics, gestures, multimodal interaction, media center.

1 Introduction

Multimodal interaction may bring many benefits for interaction with digital home appliances and their digital content. The so-called media centers, which combine the commonly used digital media (television broadcasts, music, photographs etc.) to a single device/application, form one potential area. There are plenty of such systems, both hardware solutions and software applications for personal computers and game consoles. Typically, they are used with a remote controller, which, in many cases, makes the interaction clumsy or even too complex for some users.

In this paper we present how speech input, gestures, and haptic feedback can provide an efficient interface for media centers. Furthermore, we present how they can be made accessible for different user groups. For blind users, speech output and haptic feedback provides full access to the information, while the zoomable GUI makes it accessible for visually-impaired (partly seeing) users. Finally, speech input combined with voice activity and blow detection makes the interface usable for physically impaired people who cannot use their hands at all.

2 Media Center Application

As a part of the project TÄPLÄ (Ambient Intelligence based on Sound, Speech and Multisensor Interaction) we are developing multimodal interaction methods for home

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 54–57, 2009.

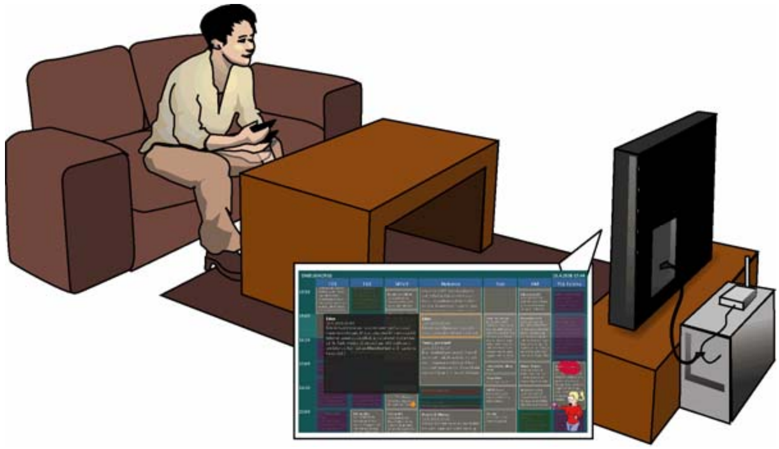


Fig. 1. Media center setup with the EPG user interface

environments. After a large consumer survey [3], we have focused on developing a multimodal user interface for a media center. This application area is rapidly becoming popular in homes, and provides opportunities and challenges for user multimodal interaction [1,2]. Our media center provides users full control over digital television content, including an advanced electronic program guide (EPG).

Our solution is based on a low-cost PC (Athlon X2 3800) running the media center server software, a mobile phone (Nokia N95) as an interface device, a wireless access point to connect these together, and a high-definition 40” television (Figure 1). The main application is written in C# and runs under Windows XP. In it includes speech recognition and speech synthesis. The mobile device has embedded gesture recognizer, speech recognizer, haptic feedback controller, speech synthesizer, and a GUI. All mobile technology components are native Symbian applications while the GUI and main logic is written in MIDP 2.0.

3 Multimodal User Interface

The system has two graphical displays; the large television screen and the mobile phone display. Both devices have also speech output capability, and the mobile device provides haptic feedback. As an alternative for the mobile phone, we have implemented a wheelchair mounted wireless microphone solution for physically impaired users. It is, however, out of the scope of this paper.

The main GUI on the television screen consists of sections for different application functionality. Here, we focus on the EPG interface. It consists of a matrix, where columns represent channels, and rows represent time slots. Cells are individual television programs. In order to utilize the high-definition resolution for different users, we apply several Focus + Context techniques to the interface. First, the columns and rows near the center of the display are slightly enlarged. Second, a strong enlargement is applied for the active program, while transparency is used to make the content under the enlarged area visible. Third, the GUI is freely zoomable from weekly overviews to extreme close-ups. Finally, overlaid animated icons give guidance and feedback for gesturing and speech input. The mobile phone GUI applies similar techniques, originally developed in different domain.

Users have full control over the media center via its speech user interface (SUI). It includes navigation in the application (e.g., "*Go to program guide*") and EPG (*"Show Monday afternoon"*), and watching media ("*Show Documentary channel*"). In EPG, it is possible to record multiple programs with a single utterance (*"Record all the Tom the Tractor shows this week*"), and to highlight programs based on their genre ("*Show me all the children programs*"). These commands were found highly appropriate among the test users in the public pilot studies and user evaluations. The SUI is implemented with context-free grammars, the vocabulary being about 900 words. There are two recognizers: an embedded recognizer running on the mobile phone, and a server-based recognizer. The choice of the recognizer is a balance between speed, vocabulary size and accuracy. Two variations of the 'push-to-talk' approach were implemented. In the first approach, the user presses a key on the mobile device to speak. The second approach combines gesture recognition with voice activity detection: the voice activity detector is activated when a user raises the phone to an upright position in front of his or her mouth.

The mobile phone keypad and gestures can be used for navigation and selections either independently or combined. Navigational gestures are made by moving the mobile phone in specific patterns. The accelerometer-based gesture recognizer supports tilting of the phone for moving the selection up and down on the screen, swinging left and right for left and right movements, swinging forward for selection, upwards for cancel, and shake to view help. Alternatively, the different orientations of the mobile phone can alter how the keypad works. In the vertical/down orientation keys are used to move selection in the EPG, in the vertical/middle orientation keys move the EPG display area, and in the horizontal orientation keys perform zooming. A combination of rule-based methods and Hidden Markov Model (HMM) based statistical model is used for gesture recognition, similar to the approach presented in [5]. The swing and shake gestures are recognized using HMMs, while the other gestures are recognized with rule-based methods.

Haptic feedback is given using the vibration component of the mobile phone. We defined a markup language to generate different kinds of haptic patterns, hapticons. Patterns are series of pulses with differences in direction of the intensity, motor, and delays between the pulses. With a set of ten control parameters (e.g., start/end intensity), it is possible to create rather sophisticated rhythmic patterns. We defined in total nine different hapticons. Recognition results, user interface actions, and the state of the application are mapped into these patterns. For example, the phone vibrates with different patterns when it recognizes a gesture, and when it receives a speech recognition result.

For visually-impaired users (who are heavy users of digital media, including digital television), the system includes a tight integration between the Focus + Context GUI and synthesized speech output. In a nutshell, the content of the active area is read out loud by the speech synthesizer when the item is activated. The spoken content is not the same as the text on display, since speech and text have different strengths and weaknesses. For example, speech outputs should use full sentences to keep the message easily comprehensible, and to allow efficient browsing, the most important information must be spoken first. The interface is based on our earlier work with mobile multimodal interfaces.

4 Pilot Studies, User Evaluations

We arranged a long-term public pilot study between June 2008 and March 2009. A mock living room with the media center was built inside the local media museum Rupriikki (Tampere, Finland). Museum visitors could freely use the system, and provide feedback using a web questionnaire. Several user studies were also conducted with museum visitors. In order to study the expectations and user experiences of different modalities, we also arranged a controlled user experiment with 26 participants in a laboratory. We used a subjective evaluation method called SUXES. The results (which are described in detail in [4]) show that the perceived quality of the speech input surpassed the upper limit of user expectations, indicating that users were very positively surprised by the performance of the speech interface. Furthermore, the low amount of out-of-vocabulary sentences shows that the restricted speech interface can be used efficiently in this domain. Users were also positive towards the application in general. In the future, we will evaluate the application in long-term pilot studies with physically and visually-impaired users. Finally, we will release the system for general public to collect feedback and usage statistics from a large amount of real users.

References

1. Ibrahim, A., Johansson, P.: Multimodal Dialogue Systems: A Case Study for Interactive TV. In: Carbonell, N., Stephanidis, C. (eds.) UI4ALL 2002. LNCS, vol. 2615, pp. 209–218. Springer, Heidelberg (2003)
2. Wittenburg, K., Lanning, T., Schwenke, D., Shubin, H., Vetro, A.: The prospects for unrestricted speech input for TV content search. In: Proceedings of the Working Conference on Advanced Visual interfaces (AVI 2006), pp. 352–359. ACM, New York (2006)
3. Soronen, H., Turunen, M., Hakulinen, J.: Voice Commands in Home Environment - a Consumer Survey. In: Proceedings of Interspeech 2008, pp. 2078–2081 (2008)
4. Turunen, M., Melto, A., Hella, J., Heimonen, T., Hakulinen, K., Mäkinen, E., Laivo, T., Soronen, H.: User Expectations and User Experience with Different Modalities in a Mobile Phone Controlled Home Entertainment System. In: Proceedings of MobileHCI 2009 (2009)
5. Schlömer, T., Poppinga, B., Henze, N., Boll, S.: Gesture Recognition with a Wii Controller. In: Proc. TEI 2008, pp. 11–14. ACM Press, New York (2008)

Comparing Gestures and Traditional Interaction Modalities on Large Displays

António Neto and Carlos Duarte

LaSIGE, Faculty of Sciences of the University of Lisbon,
Campo Grande, Lisboa, Portugal
aneto@lasige.di.fc.ul.pt, cad@di.fc.ul.pt

Abstract. Interfaces based on gesture recognition offer a simple and intuitive alternative to the use of traditional menus, keyboard and mouse. In this paper we explore the field of gestural interaction on large screen displays, presenting the results of a study where users employ gestures to perform common actions in various applications suited for large displays. Results show the actions for which gestural interaction is a best asset compared to traditional interaction.

Keywords: Gestures, Large Displays, Touch Screens, User-centered Study.

1 Introduction

The importance of gestures, despite being a central element of communication between people, has been neglected in traditional interfaces. Nowadays, with the possibilities opened up by the recent "touch revolution" and the transition we have been witnessing in the past few years to large screen displays, we are now able to explore how the use of gestural interaction can contribute to overcome the problems with the WIMP (Window, Icon, Menu, Pointer) paradigm in large screen displays.

There are two different types of research on interaction techniques for large displays. Those who seek to adapt the interaction techniques of the WIMP design paradigm to large displays and those who innovate and break from this classic paradigm. Some examples of adapted techniques are High density cursor [1], which helps users keep track of the mouse cursor, and Drag-and-pop [3], an extension of the traditional drag-and-drop. Some of the innovative techniques are Barehands [5], a free-handed interaction technique, where the user can control the invocation of system commands and tools on a touch screen by touching it with distinct hand postures, and Shadow Reaching [7], an interaction technique that makes use of perspective projection applied to a shadow representation of a user.

The field of gesture interaction has also been explored by many researchers. For example, Segen and Kumar presented Gesture VR [6], a system using vision tracking of hand gestures for spatial interaction; Cao and Balakrishnan [2] describes a gesture interface using a wand as a new input mechanism to interact with large displays.

Our current work aims at exploring the field of gestural interaction on large screen displays. For this purpose we have conducted a study where users used gestures for common actions in everyday applications, while interacting with a SMARTBoard, a

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 58–61, 2009.

large touch screen that works with a projector and a computer, and is used in face-to-face or virtual settings in education, business and government scenarios.

This paper contributions include results of gestures usage for typical actions on today's computer applications, a characterization of the applications and actions that makes them more or less suited for gestural interaction, and recommendations for gestural interfaces' designers.

In this study we used a set of gestures based in a previous study [4] that was aimed at defining a set of gestures for certain actions that can be used in non multi-touch large surface interaction scenarios.

2 Study

One of the goals of this study was to find out, in an interaction context with a single-touch large screen, which applications could benefit more from the use of gestural interaction, and which actions within those applications are able to be carried out through gestures.

Ten persons, with an average age of 24 years, participated in the study. All of the participants were regular computer users, with previous experience of the applications used in the study, and no prior experience with large displays or gestural interaction.

The experiment was conducted on a front projection SMARTBoard, with 77" (195.6 cm) active screen area, connected to a laptop PC running Windows XP SP3 with screen resolution of 1400 by 1050 pixels. The SMARTBoard supports tracking a single point in its interaction surface.

We chose four scenarios: (1) object and (2) windows manipulation, (3) image visualization, and (4) Google Earth. These everyday life scenarios are representative of settings that might benefit from the use of gestural interaction on large displays (e.g. they all can be envisioned in the context of meetings and brainstorm sessions).

Afterwards, for each scenario, we specified the actions that could be achieved by gestures, and finally, using the gestures previously defined for each action [4], implemented a gesture based interface.

Every user participated in an individual session. At the beginning of each session the purpose of the study was explained to the participants. The participants were asked to perform all actions with gestures (without being told what the "correct" gesture was) and without gestures (using the virtual keyboard and virtual mouse). All interaction was performed directly on the SMARTBoard's surface.

During each session the number of attempts and the time that the participants took to find out the right gesture for each action was measured. The timeout given to the participants to discover the gesture was one minute. After that time, if they hadn't discovered it, the gesture was exemplified.

At the end of the sessions, participants filled a questionnaire where they were asked to rate the adequateness of each action when performed with gestures and without gestures (in a 5 point scale, with 1 being the least adequate and 5 the most).

The tables bellow present, for each scenario, the action to be performed, a representation of the corresponding gesture (arrows next to the gesture indicate the direction of the movement), the average number of attempts and the average time in seconds it took participants to find out the right gesture for each action. The next two columns of the table are the questionnaire ratings, representing how adequate users

Table 1. Object Manipulation Scenario

Action	Gest	#	T	M	G	t
NewFolder	□	**2.4**	**23.3**	3.2	4.2	**3.35**
Delete	X	**1.3**	**8.2**	3.4	4.0	1.76
Copy	C	**1.2**	**8.0**	3.3	4.3	**2.74**
Paste	P	**1.1**	**6.0**	3.3	4.0	2.09
Cut		**5.3**	**55.0**	3.4	2.2	-2.57
Compress	Z	**1.3**	**8.2**	3.0	4.1	2.09

Table 2. Google Earth Scenario

Action	Gest	#	T	M	G	t
Find	┌	**1.5**	**10.9**	3.0	3.2	0.56
Placemark	P	**3.5**	**21.9**	3.4	3.8	1.31
Rotate	C	**1.1**	**6.0**	3.4	4.7	**4.99**
Zoom		**1.2**	**5.8**	3.8	4.4	1.33
Mode 3D	3	**1.9**	**9.8**	2.9	4.5	**2.95**

Table 3. Image Visualization Scenario

Action	Gest	#	T	M	G	t
Next/Prev		**1.3**	**6.9**	3.5	4.8	**3.54**
Zoom		**1.5**	**7.4**	3.5	4.5	**2.53**
Rotate	C	**1.1**	**4.4**	3.6	4.4	**3.21**
Print	P	**1.7**	**12.5**	3.9	4.1	0.56

Table 4. Windows Manipulation Scenario

Action	Gest	#	T	M	G	t
Minimize		**2.0**	**17.2**	3.6	4.5	2.21
Maximize		**2.2**	**19.2**	3.6	4.4	2.23
Close	X	**1.3**	**6.9**	3.8	3.8	0.00

feel each modality (M – mouse and keyboard, G – gesture) is to perform each action. The final column, presents the results of paired-samples t-tests comparing the adequacy of both modalities for performing the required action. Results are statistically significant for t above 2.262 (shown in bold).

3 Discussion

Based on the average time and number of attempts that participants took to find the gesture for each action, we can conclude, as expected, that these two factors are connected. The actions for which participants took less time were also the ones that needed fewer attempts to discover the right gesture (‘rotate’, ‘zoom’, ‘next/prev’, ‘copy’ and ‘paste’). For the actions participants took more time they also needed more attempts to find the right gesture (‘new folder’, ‘cut’ and ‘placemark’).

Analysing the questionnaires’ results, we can conclude that the image visualization scenario is the one with best results for gestural interaction. It can be observed that participants preferred gestural interaction when the gestures are a more direct fit and more closely related to real world interaction. On the image manipulation scenario, actions preferred are performed on the image itself as if they were done on printed photos. In the Google Earth scenario, the interactions preferred are direct, as if the user was handling non-digital objects, like an earth globe or a map.

We can verify that there is a direct relation between the factors measured in each session, time and number of attempts, and the questionnaires’ results. In general, the actions for which participants discover the gesture quicker and in fewer attempts are also the ones for which gestural interaction had best results in the questionnaires, and vice-versa.

Most of the results of this paper are in line with the results of previous studies [4]. Some of these results are concerned with the most intuitive actions to be performed using gestures and with the scenarios where gestural interaction is preferred by participants. One exception is the questionnaire results concerning the windows manipulation scenario. In previous studies participants preferred the traditional interaction to manipulate windows, which was not verified in this study. We think this difference may be due to the fact that in previous studies, participants didn't try to make these actions with traditional interaction on large screens, where despite being easily triggered by buttons, these buttons are usually small and poorly positioned in the interaction surface, which leads to gestural interaction being a better option, once it's an 'anywhere' interaction.

The classic WIMP paradigm was not originally designed for large screens or for systems without common interaction modalities such as mouse and keyboard. The use of gestural interaction does not replace these interfaces but could, if well implemented, minimize the problems and limitations introduced by their absence and improve the user interaction. One example where the direct manipulation afforded by WIMP interfaces may not be the easier option is when a user tries to delete a file. With direct manipulation the user drags the file to the trash bin. This may be the easier option when the interaction surface is small, the file is close to the trash bin, and the trash bin is not covered by a window. Performing a gesture over the iconic representation of the file is easier when these conditions are not met.

This is part of a set of studies towards the goal of a better use of gestural interaction on large screens. We plan to hold further studies with support for multiple surfaces and multi-touch technology, allowing the cooperation between multiple users on different surfaces. In the future we will develop a prototype, based on the results of this and further studies, to explore the possibilities of open cooperation between multiple surfaces through gestural interaction.

References

1. Baudisch, P., Cutrell, E., Robertson, G.: High-Density Cursor: A Visualization Technique that Helps Users Keep Track of Fast-Moving Mouse Cursors. In: Proc. INTERACT, pp. 236–243. ACM Press, New York (2003)
2. Cao, X., Balakrishnan, R.: VisionWand: Interaction Techniques for Large Displays using a Passive Wand Tracked in 3D. In: Proc. UIST, pp. 193–202. ACM Press, New York (2003)
3. Collomb, M., Hascoet, M., Baudisch, P., Lee, B.: Improving drag-and-drop on wall-size displays. In: Proc. GI, pp. 25–32. ACM Press, New York (2005)
4. Neto, A., Duarte, C.: A Study on the Use of Gestures for Large Displays. In: Proc. ICEIS. Springer, Heidelberg (2009)
5. Ringel, M., Berg, H., Jin, Y., Winograd, T.: Barehands: implement-free interaction with a wall-mounted display. In: Proc. CHI Extended Abstracts, pp. 367–368. ACM Press, New York (2001)
6. Segen, J., Kumar, S.: Look ma, no mouse! Commun. ACM 43(7), 102–109 (2000)
7. Shoemaker, G., Tang, A., Booth, K.: Shadow reaching: a new perspective on interaction for large displays. In: Proc. UIST, pp. 53–56. ACM Press, New York (2007)

Bodily Explorations in Space: Social Experience of a Multimodal Art Installation

Giulio Jacucci[1], Anna Spagnolli[2], Alessandro Chalambalakis[2], Ann Morrison[1], Lassi Liikkanen[1], Stefano Roveda[3], and Massimo Bertoncini[4]

[1] Helsinki Institute for Information Technology HIIT,
Helsinki University of Technology HUT, P.O. Box 9800, FIN-02015 TKK, Finland
name.surname@hiit.fi
[2] HTLab Dept. of General Psychology, University of Padova, V. Venezia, 8, 35131 Padova I
name.surname@unipd.it
[3] Studio Azzurro, Via Procaccini 4 C.O. La Fabbrica del Vapore , 20154, Milano, Italia
name.surname@studioazzurro.com
[4] Engineering Ingegneria Informatica, R&D Lab, Via San Martino della Battaglia 56, Roma I
name.surname@eng.it

Abstract. We contribute with an extensive field study of a public interactive art installation that applies multimodal interface technologies. The installation is part of a Theater production on Galileo Galilei and includes: projected galaxies that are generated and move according to motion of visitors changing colour depending on their voices; projected stars that configure themselves around shadows of visitors. In the study we employ emotion scales (PANAS), qualitative analysis of questionnaire answers and video-recordings. PANAS rates indicate dominantly positive feelings, further described in the subjective verbalizations as gravitating around interest, ludic pleasure and transport. Through the video analysis, we identified three phases in the interaction with the artwork (circumspection, testing, play) and two pervasive features of these phases (experience sharing and imitation), which were also found in the verbalizations. Both video and verbalisations suggest that visitor's experience and ludic pleasure are rooted in the embodied, performative interaction with the installation, and is negotiated with the other visitors.

Keywords: User studies, Art & entertainment, Multimodal interfaces, emotions.

1 Introduction

Installation art is a contemporary art form that loosely refers to a type of art-work in which the viewer is required to physically enter the work in order to experience it. Installation as a term emerged out of a series of art movements working away from creating object-centric or 'object in space' experiences for their audiences, and more towards creating experiences in spaces or purpose-made environments. Installation works stand in opposition to modern art works, where the audience had formerly been placed in a passive and non-participant role [2]. Often the themes explored or the work itself may not be considered as *art* per se by art critics, as well as the general public. Lately interactive art installations have begun to include sophisticated

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 62–75, 2009.

multimodal interfaces to track bodily interactions of visitors and their utterances. We believe these are fruitful settings to investigate in the field how such interfaces feature in public settings and how they are experienced and used in social situations.

Recently evaluation of user experience of artistic or playful interactive applications has been growing as an area of interest. This interest can also be connected to a variety of movements in HCI, such as affective computing and interfaces, gaming, and pleasurable products and interfaces. In this paper, we report a field evaluation on a public interactive art production called "Galileo all'Inferno" (Galileo in Hell). This work was performed in a public theatre environment, and the installations were then made available to the audience to experience and interpret in their own way. We utilize different methods to capture visitors' experience, with particular emphasis on understanding the nature of the users' interaction with the installation, and of the emotions that accompany it. This latter approach is motivated by the recent finding that emotions are a robust predictor of the overall enjoyment of a performance or work [18], possibly opening a perspective to interactive experiences in general that goes beyond conventional functionality. The field study aims both at evaluating the users' experience and at highlighting phenomena of interest characterizing the experience and worth further investigation.

2 Related Work

Interaction Design and HCI researchers have began working in the field of evaluating experiences with interactive art [11,13]. In 2005, Hornecker and Bruns evaluated a series of interactive installation works and noted the lack of methods to evaluate such experience dimensions as "joy of use." Gaver [7] suggested the concept of ludic engagement (or playfulness) as a way to discuss interactions with the everyday recreational use of technologies. He also proposed that these terms are useful for understanding the experience in interactive art installations [7]. As Heath et al., 2002 [10] discuss, when arriving at an interactive installation, the behaviour of those already occupying the environment sets the scene and can impact considerably on how others will act and relate to aesthetic objects in the environment. "People, in interaction with each other, constitute the sense and significance of an art work."[10, p. 11] The artist's perspective and motivation in making a playful environment for others to enjoy is also considered [23]. Of course ambiguous works often provoke unexpected and unplanned interpretations [7], in a complex interplay between the artist's intent and the visitors' response. Even though interpretation is placed in the hands of the audience as readers of the work [1], it is aware of and in various ways attempts to uncover the intention within the work. From an interaction point of view, the goals and mechanisms of an installation, may be designed with a more complex experience for its users in mind, than are conventional interactive or work-based systems. [7].

Several frameworks that can be used for studying aesthetic experiences already exist. For instance, Presence measures [9], Flow Experience [6], Dimensional models like Pleasure Arousal Dominance model (PAD) [22], Positive Affect Negative Affect Schedule [26,27], all seem feasible frameworks for the task. They also have derivates, such as the Self-Assessment Manikin (SAM) [19] which is an instrument based on the PAD model. Previous work adopting these frameworks focused on the use of questionnaires on involvement and presence to study experience of games [25], positive affect and flow experience in Internet use [3], and PAD approach to study interactive experiences [20]. However, these measures have not been extensively applied in real-world contexts to assess experiences with interactive art. Höök et. al [11] argue that the evaluation of artwork would be enriched by including an ethnographic perspective to better understand human action in-situ. Costello and Edmonds implement a method to collect experiential data with video-cued recall method to increase accuracy in recall for interview discussions [4]. Cultural probes and cultural commentators have been used in combination with ethnographic study as methods for evaluating interactive art [7,8,24]. Our methodological experimentation moves towards pluralistic evaluation approaches in allowing for divergent audiences responses from different perspectives [cf. 7,8,14,23].

3 Evaluating Installations of "Galileo all' Inferno"

Galileo all'Inferno is a theatre show developed by Studio Azzurro, it has been performed daily between 10th - 12th of July 2008 in the Teatro Arcimboldi of Milan, Italy. The show is composed of two parts, both different from an aesthetic and the technological-interactive point of view.

The first part of the show is a dance performance, during which the public attends the show in a classic way, sitting in the stalls. In the second part of the show, at the end of the performance the audience can get on the stage and interact with two interactive installations, "Galassie" ('Traces of Galaxies', Figure. 1 left) and "Ombra di stelle" (Stars' shadow', Figure. 1 right), both inspired by the writings of Albert Einstein. Our study focuses on this second, interactive part of the performance.

3.1 Description of the "Galassie" and "Ombra di stelle" Installations

"Galassie". In this installation a projector throws a beam through a transparent screen positioned on the stage. It projects a geometry of a grate of coordinates, creating a visualization of stylized shapes similar to galaxies. The software is composed of two main components: the video tracking (Retina) and the generative/reactive algorithms programmed in Processing OpenGL. The video tracking defines the position and detects the outlines of the visitors with the help of an infrared lighting system. Every person who gets on the stage generates an expanding galaxy from his body. As the user moves, (s)he's followed by his own galaxy and by a grate that visualizes persons movement in a cyclic and generative way. Moreover, by using a set of directional microphones, a component analyses acoustic features of voice based on a machine learning individualizes the emotional state of those present and influences the appearance of the galaxies. Three categories of emotions, neutral, positive or negative are detected and they modify the colour of the galaxies: a scale of grey corresponds to the neutral condition, a shade of light blue corresponds to the negative condition and a shade of red corresponds to a positive condition. Thus three semantic categories are used to send status events and the galaxies will change the dominant colour. As positive event is received all the galaxies change to "warm colours" yellow orange and red. If the status event is negative the colour ramp used by the galaxies changes to

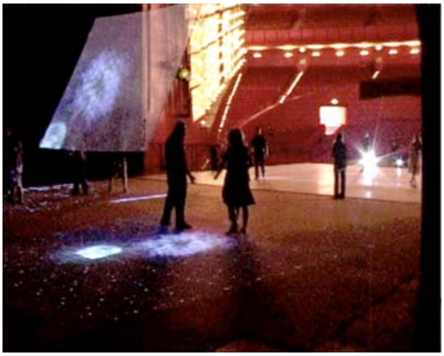


Fig. 1. Left. The art installation Galassie

Fig. 2. Right. The art installation "Ombra di stelle"

"cold colours" blue light blue violet. If the neutral event is received the galaxies turn to a grey scale. This effect will reinforce the emotional climate with introspective colours (blue Light blue) in a "negative" condition or joyful colours (orange red) in a "positive" emotional condition. The grey state should suggest the need for change stimulating reaction on the group.

"Ombra di stele". A projector transfers the image of a stellar field to a transparent vertical screen. Once passed through the screen, the beam of light is refracted and reflected, delineating stars on the stage and some other stars on the opposite side of the entrance. When the visitor gets closer, (s)he is lighted by infrared rays, creating a shadow on the ground invisible to the visitor. This shadow is detected by a camera equipped with a IR filter. The signal is analyzed by a video-tracking algorithm that identifies the shapes of the shadows throughout a sequence of coordinates. The data are elaborated by a software that reacts in real time and generates the graphics. The image of the stellar field changes depending on the graphics and the stars concentrate around the shape of the infrared shadow based on two parameters: presence and persistence. As the visitor moves, the stars move with her/him with a certain inertia. Looking at the ground (or at the backcloth), the visitor sees a constellation of stars surrounding his/her silhouette. (see Figure 1).

3.2 Approach and Method

The goal of the present work was to understand how visitors experienced Galassie and Ombra di Stelle, and to highlight specific phenomena characterizing their experience that could be subject of future work. The approach we employed to study user experience was a triangulation of three techniques:

- Quantitative data about emotions collected by administering the Positive and Negative Affect Schedule (PANAS; [26,27]).
- Written interviews to collect descriptions about the views of visitors
- Video-recording of users interacting with the art installations

We collected data on three consecutive performance days. On each day, we video-recorded users' interactions on stage. On the first day (10th July 2008) we administered the qualitative questionnaire. On the second day we administered the PANAS, and on the third day both the qualitative questionnaire and the PANAS in a balanced way. The qualitative and PANAS instruments also included a section inquiring peoples' demographic details.

The rationale of the evaluation was to contrast the user's self-reported experience against our expectations that Galassie and OdS triggered positive feelings connected to interaction and co-presence. In case of negative feelings, we wanted to identify their nature and the circumstances associated with them. The videoanalysis was aimed at integrating the outcomes of the other techniques, and to explore the nature of interaction and co-presence in the installation. Some recurrent phenomena emerging from the videoanalysis could be worth further investigation or could be used to inform subsequent evaluation approaches. In the rest of this section we described each of technique used.

The Positive and Negative Affect Schedule. The PANAS is a self-report measure of Positive Affect and Negative Affect developed by Watson and Tellegen [26,27]. It presents a two-dimensional model of affect including independent positive and negative affect dimensions. PANAS is composed by positive affect and negative affect subscales, each consisting of 10 terms. Respondents are asked to rate the extent to which they have experienced each emotion in a 5-point scale. [29]. Positive affect (PA) represents the extent to which a person feels enthusiastic, active and alert. A high PA score reflects the state of full concentration, high energy and pleasurable engagement, whereas low positive affect [26,27]. is characterized by sadness and lethargy. Negative affect (NA) is a dimension of subjective distress, in which high level of NA is described by subjective distress and unpleasurable engagement, with low NA a state of calmness and serenity [5,27]. The analysis of the PANAS data was carried out using SPSS Version 15.1 for Windows. (SPSS Inc., Chicago, IL).

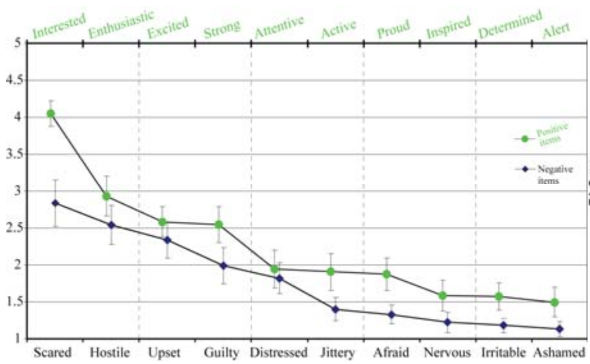
The qualitative questionnaire. The questionnaires included a range of questions, some adapted from [11]. These questions probed which emotions were raised and what provoked those emotions. We also measured dominance-submissiveness dimension - the extent to which the visitors felt they could influence the installation. Dominance was defined as a feeling of control and influence over one's surroundings and other people, versus feeling of being controlled or influenced by the situation or other people on stage [21].

Video-recordings. The video-based observations present the ethnographic approach in our methodology. This method enables understanding more fully the types of behavior—including the physical movements and the social interactions—that can occur in a particular context. Ethnographic studies are commonly conducted with new technologies to examine how people interact and which kind of emerging collaborative practices can be found. Here, we planned to video-based interaction analysis to investigate four topics a) How users discover the way to interact with the art installations, b) How they interact, c) Which are the activities that visitors are going to perform when in the same place, and d) how people orient themselves [10]. In the installation space, cameras were positioned at the rear of the stage, 90 degrees apart, pointing at each of the 2 installations on the stage and effectively capturing the whole wide area of interaction.

4 Analysis and Results

4.1 Measures of Positive and Negative Affect

A total of 115 subjects (52 male, 63 female, age 11-77) filled the PANAS questionnaire. In initial analysis, we inspected the responses to single items of the instrument. Mean values of Positive Affect items and Negative Affect items are show in Figure 3. Scores are derived from a scale 1 “very slightly or not at all”, 2 “a little”, 3 “moderately”, 4 “quite a bit” and 5 “very much”.



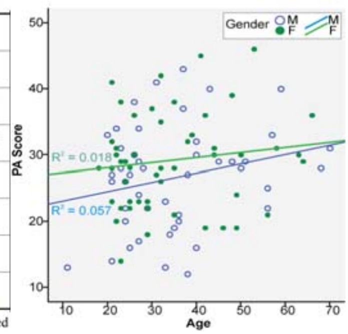


Fig. 3. Left. Means and 95 % confidence intervals for the items of positive (top) and negative (bottom) affect values. Scale 1=not at all, 5=very much.

Fig. 4. Right A scatter plot about the relation of Age and PA score. Women PA and men PA are represented by distinct markers and dedicated linear regression lines.

The majority of the items shows moderate or weak agreement, people generally agreeing more on the positive items, than the negative. From individual items we calculated the PANAS PA and NA scales and estimated their reliability. Cronbach's α was .78 for the PA scale, and .78 for the NA scale, showing that the translation of the scale into Italian language had produced a comprehensible instrument. The mean scores on the PA and NA scales were 27.90 (SD = 7.56) and 12.62 (SD = 4.14) respectively. In order to evaluate the influence of background variables on the PANAS scores, we performed two univariate analyses of variance (ANOVA). Fixed factors included gender, age and age and the interactive art installation used on the stage. Daily computer use hours and the number of interactive art installations visited in the past were used as covariates. Because of missing answers, the amount of subjects for ANOVAS was 100. ANOVA results revealed that females obtained significantly higher scores than males on the PA scale ($F = 4.124$, $p = .045$). This result is visualized in Figure 4. However, no significant gender difference was found for NA. A significant effect of age was found for PA ($F = 4.028$, $p = .048$), Figure 4 shows that increasing the age of visitors is increasing the PA score as well (regression B = .120). No significant differences were found between the two interactive art installations.

4.2 Visitors' Verbalizations

A total of 100 qualitative records were collected, 49 from users interacting with "Galassie" and 51 from users interacting with "Ombra di stelle". The data from the responses was analyzed in a bottom-up manner. The objective of this approach is to branch the descriptions of feelings in categories that result directly from the answers of the visitors. The first step was to identify terms describing a particular emotional state, following the categorization of emotional terms made by [28]. The second step was to group the descriptions in logical semantic categories. After an iterative analysis of the corpus, we identified 14 categories of feelings: interest, transport, ludic pleasure, amazement, involvement, creation, serenity, freedom, confusion, irritation, indifference, frustration, boredom, distressed. Moreover it was possible to split the 14 emotional categories in two macro-categories: Positive feelings and Negative feelings.

Table 1. Positive and negative Categories, characteristic terms and frequency in descriptions of feelings

Positive	Typical terms	Freq.
Interest	Curiosity, research, try to understand, interesting	19
Transport	In a surreal world, abducted, altered time dimension, lack of reference, sensorial isolation, "floating in a dream"	19
Ludic pleasure	Sense of play, amusing, creative game maker	11
Amazement	Astonished, surprised, amazed	5
Involvement	Participation, involvement	4
Creation	To create, to generate	4
Serenity	Peacefulness, peace, lightness	4
Freedom	Sence of freedom, "I was feeling to be on the sky, to be free"	4
Misc.	Attentive, happy	1 each
Negative	Typical terms	Freq
Confusion	Feel confused, disoriented	6
Irritation	Annoyance, irritated	4
Indifference	Not involved, not interested	4
Frustration	Frustrated, feel nitwit	3
Boredom	Feeling bug, boredom	2
Distressed	Upset	2
Misc.	Unsure, disquiet, Embarrassed, Fear, loneliness	1 each

Table 1 shows the 14 categories of feelings described by the visitors and the frequency of different terms. The semantic categories are depicted in a decreasing order of occurrence in the descriptions collected.

Positive and Negative feelings. A total of 99 visitors described their feelings. Positive feelings were described 72 times and negative 26 times. 14 subjects described more than one feeling, 13 subjects reported that they were not feeling emotions.

Perception of others. A total amount of 72 visitors answered questions about feelings perceived in others. 21 respondents described more than one feeling. Positive feelings were described 72 times, negative feelings were described 11 times. We found that Interest (23 times) and Ludic pleasure (23 times) were the feelings most often described, visitors referring to Interest with terms such as "curiosity". Ludic pleasure was described with sentences such as "they were playing like children", they were playing… who was dancing, who something else". Amazement was present 13 times in the descriptions with variable terms, for instance "surprise" and "amazed". Transport was described only 2 times. The Negative feelings that were more perceived in others were Confusion (6 times) and Indifference (3 times). 11 visitors answered that they didn't know the feelings of others.

Influencing the artworks and dominance. 54 visitors (48.8 %) described that they were able to influence the artworks and 32 (37.2 %) indicated the opposite (N=44 of "Ombra di stelle", N=42 for "Galassie"). Responses were similar for both installation, With "Ombra di stelle", 26 visitors interacting indicated being in control, With "Galassie" 27 visitors responded being able to influence the stage. Less data was acquired about Dominance. A total of 44 visitors (N=59) answered that they were not controlled by the artworks. 15 visitors felt being controlled by the system and 10 visitors explained that they felt dominated in the absence of feedback from the interactive installation.

Influence of others in the same space. Near half of all visitors acknowledged other people's influence. 31 visitors (44.3 %, N=70) described having being influenced. 10 people told that they were positively influenced by others and 3 of them explicitly mentioned imitation as adding to the positive value of their experience (e.g. "yes, observing them simplified my interaction"). 10 visitors responded to be influenced in a negative way by others, four of them stated that they would prefer a private experience.

Self-reported emotion determinants. Visitors described that Interest was provoked by curiosity, by the originality and the unfamiliarity of the artwork, and the attempt to understand the functioning of the artwork. Visitors who felt Transport used different sentences to explain what provoked that feeling: "the response to the interaction with the artworks, because of introduced a dimension changing in the silence", by the context of the particular environment, "the darkness around me and the background were producing a sense of isolation from what and who were around me ", "lights and colors", "I was feeling attracted without point of reference, like floating". Visitors explained that Ludic Pleasure was provoked by the "emotion of play", by the interaction with images that were following them and by the possibility to control the projections. Visitors described that Amazement was provoked by the surprise and by the experience of trying something that they have never done before. Visitors described that confusion was provoked by not understanding the work and the surrounding darkness. Visitors reported that Irritation was provoked by the incorrect feedback from the installations, by the scanty response of the software and the presence of other people in the same space.

4.3 Phases of Exploration towards Play

The videoanalysis, as well as the other techniques deployed in this field study, focused on two aspects of the installation, namely the interactivity and the co-presence with other visitors. The interaction analysis of people exploring the installations identified three recurrent phases: Circumspection, Testing, Play. All of them are based on the exploration of the installation affordances, showing what visitors recognized as the possible way to interact with it.

Circumspection is the phase in which the visitor is entering the interactive area, observing the current setting and selecting a point to start from.

Testing is the phase in which the visitor starts to try to interact with the artwork by making a particular bodily movement such as "moving an arm" in order to find out which movements have a consequence on the configuration of the installation. In this phase, visitors usually remain within a portion of the installation, and appropriate it by exploring and testing.

Play is the phase in which the visitor interacts with the art installation in an aware, active and involved way usually after having discovered the "working of a principle" behind the installation. In this phase, they do not just to wait for the artwork reactions, but also try to provoke those reactions by using creatively the movements that were tested in the previous phase and new ones. In the following, we will highlight two characteristics of these three phases emerged from the analysis, namely sharing the art experience and relying on imitation as a guiding principle for interaction.

Co-testing and Co-playing. The social component of the art experience is apparent from episodes that we called co-testing and co-playing. When visitors came on the stage with friends or family, they experienced the artworks by taking into account both the installation and the other people accompanying them. Entrance in the installation space, testing the artwork possibilities and finally playing with it developed as a common activity, where the users oriented both to the artwork and to their accompanying people. People in these groups tended to focus on a same portion of the installation, and to take turns into testing or playing with the artwork.

Figure 5. Circumspection

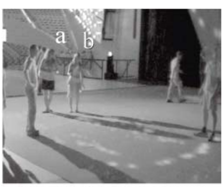


Figure 6. Circumspection 2


Figure 7. Testing 1

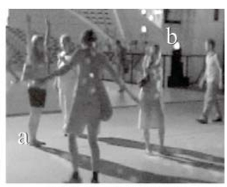


Figure 8. Testing 2

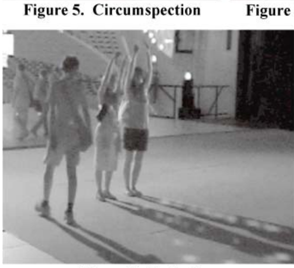


Figure 9. Co-testing

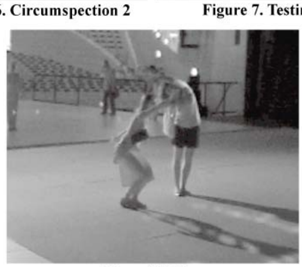


Figure 10. Play

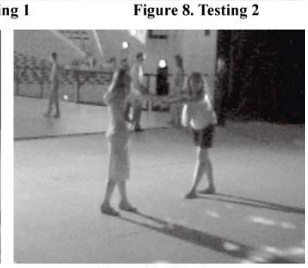


Figure 11. Co-play

We provide an example of an interaction episode in Figures 5-11. Two girls are entering the interactive area of Ombra di stelle, while others visitors are already present. Girl A is walking on the stage focusing on the transparent wall where the stars are projected, and Girl B follows her (see Figure 5). Girls A and girls B are still in the circumspection phase, moving in front of the interactive installation art and focusing on the interactive environment (Figure. 6). The circumspection phase of Girl B ends when she starts interacting by raising her arm and waiting for the response from the transparent wall (Figure 7). Girl A starts her testing phase by following the movement of the Girl B (Figure 8). Afterwards, the two girls remain positioned side by side, make joint actions or taking turns at interaction attempts (Figure 9). Girl A turns her body and raises her arms, Girl B brings down her body and moves her arms (see Figure 10). Figure 11 depicts Girl A and Girl B joining their hands, as they start interacting in an involved, creative way together, or co-playing.

Imitation. To follow other people's behaviour is a way to discover the actions that are possible in the installation and that have a consequence there. "The conduct of others within the same space can feature in how people orient, what people choose to look at and how they experience particular objects, artefacts and events" ([10], p. 23). We found that imitation was a very common strategy, especially since the installation was a little enigmatic, so that, after a while, a certain movement was adopted contagiously in different locations by people already in the installation space, and was inherited by people just entering the space, as a legitimate way to act in that context. An episode of imitation, where visitors learn how to interact with an artwork by observing others, is apparent in the interaction fragment represented by Figures 12-21.


Figure 12. Circumspection 1 (on the right)


Figure 13. Circumspection 2


Figure 14. Subject stamps the floor, girl observes


Figure 15. Girl imitation


Figure 16. Subject claps, girl and boy observe


Figure 17. Boy imitation


Figure 18. Girl testing moving her leg


Figure 19. Girl play 1, spinning


Figure 20. Girl play 4, spinning


Figure 21. Girl play 4, dancing

A young couple, a women W and a man M, enters the installation space and witnesses a girl clapping her hands, and a circle of light moving in her direction after that. This 'input' gesture is immediately adopted by the man M, and soon afterwards by the woman W, but without success. Then, they try with some feet movements, and stick to them (Figures 14-20), testing several variants and treating the floor as main interface between them and the installation.

Feet-sliding is also tried but is also soon abandoned, whereas feet-hitting-the-floor becomes the favourite movement, experimented in several variants during the play phase. The light circles projected on the floor are constantly looked for and oriented to as the active part of the environment, the one responding to their movements and validating them. Sometimes visitors acted like spectators, namely they observed other people and their behaviours without engaging in any interaction themselves.

5 Discussion and Conclusions

In this paper we presented a field study with two interactive art installations that utilise multimodal interface technologies. Using different techniques allowed to explore several different but interconnected aspects that are considered as crucial in the aesthetic experience of an interactive art installation in its emotional, physical and social aspects [13, 15, 16, 17]. The emotional dimension was considered a core part of the aesthetic experience [18], and was captured by two different techniques returning complementary and congruent outcomes. The core aspects around which the study gravitates are the visitors' interaction with the installation and with other visitors. The emotional connotations of these aspects constituted the basis for the evaluation of the experience, while the analysis of the specific actions performed with the installation and with the other visitors represented an additional resource for the evaluation and a means to explore some characteristics of the users' experience worth further investigation.

5.1 Visitors' Evaluations

The key finding of the different emotional probes is a strong tendency to experience, or at least report, positive feelings. The results of the PANAS show that the art installations evoked more positively feelings than negative ones. Interestingly, PANAS scores are not affected by the kind of art installation used, by the number of hours spent at the computer or by the experience with other artworks. It is influenced instead by age and gender.

The qualitative analysis of the questionnaires showed that a Positive interactive experience was mainly characterized by Interest, Transport and Ludic pleasure. Asking visitors to describe the feelings that they perceived in other visitors confirmed the relevance of the categories of Interest and Ludic pleasure. Visitors described to be interested, curious, involved in a creative game and to have felt Transport. When negative feelings were reported, the verbalizations suggest that they were motivated by frustration when the interaction did not develop smoothly, again showing the deep connection between aesthetic experience emotions and interaction. Merging the results of qualitative analysis and the interaction analysis we can argue that the bodily interaction with art installations has had a central role in the users experience but also in what people perceived in others who were in the same space. Visitors reporting what they perceived in others used expressions such as "they were playing like children, it is a wonderful thing", "they were playing, everyone showing his proper art, who was dancing, who something else", "they feel play".

Considering the evidence from the qualitative analysis, it seems that Ludic pleasure is related to visible, activities such as spinning, jumping and dancing, which are also able to capture the interest of others. This would also explain why the 'Ludic pleasure' category obtained higher frequency in the perception of feelings of others. What remains difficult to investigate with video observations are feelings that we summarized in the category of Transport, which seems to be a deep, subjective and interior experience. Only the qualitative questionnaire was able to shed light into this aspect of the experience. The category of Transport had frequency 19 when users were describing their emotions, but obtained only frequency 2 in the feeling perceived in others. This category seems to reflect a personal, private, and subjective experience.

Asking what provoked feelings during the interaction revealed the influence of the context and underline the important of the feedback/response of the systems, it is important to consider that not all the visitors were explaining the reason why they were feeling certain emotions/sensations in an understandable way, (Interest was provoked by "the sensation of enlarged space"). Here it did not seem reasonable to try to understand emotion in a cause-effect way. We can conclude that most of all the visitors enjoyed the experience on the stage and that they were engaged with the art installations.

5.2 Phenomena Highlighted in the Visitor's Interaction

The video-based interaction analysis allowed to explore the unfolding of the interaction with the installation and the other visitors. This more exploratory part of the field study returned some interesting insights that – besides adding some clarifications to the outcomes from the PANAS and questionnaire– represent phenomena worth been included in further studies with interactive installations.

The analysis of the video-recordings showed the existence of three main phases of interaction: circumspection, testing and play. These results are aligned with the visitors themselves describing the experience as being 'Ludic', referring to the amusing sense of play, to the emotion of play, and to the interaction with the system that was able to give them feedback. These phases show that the discovery of the action possibilities is at the core of the experience [15,16,17]. It is here we see that the participant who having learnt what the environment does, need no longer be concerned if their reading of the intention of the work is correct or not. At this stage, as well as being comfortable with responding to others in this social and performative setting, participants are free to enjoy the designed enhanced ambience and interactivity of the installation in its full force. Imitation also shows that this sequence is influenced by co-present people: users do not just study the physical environment, but the social environment as well in order to find out what action is possible. Co-testing and co-playing show that people organize their experience by including other people in the dialogue between them and the installation. Echoing the observations of a previous study [10], the social dimension revealed to be constitutive of the visitors' experience, even if the installation was not expressively meant to be experienced collectively. The presence of others was both a resource and a constraint: a resource because it helped creating a practical meaning for the space; a constraint because the space had to be conquered, configuring/marking a personal area on the floor to experience the installation.

5.3 Future Work

The results of the evaluation suggest that interaction and co-presence were the reasons at the bases of both positive feelings and negative feelings reported by visitors. The field study also allows to highlight some phenomena that can be included in future studies on interactive installations. Future evaluations could include the circumstances under which visitors give up and leave the installation, to understand what hampers the fruition of the art, or what defines the experience as completed. The equilibrium between engagement and disengagement with others in the public spaces of the exhibitions could also represent a way to evaluate the experience. Finally, other installations could be studied to check whether the same progression over three phases emerges and to go more deeply inside the practices that create a subjective meaningful place out of a spatial installation.

Acknowledgements

This work has been co-funded by the European Union (EU) project CALLAS (Co. 034800) within the 6th Framework Programme (IST).

References

1. Barthes, R.: Image/Music/Text. Trans. Stephen Heath. Noonday, New York (1977)
2. Bishop, C.: Participation. MIT Press, London (2006)
3. Chen, H.: Flow on the net–detecting Web users' positive affects and their flow states. Computers in Human Behavior 22, 221–233 (2006)
4. Costello, B., Ernest Edmonds, E.: A study in play, pleasure and interaction design. In: Proc. DPPI 2007 (2007)
5. Crawford, J.R., Henry, J.D.: The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normatives data in a large non-clinical sample. British Journal of Clinical Psychology (2004)
6. Csikszentmihalyi, M., Robinson, R.: The Art of Seeing. Pau Getty Museum, Malibu (1990)
7. Gaver, W., Boucher, A., Pennington, S., Walker, B.: Evaluating Technologies for Ludic Engagement. In: Proc. CHI 2005. ACM Press, New York (2005)
8. Gaver, W.: Cultural commentators: Non-native interpretations as resources for polyphonic assessment. Intern. Journal of Human-Computer Studies 65(4), 292–305 (2007)
9. Gemeinboeck, P., Blach, R.: Spacing the Boundary: An Exploration of Perforated Virtual Spaces. PsychNology Journal 3(1), 74–89 (2005), http://www.psychnology.org/
10. Heath, C., Luff, P., Vom Lehn, D., Hindmarsh, J., Cleverly, J.: Crafting participation: designing ecologies, configuring experience. SAGE 1(1), 9–34 (2002)
11. Höök, K., Phoebe Sengers, P., Gerd Andersson, G.: Sense and sensibility: evaluation and interactive art. In: Proc. CHI 2003 (2003)
12. Hornecker, E., Sifter, M.: Learning from Interactive Museum Installations About Interaction Design for Public Settings. In: Proc. OzCHI 2006. ACM Press, New York (2006)
13. Hornecker, E., Bruns, F.W.: Interactive Installations Analysis - Interaction Design of a Sensory Garden Event. In: Proc. ECSCW (2005)
14. Hughes, J.A., King, V., Rodden, T., Andersen, H.: Moving out from the control room: ethnography in system design. In: Furuta, R., Neuwirth, C. (eds.) Proc. CSCW 1994, pp. 429–439. ACM Press, New York (1994)
15. Jacucci, G.: Interaction as Performance. Cases of Configuring Physical Interfaces in Mixed Media. Academic Dissertation. Oulu, University of Oulu (2004)

16. Jacucci, C., Jacucci, G., Wagner, I., Psik, T.: A Manifesto for the Performative Development of Ubiquitous Media. Accepted for publication in Critical Computing, Between Sense and Sensibility, The 4th Dec. Aarhus Conference, Denmark, pp. 19 –28 (2005)
17. Jacucci, G., Wagner, I.: Performative Uses of Space in Mixed Media Environments. In: Davenport, E., Turner, S., Turner, P. (eds.) Spatiality, Spaces and Technologies, The Kluwer International Series on Computer Supported Cooperative Work, vol. 5, pp. 191–216 (2005)
18. Jobst, J., Boerner, S.: The Enjoyment in Opera – An Empirical Study of Visitors'Experience in Music Theratre. In: Proc. ICMPC 10, pp. 207–212 (2008)
19. Lang, P.J.: The Cognitive Psychophysiology of Emotion: Anxiety and the Anxiety Disorders. Lawrence Eribaum, Hillsdale (1985)
20. Laarni, J., Ravaja, N., Kallinen, K., Saari, T.: Trascendent experience in the use of computer-based media. In: Proc. Of NordCHI 2004, pp. 409–412 (2004)
21. Mehrabian, A.: Pleasure-Arousal-Dominance: a general framework for describing and measuring individual differences in temperament. Current Psychology, Winter 96 14(4) (1999)
22. Mehrabian, A., Russell, J.A.: Evidence for a Three-factor Theory of Emotions. Journal of research in Personality 11, 273–294 (1977)
23. Morrison, A., Mitchell, P., Viller, S.: Evoking Gesture in Interactive Art. In: ACM Multimedia 2008. ACM Press, New York (2008)
24. Sengers, P., Gaver, B.: Staying open to interpretation: engaging multiple meanings in design and evaluation. In: Proc. of Designing Interactive Systems, pp. 26–28 (2006)
25. Takatalo, J., Hakkinen, K.J., Sarkela, H., Nyman, G.: Involvement and Presence in Digital Gaming (2006)
26. Watson, D., Clark, L.A., Carey, G.: Positive and negative affectivity and their relation to anxiety and depressive disorders. Journal of Abnormal Psychology 97, 346–353 (1988)
27. Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect: The PANAS Scales. Journal of Personality and Social Psychology 47, 1063–1070 (1988)
28. Zammuner, V.L., Galli, C.: La conoscenza delle emozioni negli adolescenti, e in giovani adulti. Tre studi con il compito di produzione spontanea di parole. In: Matarazzo, O. (ed.) Emozioni e adolescenza,Napoli, Liguori, pp. 197–224 (2001)
29. Zevon, M.A., Tellegen, A.: The structure of mood change: An idiographic/nomothetic analysis. Journal of Personality and Social Psychology 43 (1982)

Advanced Maintenance Simulation by Means of Hand-Based Haptic Interfaces

Michele Nappi, Luca Paolino, Stefano Ricciardi, Monica Sebillo, and Giuliana Vitiello

Dipartimento di Matematica e Informatica, Università degli Studi di Salerno, 20186, Fisciano (SA), Italy

Abstract. Aerospace industry has been involved in virtual simulation for design and testing since the birth of virtual reality. Today this industry is showing a growing interest in the development of haptic-based maintenance training applications, which represent the most advanced way to simulate maintenance and repair tasks within a virtual environment by means of a visual-haptic approach. The goal is to allow the trainee to experiment the service procedures not only as a workflow reproduced at a visual level but also in terms of the kinaesthetic feedback involved with the manipulation of tools and components. This study, conducted in collaboration with aerospace industry specialists, is aimed to the development of an immersive virtual capable of immerging the trainees into a virtual environment where mechanics and technicians can perform maintenance simulation or training tasks by directly manipulating 3D virtual models of aircraft parts while perceiving force feedback through the haptic interface. The proposed system is based on ViRstperson, a virtual reality engine under development at the Italian Center for Aerospace Research (CIRA) to support engineering and technical activities such as design-time maintenance procedure validation, and maintenance training. This engine has been extended to support haptic-based interaction, enabling a more complete level of interaction, also in terms of impedance control, and thus fostering the development of haptic knowledge in the user. The user's "sense of touch" within the immersive virtual environment is simulated through an Immersion CyberForce® hand-based force-feedback device. Preliminary testing of the proposed system seems encouraging.

Keywords: Haptics, Virtual Reality, Multimodal Interfaces.

1 Introduction

In the last years the industrial world has been increasingly adopting computer-aided solutions for design-for-maintainability and maintenance training tasks with the goal to reduce development costs and to shorten time, and to improve product and service quality. Indeed, when technical systems are not designed and realized taking into account future maintenance operations, they imply high times and costs even for a simple replacement intervention [1]. For these reasons maintainability has to be considered a design added value and a competitive factor [2]. The term maintainability represents the synthesis of several characteristics that are settled on during the design

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 76–88, 2009.

phase. In fact, in this phase the critical areas to carry out maintenance tasks can be identified, the necessity of using instruments for failures diagnosis can be highlighted, and, at last, special equipment for maintenance can be devised and developed. For these reason it is clear that the design for maintainability has a positive effect on the operational costs that the user will face using the system. This can probably imply a bigger purchase price, but the final objective is to minimize the overall life cycle cost [3]. Carrying out simulations of maintenance activities within a virtual environment gives a person the ability to directly interact with 3D virtual models for maintenance purposes. Engineers can employ it to evaluate aspects of human-centered design for maintainability (accessibility, reachability, tool usability, part mount/dismountability). With respect to well-established techniques based on the intermediation of virtual characters, the first-person approach offers a more direct, intuitive control over the interaction activity, thus speeding up the maintenance checks, along with the opportunity to find out better design solutions from the maintainability point of view. Furthermore, maintenance operators (mechanics, technicians) can be trained within a highly interactive, realistic virtual reality simulator, thus combining advantages of a safe training environment with the value of the "learning by doing" practice.

To this regard, ViRstperson is a virtual reality software system developed at the Italian Center for Aerospace Research (CIRA) Virtual Reality Laboratory for carrying out digital maintenance simulations by a first-person approach [4, 5]. It is aimed at supporting engineering and technical activities such as design-time maintenance procedure validation and maintenance training. Techniques employed for improving the realism of the interactive experience include advanced lighting and shadowing to improve the user's spatial awareness within the virtual environment and a complete dynamics simulation facility which rules the interaction of all bodies within the environment, including sensor-attached anthropomorphic parts (e.g. a digital glove) [6, 7]. Both a desktop and an immersive approach can be employed by means of display systems such as auto-stereoscopic screens, mono/stereoscopic projection systems, and head-mounted displays. Computer Based Training systems created to simulate machine assembly maintenance are typically operated by means of ordinary human-computer interfaces (keyboard, mouse, etc), but this usually results in systems that are far from the real procedures, and therefore not effective in terms of training.

This study presents an extension of ViRstperson to address the simulation of the haptic feedbacks provided during the manipulation of components and equipments involved during service or maintenance operations for aerospace industry. The user-system interaction is accomplished by means of commercially-available head/hand trackers, digital gloves (CyberGlove®) and Immersion CyberForce® force-feedback system [8]. The goal is to allow techinical people to be trained to maintenance tasks, enabling them to perceive the physical characteristics of the simulated environment in a realistic way, thus improving their overall knowledge of the procedures as well as their efficiency.

This paper is organized as follows. In section 2. the relationships between haptic interfaces and system dependability are briefly exposed. In section 3. the proposed approach to visual-haptic maintenance simulation is presented in detail, while in section 4. the first experiences with the training simulator are exposed and discussed. The paper concludes in section 5.

2 Haptic Interfaces and System Dependability

When designing haptic systems, two different aspects may cause difficulties. One is related to the haptic device's functional requirements, the other is related to the associated user interface (user interface requirements). Haptic devices should be designed to guarantee dexterity (no motion constraints causing uneasiness in the extended use), human safety and portability (robust, dependable, easy to fit and adjust and be self-contained with wireless communications) [9, 10]. However, an effective use of those devices may be strongly affected by the lack of dependability in the user interface, which is the second important property haptic system designers should strive for.

Dependability of a system is commonly referred to its Reliability, its Availability and its Maintainability (RAM). However, when focusing on the user interface, there is no common agreement on what aspects of user-system interaction are related to a satisfactory RAM level of the whole system. According to [11], a dependable user interface should support speedy and accurate completion of user tasks. One of the major causes for failure with respect to accuracy and speed specifications is in fact human error. If a user makes an error while performing a task, this will propagate to the subsequent steps and result in incorrect task completion. Therefore, human computer interaction researchers strongly recommend that interface designers put special focus on error prevention, besides error recovery activities [12, 13]. Designers should investigate the most common errors users of the target domain may perform on the interface, and the most common causes for those errors, in order to understand how to prevent them, and consequently reduce the risk of mission failures in the underlying system. When dealing with haptic systems, interface dependability may become a crucial issue, especially when:

a. life-critical systems are to be manipulated, e.g., in medical and in military domains;
b. costly remote control operations are to be performed, such as product manufacturing, automotive or aerospace engineering.

Several cognitive studies on the human perceptual system related to the sense of touch, have tried to identify what human factors should be considered to connect the physical hand movements with a virtual world, as displayed, e.g., by visual interfaces or by more complex immersive environments [14,15,16]. Some of those studies focus on the tactile feedback, dealing with devices able to transmit to user's hand nerve endings heat, pressure or texture sensations. Other studies have instead considered the area of force feedback, usually adopted for devices such as robotic manipulators which interact with user's kinesthetic system (muscles, tendons and joint), applying forces as if they were generated by the virtual environment the user is operating in. The latter category also encompasses force feedback hand-based devices like the Immersion CyberForce® which has been adopted as haptic interface for the proposed simulator. Some authors already point out the importance of physical interaction issues for an accurate design of mechanism, sensors, actuators and control architecture in haptic systems [17]. They also observe that when operating with haptic devices, cognitive and physical interactions tightly depend on each other: "physical interaction can help in setting rules for cognitive evaluations of the environment, while cognitive aspects may improve the physical interaction by setting suitable control interaction parameters." If, for example, haptic interaction is used to explore the shape or the consistency of an object cognitive-based inference rules can be considered for compliance control of the device which physically interacts with humans during manipulation tasks. Conversely, the force/tactile feedback is used to convey to the user a comprehension of the physical characteristics of the virtual environment, and of the manipulated objects.

3 Haptic-Based Maintenance: The Simulation Engine

When a human has to perform a real time simulation based on the direct manual interaction approach, some parameters determine the usability of the VR system and, consequently, the reliability of the tests. Some aspects, such as the care in the simulation modelling and the visual quality, have an effect on how much realistic the user perceives the behaviour of the simulated environment, for example in terms of the reaction of the objects to the direct interaction. Even if with a more soft threshold, the parameter related to the realism influences the information obtainable from the simulation. With regard to the realism of a VR simulation, determinant aspects are basically three: the visual realism, the behaviour of the simulated world and the conveying of the sensations. In order to improve the visual realism of a virtual environment and to raise the acceptability threshold of the experience, shadows run a decisive rule. In particular, shadows are very useful to perceive the spatial relations between the objects within the scene, such as, for example, the position of the virtual hand, controlled by means of a digital glove, with regards to those objects in the scene that have to be manipulated. In order to improve the realism in the interaction the key is the correct modelling of the physical laws ruling object interaction. In this way, the virtual objects can have a reaction very similar to that one they should have in the reality, both in the direct interactions and in those deriving from the transmission of forces through collisions and/or pre-established kinematical chains.

The immersive 3D visualization system, named ViRstperson, is based on an open-source scene graph library, the Plib [18], whose main part, the SSG (Simple Scene Graph), is, in turn, based on OpenGL for real-time 3D rendering. The SSG creates and manipulates a hierarchical data structure describing the scene content in terms of nodes representing graphical objects (triangle meshes), attributes and materials, or transformations.

Aiming at developing a realistic visual and interaction virtual environment, ViRstperson has been improved with many visual and interactive functionalities, including advanced lighting, soft shadowing, stereoscopic rendering and its overall architecture is exposed in Fig. 1.

A generalizing approach to get realistic object behaviour relies on the simulation of the dynamics of objects considered as rigid bodies [19]. According to classic physic's laws, such a simulation calculates the motion of these bodies, each with assigned mass properties and shape, subjected to forces such as gravity, contacts (collisions between bodies), user-imposed, and various types of constrains. Contact forces are also treated as constraint reactions originated by the collision of two bodies. The collision detection phase, where contacts are located in, updates the world of simulated bodies with contact constraints, which subsequent simulation translates in corresponding repulsion forces.

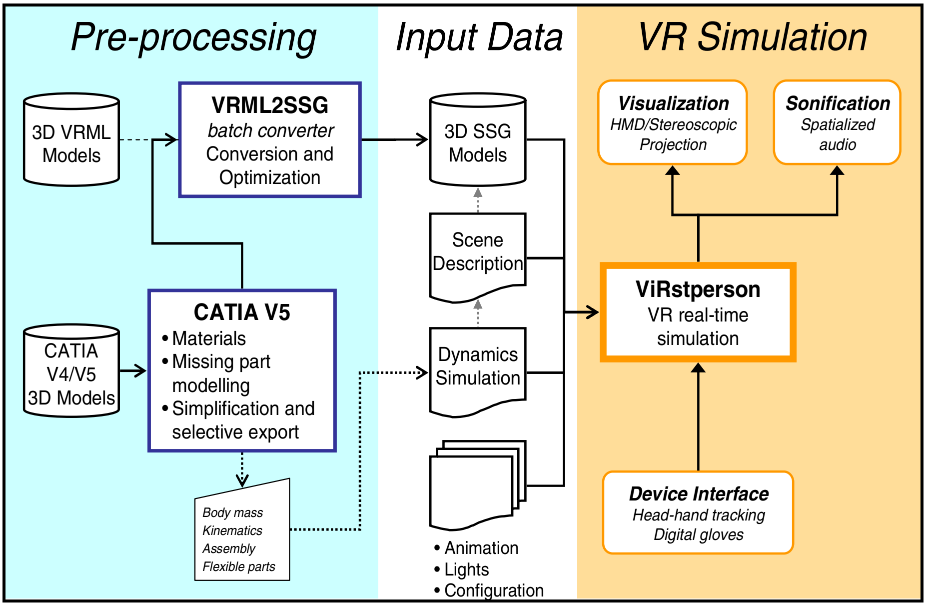
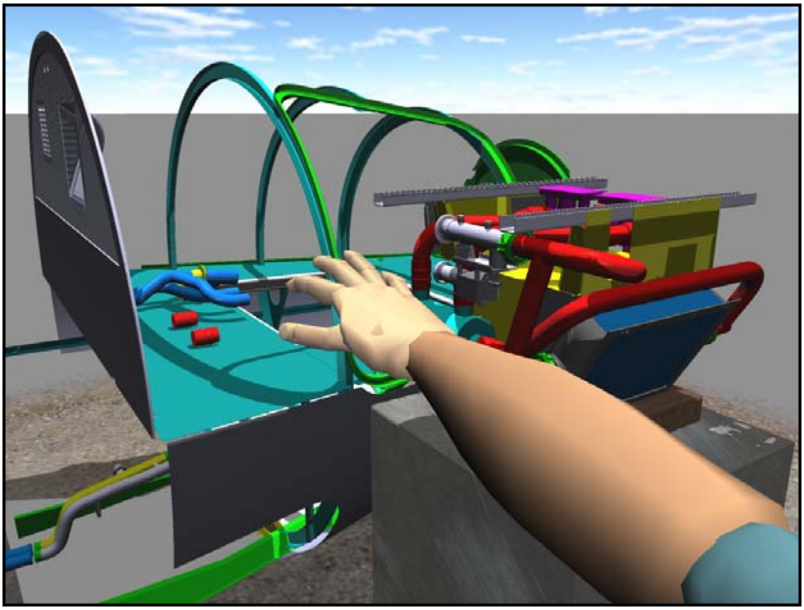


Fig. 1. A schematic view of the ViRstperson simulation engine

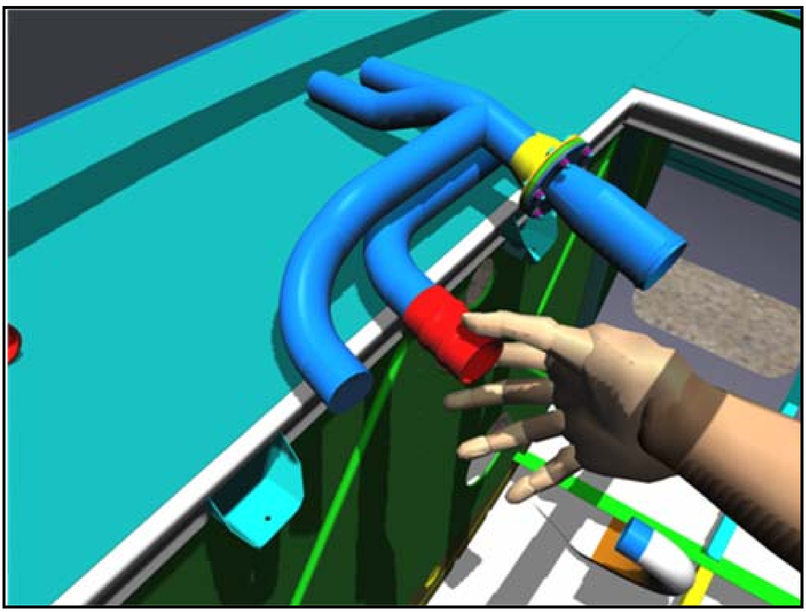
The Open Dynamics Engine (OpenDE) open-source library has been integrated in ViRstperson for adding dynamics simulation functionalities [20]. An OpenDE-based simulation is not strictly real-time, in the sense that it doesn't guarantee to update the system in a certain fixed time, yet it's sufficiently fast for interactive, man-in-the-loop application such as a VR simulation (Fig. 2a-2b). Performances may vary depending on the complexity of the virtual world (number of bodies and their geometric complexity, number of constraints): given a world, one can choose a combination of the solver type ("big matrix" or iterative) and the integration step (time) in order to find a suitable balance between accuracy and speed. For maximum computing efficiency on multiprocessor computers the dynamics simulation is implemented as a separate thread which shares the scene graph access with the visualization thread.

A further performance improvement is expected from the parallelization of the collision detection. In fact while the computing complexity of the simulation (integration) step is related to the sum of the number of bodies and joints, the complexity of the collision detection is related to the number of objects (which are either the "skin" of bodies or static parts of the dynamics world) and their geometry complexity (i.e. the number of triangles of their tri-meshes).

Normally when dealing with industrial, CAD-originated parts the collision detection occur among thousand of objects and some millions of triangles, whilst the number of bodies (i.e. the "movable" parts of the world) and the number of joints (the "behavioural" constraints between bodies) stay below two tens. Thus since the collision detection between pairs of objects is an inherently parallel task, a further threaded architecture will take advantage of recent multi-core processors with two or more processor per machine.



(a)



(b)

Fig. 2. Sample screens captured during direct first-person hand interaction in a virtual maintenance application using the ViRstperson VR engine

At the end of the whole collision detection task the resulting contacts calculated by each thread would be recollected in a single array for feeding the simulation step. Thus since the collision detection between pairs of objects is an inherently parallel task, a further threaded architecture will take advantage of recent multi-core processors with two or more processor per machine. At the end of the whole collision detection task the resulting contacts calculated by each thread would be recollected in a single array for feeding the simulation step.

3.1 Adding the Sense of Touch

The development of a haptics-assisted virtual aircraft maintenance environment is very complex because of the difficulty to simulate realistic physical processes and due to limitations of the currently available VR devices. The research activity carried out in this field focused on particular tasks of the maintenance activity such as disassembly, accessibility and manipulability assessment of geometrically-complex mechanical systems. The virtual aircraft maintenance environment reported in this paper is focused on both main industrial maintenance fields that are design for maintainability evaluation and training of maintenance operators. During the execution of maintenance tasks the contact among virtual objects, and between hand and object, has to be as much as possible compliant to dynamic behaviour which rules the interaction between physical objects in the real world.

In this context, one of the most important sensation to be transferred to a human, is the force feedback. It allows to transfer to the user's hand the sensation to touch an object, to perceive the nature of its surfaces and/or to feel the resistance that the manipulated body oppose to manipulation because of its mass properties. This feature is particularly important in VE's for maintenance training application. In fact, in the reality an operator gets used to employ tools and parts to be maintained that have mass and stiffness and the VR system has to transmit him/her the same feeling. If these factors are not taken into account during the virtual training, such training risks to be useless or having negative impact [21].

The system proposed computes contact force rendering by making its intensity dependent on the depth of penetration of the tracked hand model into the grasped object [22, 23]. The force calculated is felt by the operator through the haptic interface when the hand touches an object and/or parts come in contact during the simulation phase of the maintenance task. The penetration of objects into other objects in the scene is thus prevented by the combination of the collision detection system implemented in ViRstperson and the consequent force actuation by the force feedback device.

A CyberForce® force feedback system by Immersion Corporation has been utilized in order to provide the user with haptic sensations during the execution of maintenance tasks. The Cyberforce is composed by an articulated exoskeleton ending with a CyberGrasp system including a CyberGlove 22 sensors dataglove (see Fig. 3). The CyberGrasp provides the transmission of contact sensations to the operator's fingertips while the whole CyberForce simulates the weight and the inertia of the manipulated objects within its operative volume. The weight of the grasped object can be rendered by the CyberForce through the application of a force on the back of the user's hand. The operator can, therefore, estimate the accessibility of the tools and the possibility to remove components directly during the manipulation [24], since the

Fig. 3. The Immersion CyberForce articulated exoskeleton and the related CyberGrasp

penetration among objects gets cancelled by the reaction forces activated in relation to contacts among an object and the objects of the virtual hand [25]. Moreover the implementation of a system that allows to obtain a force feedback effect of the collisions helps the operator to notice also contacts that possibly occur outside of his field of view (e.g. contacts between elbow and objects), as long as a credible way to convey such kind of alert through the hand will be devised.

From a software point of view the integration of the CyberForce suite and ViRstperson (see Fig. 4) relies on the the product's Virtual Hand SDK for accessing device functionalities through the network [26]. On the application (client) side ViRstperson reads/writes hand-relevant data to specific shared memory segments, in order both to be general and to avoid linking with device-specific libs. External, device-specific interface programs link to the actual devices. For the force-feedback glove ViRstperson reads the sixteen 4x4 transformation matrices for the virtual hand's objects, plus the wrist position/orientation supplied by the hand tracker. It also outputs force-relevant data, which consist in one penetration depth value for each finger, plus a 3-component vector of the force vector to be actuated by the hand back force transmission arm. Then the ViRstperson CyberForce Hand Device Interface program translates such data in CyberForce-compatible force ranges for fingers and hand back actuators. The hardware configuration includes a dual 2,4 GHz Opteron based workstation equipped with a NVidia Quadro 4500 graphics board, a CyberMind Visette SXGA stereoscopic Head Mounted Display and CyberTrack hand tracking system as part of the CyberForce (see Fig. 5). The average scene load on which the system has been tested is in the range of 1 to 2 millions triangles, as the number and the geometrical complexity of many items results from the importing of CAD data.

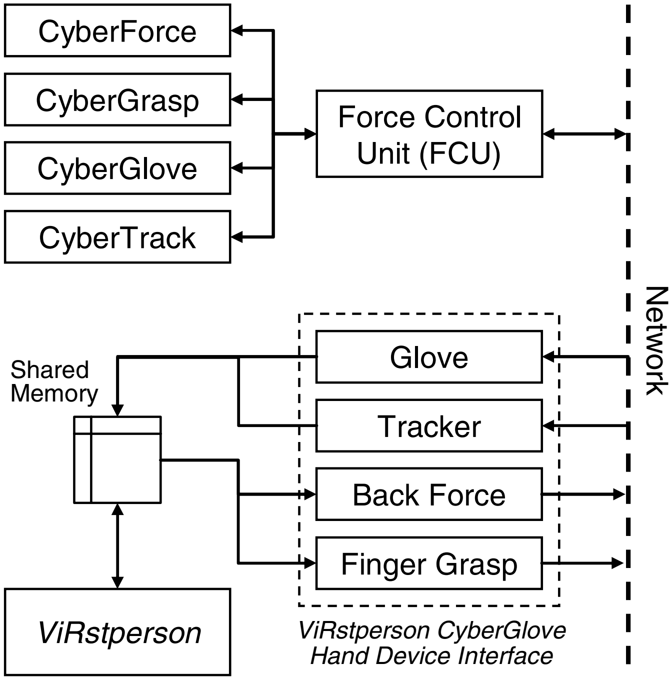


Fig. 4. Schematic view of the data channels involved in the communication between the CyberForce system and ViRstperson engine

4 First Comments on Using the Haptic-Enabled Maintenance Simulator

The simulation of maintenance tasks in a virtual environment heavily relies on the environment awareness gained by users through intuitive and natural interaction. The execution of virtual maintenance task through specific 3D interaction devices significantly affects the performance of the users, and contact force sensation can enhance the user presence in the virtual environment, and consequently his performance during the task execution [27, 28, 29]. Nevertheless the sensation of contact forces is not the only factor that affects human performance. The nature of interaction with virtual environment, the difference in perception levels between maintenance task execution in real world and in computer generated environment, human factors, ergonomic issues, and other considerations related to the effects of human computer interaction have to be taken into account for an objective evaluation of the developed interaction technique and to measure its effect on user performance in virtual maintenance tasks.

The development of a haptics-based interaction paradigm can contribute to the improvement of maintenance task execution, particularly when the manipulation of the objects requires a high level of perception and the execution of virtual maintenance tasks is not simply assembly-disassembly 3D models of parts. An important aspect of this application is the capability of the whole haptic device plus haptic rendering to

Fig. 5. Direct first-person hand interaction in the proposed virtual maintenance application using the CyberForce force feedback system

faithfully replicate the forces involved in the manipulation of technical components aboard aircraft. Indeed the mass and weight of each component as perceived by the system's user should be comparable to the real counterparts (in a range varying from hundreds of grams to several kilos). Even the operative volume can have a direct impact on the realism of the simulation if it does not allow to reach some equipments, but in this case a simple yet adequate solution could be provided by virtually shifting the haptic-device location in the virtual environment. Unfortunately, at this moment the CyberForce is one of the very few commercially available example of hand based haptic device, so its design implies some mechanical and operational limitation (Force Generation: 8.8 N max / 6.6 N min). One interesting option is the possibility to use two CyberForce arms in conjunction, to simulate two hands manipulation and coordination which could be crucial for many kind of real activities and thus indirectly doubling the forces generated during interaction. As this study is still in a early stage we can not objectively compare the "conventional" VR based training simulation to the haptics-based approach proposed.

Indeed, this will be one of the main purposes of a comprehensive evaluation study which is currently in progress, taking into consideration aspects such as usability, acceptability, level of correspondence of haptics based perception of the virtual working environment to the "sense of touch" based knowledge of the real environment.

Another usability factor which may directly affect dependability of a haptic system is robustness. Robustness is defined as the support the system gives users to let them figure out if they have achieved their goals, and how well [30]. Considering the type of certain critical haptic interaction tasks, a sufficient degree of robustness is needed to safely and effectively achieve the initial interaction goals. Several features can be considered important to achieve robustness when dealing with a haptic interfaces. These are again related to both the haptic and the visual/audio input/output modes that the system is adopting. In particular, user's ability to evaluate the internal state of the system from its perceivable representation, and the persistence of the effect of an action on the interface (e.g., the weight of the manipulated rigid body perceived as long as it is held by the cyberglove). Also forward/backward recovery actions should be allowed as far as possible during task performance. Anyway, preliminary testing conducted on the virtual model of an avionic system performing simple to medium complexity manipulation tasks (e.g. to inspect equipments, to mount/unmount components, to replace items, etc.) with and without the haptic interface resulted in an explicit preference for the haptics-assisted modality encouraging the authors to proceed in this research.

5 Concluding Remarks

In this paper a framework for haptics-assisted VR based aircraft maintenance training has been presented. The system integrates a hand based force-feedback haptic device with a real-time VR simulation engine supporting realistic interaction through dynamics simulation. Although an extensive usability study has not been completed yet, the first impressions of prospective users seem to confirm the added value of haptic-based interaction compared to an exclusively visual-based training simulation. Moreover, the use of a hand-based haptic device versus a stylus-based end effector has been very favourably acknowledged by technical personnel already accustomed to VR-based training procedures, proving that for real world operations at least a force-feedback on the whole hand is required. Future work will include an extensive experimentation aimed to measure the advantages of the haptic-based approach applied to virtual training and the extension of the framework to a two-hands system which could bring even more realism to the interaction.

Acknowledgements

We gratefully acknowledge the contribution of the Italian Aerospace Research Center (CIRA) to the development of this research.

References

1. Hubka, V.: Scienza della progettazione. In: Masson Editoriale ESA, pp. 23–56 (1994)
2. Ivory, C.J., Thwaites, A., Vaughan, R.: Design for maintainability: the innovation process in long term engineering projects. In: Proc. of the Conference The future of innovation studies, Eindhoven, Netherlands (2001)

3. Serger, J.K.: Reliability Investment and Life-Cicle Cost. IEEE Transaction on Reliability 32, 259–263 (1983)
4. Liang, J., Shaw, C., Green, M.: On temporal-spatial realism in the virtual reality environment. In: proceedings of the ACM Symposium on User Interface Software and Technology, pp. 19–25 (1991)
5. Ware, C., Rose, J.: Rotating virtual objects with real handles. ACM Transactions on CHI 6(2), 162–180 (1999)
6. Rezzonico, S., Huang, Z., Boulic, R., Magnenat Thalmann, N., Thalmann, D.: Consistent Grasping Interactions with Virtual Actors Based on the Multi-sensor Hand Model. In: Virtual Environments 1995, pp. 107–118. Springer, Wien (1995)
7. Poupyrev, I., et al.: Go-Go Interaction Technique: Non - Linear Mapping for Direct Manipulation in VR. In: ACM UIST 1996, pp. 79–80 (1996)
8. Immersion Corp 3D Interaction Hardware and Software Products, `http://www.immersion.com/3d/`
9. Biggs, S.J., Srinivasan, M.A.: Haptic Interfaces. In: Stanney, K.M. (ed.) Handbook of Virtual Environments: Design, Implementation, and Applications, pp. 93–115. Lawrence Erlbaum, Mahwah (2002)
10. Hale, K.S., Stanney, K.M.: Deriving Haptic Design Guidelines from Human Physiological, Psychophysical, and Neurological Foundations. IEEE Computer Graphics and Applications 24(2), 33–39 (2004)
11. Reeder, R.W., Maxion, R.A.: User Interface Dependability through Goal-Error Prevention. In: Proceedings of the 2005 International Conference on Dependable Systems and Networks, DSN 2005 (2005)
12. Nielsen, J.: Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods. John Wiley & Sons, New York (1994)
13. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, vol. 4/E. Addison-Wesley, Reading (2005)
14. O'Malley, M., Goldfarb, M.: The Implications of Surface Stiffness for Size Identification and Perceived Surface Hardness in Haptic Interfaces. In: Proc. IEEE Int'l Conf. Robotics and Automation, pp. 1255–1260 (2002)
15. Sherrick, C.E., Cholewiak, R.W.: Cutaneous Sensitivity. In: Boff, K., Kaufman, L., Thomas, J. (eds.) Handbook of Perception and Human Performance v.1: Sensory Processes and Perception. John Wiley & Sons, Chichester (1986)
16. Tan, H.Z., et al.: Human Factors for the Design of Force-Reflecting Haptic Interfaces. Am. Soc. of Mechanical Engineers DSC 55-1 (1994)
17. De Santis, A., Sicilian, B., De Luca, A., Bicchi, A.: An atlas of physical human–robot interaction. Journal of Mechanism and Machine Theory 43, 253–270 (2008)
18. Baker, S.J.: PLIB: A Suite of Portable Game Libraries, `http://plib.sourceforge.net`
19. Garbaya, S., Zaldivar-Colado, U.: The affect of contact force sensations on user performance in virtual assembly tasks. Journal of Virtual Reality 11(4) (October 2007)
20. Smith, R.: ODE: Open Dynamics Engine, `http://ode.org`
21. Liang, J., Shaw, C., Green, M.: On temporal-spatial realism in the virtual reality environment. In: Proceedings of the ACM Symposium on User Interface Software and Technology, pp. 19–25 (1991)
22. Borst, C.W., Indugula, P.: Realistic virtual grasping. In: Proceedings IEEE virtual reality conference (VR 2005). Bonn, Germany, vol. 320, pp. 91–98 (2005)
23. Borst, C.W., Indugula, A.P.: A spring model for whole-hand virtual grasping. Presence Teleoperators Virtual Environ 15(1), 47–61 (2006)

24. Chryssolouris, G., Mavrikios, D., Fragos, D., Karabatsou, V.: A virtual reality-based experimentation environment for the verification of human-related factors in assembly processes. Robot Comput. Integr. Manuf. 16(4), 267–276 (2000)
25. Colgate, J.E., Grafing, P.E., Stanley, M.C., Schenkel, G.: Implementation of stiff virtual walls in force-reflecting interfaces. In: Proceedings IEEE virtual reality annual international symposium (VRAIS), Seattle, WA, pp. 202–208 (1993)
26. Carmel, R., Ullrich, C., Silver, J.: VirtualHand v2.5 programmer's guide. Technical Report, Virtual Technologies Carpenter ID, Dewar RG, Ritchie JM, Simmons JEL (1996) (2001)
27. Zachmann, G., Rettig, A.: Natural and robust interaction in virtual assembly simulation. In: Proceedings 8th ISPE international conference on concurrent engineering: research and applications, Anaheim, CA (2001)
28. Zorriassatine, F., Wykes, R., Parkin, R., Gindy, N.: A survey of virtual prototyping techniques for mechanical product development. Proceedings Inst. Mech. Engineers. Part B: J. Eng. Manuf. 217(4), 513–530 (2003)
29. Garbaya, S., Coiffet, P., Blazevic, P.: Integrating computerised assembly planning with virtual assembly environment. In: Proceedings virtual reality in mechanical and production engineering (VR- Mech 2001), Brussels, Belgium (2001)
30. Dix, A., Finlay, J., Abowd, G., Beale, R.: Human-Computer Interaction, 3rd edn. Prentice-Hall, Englewood Cliffs (2004)

Multimodal Interaction within Ambient Environments: An Exploratory Study

Yacine Bellik[1], Issam Rebaï[1], Edyta Machrouh[1], Yasmin Barzaj[1],
Christophe Jacquet[2], Gaëtan Pruvost[1], and Jean-Paul Sansonnet[1]

[1] LIMSI-CNRS, B.P. 133, 91403, Orsay cedex, France
{yacine.bellik,issam.rebai,edyta.machrouh,yasmin.barzaj,
gaetan.pruvost,jps}@limsi.fr
[2] Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France
Christophe.jacquet@supelec.fr

Abstract. Inputs and outputs are not two independent phenomena in multimodal systems. This paper examines the relationship that exists between them. We present the results of a Wizard of Oz experiment which shows that output modalities used by the system have an influence on the users' input modalities for a large category of users. The experiment took place in a smart room. This kind of environment does not require any particular knowledge about computers and their use and thus allowed us to study the behavior of ordinary people including subjects who are not familiar with computers. The experiment also shows that speech is a favorite modality within smart room environments for a large part of users. We think that the results presented in this paper will be useful for the design of intelligent multimodal systems.

Keywords: Multimodal, Interaction, Wizard of Oz, Modality, Ambient Environment, Ordinary People.

1 Introduction

Multimodal interfaces have been extensively studied for several years. At the beginning, most of the studies addressed the input side of multimodal interfaces, mainly trying to solve the problem of information fusion, to structure the design space of multimodal interfaces and to study the multimodal behavior of users. More recently, researchers have been more interested in the output side. Different issues have been addressed such as the intelligent presentation of information and the adaptation of output multimodal interfaces. However, the input and the output sides of multimodal interaction are not independent phenomena and we need to know more about the relationship that exists between them and how each side may affect the other side. In this paper, we focus on the influence of system output modalities on user input modalities in an ambient environment. We present a Wizard of Oz (WOZ) experiment which took place inside a smart room. This kind of environments allows one to interact with familiar everyday-life objects such as lights, windows, etc. Furthermore, ambient environments do not require particular knowledge about computers and thus are well adapted to study the behavior of ordinary people who do not have particular knowledge about computers.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 89–92, 2009.

2 Related Work

Previous work showed the benefits of using multimodal interaction within ambient environments [1]. However, related research work mainly addressed the study of multimodal user behavior in order to evaluate the usefulness of multimodality and to extract patterns allowing the development of more robust multimodal systems.

In [2], the authors showed how speech and gesture were combined by users. They showed that the task can influence users' choice of input modalities. Tasks without spatial components were almost always completed by users using speech only, whereas those with spatial components were performed using multimodal combinations. In [3] the authors studied the stability of multimodal patterns depending on the user's age. The patterns of children presented many similarities with those of adults even though children follow a simultaneous multimodal pattern more often. In [4], the authors showed that users switch modes under certain contexts. For instance, when recognition errors occur, users will shift from one mode to another in order to recover. In another interesting work, Oviatt & al. [5] focused their study on the speech modality. They showed how users adapt their speech signal input to converge with a text-to-speech output system. However this study is mainly concerned with how users adapt the attributes of a given modality (speech) rather than how users are influenced when choosing between different input modalities. To sum up, previous works showed that different factors may influence the users' choice of input modalities. Task, context, users' age and errors have been the most studied parameters. In this paper, we focus on system output modalities. In our study we preserve all the previous parameters from any changes and we vary only the system output modality (tasks follow the same schema, context does not change, all users are adults and error conditions are minimized).

3 Experiment

We conducted a Wizard of Oz experiment to compare the input modalities used by the subjects when the output modality of the system changes. The experiment took place inside a smart room in our lab. This smart room is a testbed for "ambient intelligent" environments, in which people can interact with the ambient system in a natural way, through various modalities. Fifteen unpaid adult volunteers, 5 males and 10 females, aged 35 years on average, served in the experiment. There were three different output modalities for the system: text, graphics (icons) and speech synthesis. For the user, three input modalities were available: speech, pointing gestures on a touch screen and button presses on a remote control. In the remainder of this paper we will refer to these three modalities, respectively by speech, touch screen and remote control (even though the touch screen and the remote control are devices, not modalities). Speech recognition and the detection of touches on the screen were simulated by the operator of the WOZ system (who was located in another room). Speech recognition was simulated because we didn't want to use an intrusive device (microphone) and we wanted to minimize recognition errors without constraining users to a limited vocabulary. We induced the users into believing that microphones were embedded into the room walls and that the screen was a touch screen. These three modalities were chosen because of the ambient context where the experiment took place. For instance, using a keyboard to switch the light on would not have been very relevant to a home living setting. For a first study, the issue of combined modalities is not considered here. Studying the influence of system output modalities on user input *combined* modalities, is planned as one of our future experiments.

Therefore we defined simple tasks to allow the subjects to perform them easily by using any single modality chosen among the three available ones. The subjects had to perform six tasks: switch the light on, listen to music, increase the sound level, decrease the light level, stop the music and switch the light off. As said before, the studied factor is the influence of system output modalities. Any other factor such as task complexity has to be kept static. Hence, every task has the same 5-steps structure. Each task begins by a task presentation step (1). During this step and to avoid influencing the subject with a particular modality, the system introduces the task using all output modalities (Text + Graphics + Speech) in a redundant way. Then we observe which input modality the subject uses to start the task (2). Then the system asks a question (3) to get more details about the task parameters and proposes two possible answers. Now, the system uses only one modality (speech, text or graphics). Contrary to step (1), this step of the experiment aims at trying to influence the user with a particular modality. We then observe which input modality is used by the subject to answer the question (4). Finally and after getting the answer from the subject, the system performs the action (5) (switches the light on, plays music, etc.).

We predicted that the output modality used by the system would have an influence on the input modalities used by the subject. In other words, we predicted that the subject would use different input modalities depending on system output modality.

4 Results

We analyzed the input modalities used by the subjects depending on system output modalities (Fig. 1). The results show that speech input is a favorite modality within

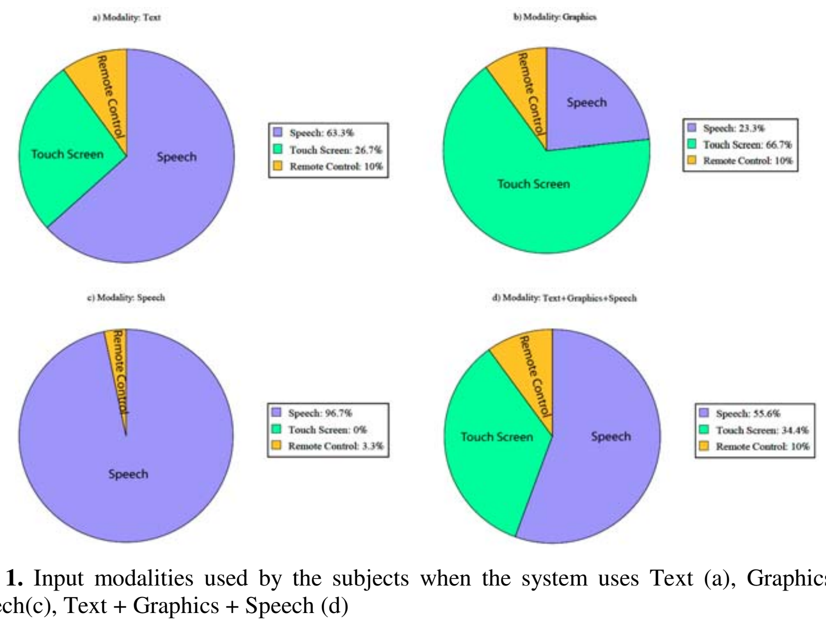


Fig. 1. Input modalities used by the subjects when the system uses Text (a), Graphics (b), Speech(c), Text + Graphics + Speech (d)

smart room environments for a large part of users, except when graphics modality is used by the system. In this case, the touch screen is preferred to speech. These results are confirmed by the answers to the questionnaires. Moreover we can observe a strong relation, on the one hand, between verbal modalities, and on the other hand between non-verbal modalities. When the system uses text or speech, users tend to use speech. However, when the system uses graphics then users tend to use pointing gestures on the touch screen. Finally, we can also observe that when the 3 output modalities are used simultaneously then speech is the preferred modality but with a lower percentage than when text or speech synthesis are used alone. It seems that each modality has a kind of *influence power* and when the modalities are combined their influence powers are combined too. This is a interesting phenomenon which needs to be investigated and confirmed in further experiments.

5 Conclusion and Future Work

We have developed a smart room and WOZ platform to study Human-System interaction within ambient environments. Our first experiment concerned the multimodal interaction within such environments. This experiment allowed us to study the relation between system output modalities and user input modalities. It shows that system output modalities have an important influence on user input modalities. The experiment shows also that speech is a favorite modality within smart room environments for a large part of users, except when graphics modality is used by the system. We think that these results will be useful for the designers of intelligent multimodal systems. In future work we will build another WOZ experiment where the tasks can be performed either by using single modalities or combined modalities so that we can study if the monomodal-multimodal system behavior may influence the mono/multimodal user behavior. This would allow us to formalize the notion of *influence power* of a modality and study the influence of combined modalities to see if there is a law which explains how these influence powers are combined.

References

1. Burzagli, L., Emiliani, P.L., Gabbanini, F.: Ambient Intelligence and Multimodality. In: Proc. of UAHCI 2007 held as part of HCI International 2007, part II, pp. 33–42 (2007)
2. Oviatt, S.L., De Angeli, A., Kuhn, K.: Integration and Synchronization of Input Modes during Multimodal Human-Computer Interaction. In: Proc. of CHI 1997, pp. 415–422 (1997)
3. Xiao, B., Girand, C., Oviatt, S.L.: Multimodal Integration Patterns in Children. In: Proc. of ICSLP 2002, pp. 629–632 (2002)
4. Oviatt, S.L., Bernard, J., Levow, G.: Linguistic adaptation during error resolution with spoken and multimodal systems. Language and Speech, Special Issue on Prosody and Conversation 41(3-4), 415–438 (1999)
5. Oviatt, S.L., Darves, C., Coulston, R.: Toward Adaptive Conversational Interfaces: Modeling Speech Convergence with Animated Personas. Transactions on Human Computer Interaction, Special Issue on Mobile and Adaptive Conversational Interfaces 11(3), 300–328 (2004)

Multimodal Interaction: Intuitive, Robust, and Preferred?

Anja B. Naumann[1], Ina Wechsung[1], and Jörn Hurtienne[2]

[1] Deutsche Telekom Laboratories, Berlin Institute of Technology, Ernst-Reuter-Platz 7, 10587 Berlin, Germany
{anja.naumann,ina.wechsung}@telekom.de
[2] Chair of Human-Machine Systems, Berlin Institute of Technology, Franklinstr. 28-29, 10587 Berlin, Germany
hurtienne@acm.org

Abstract. We investigated if and under which conditions multimodal interfaces (*touch*, *speech*, *motion control*) fulfil the expectation of being superior to unimodal interfaces. The results show that the possibility of multimodal interaction with a handheld mobile device turned out to be more intuitive, more robust, and more preferred than the interaction with the individual modalities *speech* and *motion control*. However, it was not clearly superior to *touch*.

Keywords: multimodal interfaces, intuitive use, usability.

1 Introduction

Multimodal interfaces allow for using multiple communication channels simultaneously or one after another [1] which is assumed to be analogue to natural human-to-human communication [2]. This led to the expectation of multimodal systems enabling flexible, robust, efficient, and intuitive interaction [1]. However, there are results both supporting this assumption [e.g. 3] and showing the contrary [e.g. 4]. In this study we investigate if multimodal interaction really is better than interacting with single modalities in a handheld mobile device (*touch*, *speech*, *motion control*) and how different modality conditions support intuitive interaction (defined as the users' effective and subconscious application of prior knowledge [5]) and robust interaction (i.e. fast and error free interaction). In addition, we want to find out if multimodality is utilized fully and if modalities which are preferred are also used more often.

2 Method

30 participants (15 male/15 female) aged between 22 and 78 years took part in the study. The device used for testing was a smartphone (HTC Touch Diamond; IBM embedded ViaVoice) controllable via *motion* (tilt and twist), *speech* and *touch* input. System output was graphical for all modalities. For motion control, additional tactile feedback and for speech control, additional auditive feedback was given. The application tested was a multimodal mailbox system capable of handling speech-, e-mail- and fax-messages, call forwarding, and notifications of mailbox messages.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 93–96, 2009.

For measuring the consequences of intuitive use according to [5] we used a questionnaire with the subscales: *Perceived Cognitive Load* (6 items, e.g. "The use of the system was not complicated."), *Perceived Task Completion*: (6 items, e.g. "I was able to reach all my goals with the system."), *Perceived Error Rate*: (3 items, e.g. "No problems occurred while using the system."), *Perceived Effort of Learning*: (6 items, e.g. "I immediately knew how to operate the system."), *Familiarity*: (6 items, e.g. "The operation of the system always met my expectations.", *Global*: (1 item: "The use of the system was intuitive."). Items were answered on a five-point Likert scale (0= strongly disagree, 4= strongly agree). All results show scale means. Higher values indicate a higher probability for intuitive use of the system tested. For measuring the robustness, performance data (successful task completion and aborts of task execution) were logged while the participant was dealing with the tasks. Also, the task duration and the chosen modality were recorded. The participants had to execute 4 blocks of tasks with a total of 14 tasks (get messages, reply to them, forward, and sort messages as well as changing notification options). The tasks were chosen to equally favour the different modalities. If the task goal was not achieved within three trials, task execution was aborted and the next task started. First, participants were asked to solve all tasks with a given modality. Afterwards, participants evaluated the interaction via the questionnaire. This was repeated for all three modalities, the sequence of the modalities was balanced between the participants. In the final block, participants were free to choose the modalities they used for solving the task. In this condition it was always possible to switch or to combine modalities. Again, the participants evaluated the interaction after solving all tasks in this condition.

3 Results

Questionnaire for measuring intuitive use: An ANOVA showed significant differences between the four test blocks. For most subscales, the lowest value and thus the worst rating was given to *motion control*. The best rating was always given to the *multimodal* condition (see Figure 1). As post hoc tests show, the *multimodal* test block was better rated than the interaction in all three unimodal blocks on the scales *Perceived Cognitive Load*, $F(3, 81)=12.27$, $p=.000$, part. $eta^2=.31$, *Perceived Effort of Learning*, $F(2.37, 61.74)=10.16$, $p=.000$, part. $eta^2=.25$, and *Global*, $F(2.20, 57.16)=16.34$, $p=.000$, part. $eta^2=.39$. Regarding the scales *Perceived Task Completion*, $F(3, 81)=21.59$, $p=.000$, part. $eta^2=.44$, *Perceived Error Rate*, $F(3, 78)=14.95$, $p=.000$, part. $eta^2=.37$, and *Familiarity*, $F(3, 78)=22.81$, $p=.000$, part. $eta^2=.47$, the test blocks *speech* and *motion control* were rated worse than the *touch* and the *multimodal* test condition. *Touch* was rated as good as *multimodal*.

Robustness (successful first trials and aborts): The participants had more successful first trials in the *multimodal* and in the *touch* condition, $F(3, 81)=26.58$, $p=.000$, part. $eta^2 =.50$. In both conditions, more than 60 percent of the tasks were solved in the first trial (see Figure 2).

Furthermore, in the *multimodal* condition, less tasks were aborted than in all other conditions, $F(2.25, 60.67)=20.18$, $p=.000$, part. $eta^2 =.43$ (see Figure 2).

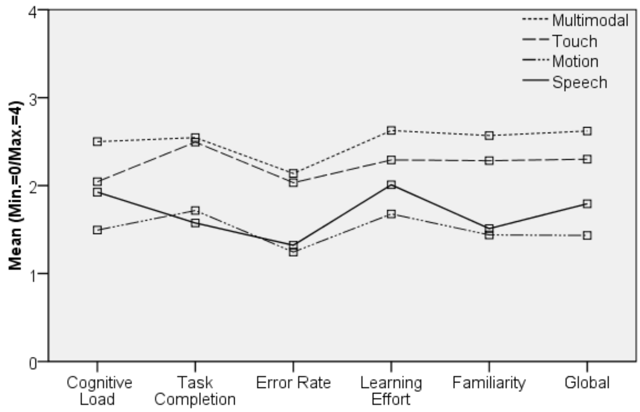


Fig. 1. Results on all subscales of the questionnaire measuring intuitive use by modality

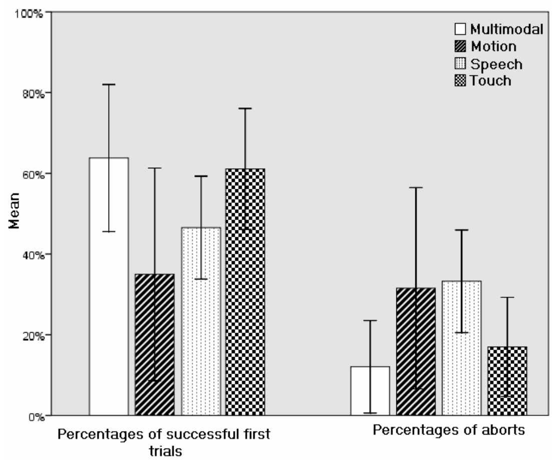


Fig. 2. Percentages of tasks successfully completed in the first trial and percentages of tasks aborted. Error bars display one standard deviation.

Task duration: The task duration varied between the four conditions, $F(2.29, 61.94)=76.42$, $p=.000$, part.eta²=.74. Most time was needed for *motion control* and least for the *multimodal* condition.

Frequency of modality usage: In the *multimodal* condition, *touch* was the modality most preferred by the participants, $F(1.35, 36.41)=28.42$, $p=.000$, part. eta²=.51. In 68 % of the tasks, *touch* was the first choice. For 19 % of the tasks, the participants first chose *speech* input. For 7 % of the tasks *motion control* was used. For only 6 % of the tasks a combination of the different modalities was the first choice (i.e. solving different steps of the task with different modalities). In case the participants were not successful they could try to solve the task for a second and a third time. Also in the second trial, there are differences regarding the preference of modalities,

$F(1.76, 47.65)=13.57$, $p=.000$, part. $eta^2=.33$. Again, most users preferred *touch* (59%), followed by *speech* (20%), a combination of modalities (16%), and *motion control* (5%). In the third trial, 52% of the users preferred *touch*, 21% preferred a combination of modalities, 17% *speech*, and 9% *motion control*, $F(3,69)=5.75$, $p=.001$, part. $eta^2=.20$.

Modality switches: Only few participants used a flexible interaction strategy. In case they were not successful with a task in the first trial, only 21% of the participants switched to another modality in the second trial, $\chi^2(4, N=147)=323.4$, $p=.000$. Even in the third trial, 73% of the participants kept using the modality they had chosen before. Of the remaining participants most switched to a combination of modalities in both the second (12%) and the third trial (11%).

4 Discussion and Conclusion

Overall, the *multimodal* condition always showed results at least as good as all unimodal conditions regarding performance and subjective ratings. It was rated as more intuitive and less straining than *speech* and *motion control*. Also, the percentage of successful first trials was higher and the percentage of aborts lower than for *speech* and *motion control*. The results show that the possibility of multimodal interaction turned out to be more intuitive, more robust, and more preferred than the interaction in the individual modalities *speech* and *motion control*. However, even if a combination of modalities was possible in the last test block, most participants stayed with *touch*. Thus, for the subjective ratings, there were only small differences shown between the *touch* and the *multimodal* condition. Switching to another modality should pay off the most when errors are made during the first trial solving a task [1]. However, in our study, participants did not switch the modality. They did not utilize the advantage of multimodality for compensating usability problems in other modalities. A long term observation of modality usage and switches is planned in order to get more insight to this question.

References

1. Oviatt, S.: Ten myths of multimodal interaction. Comm. of the ACM 42(11), 74–81 (1999)
2. Chen, F.: Multimodal Interaction Systems. In: Chen, F. (ed.) Designing Human Interface in Speech Technology, pp. 212–223. Springer, Berlin (2006)
3. Sarter, N.B., Woods, D.D.: Teamplay with a powerful and independent agent: A full-mission simulation study. Human Factors 42(3), 390–402 (2000)
4. Wechsung, I., Naumann, A.B.: Evaluation Methods for Multimodal Systems: A Comparison of Standardized Usability Questionnaires. In: André, E., Dybkjær, L., Minker, W., Neumann, H., Pieraccini, R., Weber, M. (eds.) PIT 2008. LNCS (LNAI), vol. 5078, pp. 276–284. Springer, Heidelberg (2008)
5. Naumann, A., Hurtienne, J., Israel, J.H., Mohs, C., Kindsmüller, M.C., Meyer, H.A., Husslein, S.: Intuitive Use of User Interfaces: Defining a Vague Concept. In: Harris, D. (ed.) HCII 2007 and EPCE 2007. LNCS (LNAI), vol. 4562, pp. 128–136. Springer, Heidelberg (2007)

Sharing Map Annotations in Small Groups: X Marks the Spot

Ben Congleton[1,*], Jacqueline Cerretani[1,*], Mark W. Newman[1,2], and Mark S. Ackerman[1,2]

[1] School of Information
[2] Department of Electrical Engineering and Computer Science
University of Michigan
1075 Beal Avenue
Ann Arbor, MI 48109, USA
{bcx,jackiejc,mwnewman,ackerm}@umich.edu

Abstract. Advances in location-sensing technology, coupled with an increasingly pervasive wireless Internet, have made it possible (and increasingly easy) to access and share information with context of one's geospatial location. We conducted a four-phase study, with 27 students, to explore the practices surrounding the creation, interpretation and sharing of map annotations in specific social contexts. We found that annotation authors consider multiple factors when deciding how to annotate maps, including the perceived utility to the audience and how their contributions will reflect on the image they project to others. Consumers of annotations value the novelty of information, but must be convinced of the author's credibility. In this paper we describe our study, present the results, and discuss implications for the design of software for sharing map annotations.

Keywords: Map Annotation, Social Media, Sharing, Location Based Computing, HCI, Requirements Analysis.

1 Introduction

The boundary between "cyberspace" and physical space is eroding. Advances in location-sensing technology (e.g., GPS), coupled with an increasingly pervasive wireless Internet, have created a new class of applications where users create, share, and access location-based information linked seamlessly to their physical context.

Traditionally, when we think of connecting information to a geospatial location, we think in terms of maps. Although maps have been used for thousands of years, only recently have end users been able to create and annotate geo-spatially indexed information to easily share with others. Google Earth, Google Maps, Yahoo Maps, and Microsoft Live Maps allow one to render geospatial annotations, and also provide tools to create and add annotations directly on maps.

At present, these services are the most popular platforms for sharing and annotating geo-coded information. However, none were built from the ground up with the

[*] Co first authors.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 97–110, 2009.

intention of creating maps for sharing. In most cases, sharing annotated maps was a feature added on top of platforms primarily designed to help people navigate from place A to B or find local businesses. In this study we wanted to step back and explore the creation and sharing of annotated maps without the ingrained technological constraints of platforms that were built with other goals in mind.

Moreover, these systems have already begun to establish patterns of use that may discourage or preclude the wide range of applications we feel are possible and perhaps likely to emerge in this space. Geospatial annotation systems intended to support publicly shared annotations will necessarily be designed differently from those intended to support small intimate groups. Similarly, those intended to support large geographical areas will differ from those focused on specific communities. In order to illustrate this area and its potential more fully, we provide a brief sketch of three hypothetical applications involving spatial annotation in order to illustrate the potential diversity of applications within this category.

Orientator allows new university students to share the burden of exploring a new physical environment, while getting to know each other more rapidly. Groups, such as dormitory floormates or incoming PhD students in a large department, are tasked with annotating a shared map of their new campus. As each individual explores the area, they contribute to a shared representation of "interesting" and "useful" locations. Identities are linked to annotations, revealing participants interests and tastes to fellow students, facilitating the formation of personal relationships around shared interests.

PubCrawler supports the cataloging and annotation of a particular location type within a geographic area—i.e., places to drink. The community of users is defined by mutual interest rather than membership in an organization, and may be subdivided by more narrow interests within the topic (e.g., martini vs. microbrew enthusiasts). Self-appointed domain experts as well as less specialized community members select, rate, and annotate pubs and bars, and the shared record is used by all members to identify new places to explore.

LocalExpert allows users in a particular geographic area to pose queries to anonymous domain "experts" and receive replies from those experts or from a cache of previous responses. This application can be seen as a geospatial instance of Answer Garden [1], in which the determination of expertise is generated by privacy-aware automated location traces of users' movements and/or manual annotations of favorite locations. Queries might be specified using natural language (e.g., "Which store in Manhattan has the best selection of shoes?") or by identifying points on a map (e.g., "What does [Café Zola; 112 Washington St.] serve for lunch?").

Each of these applications shares a set of features: users must be able to indicate (either explicitly or implicitly) locations of interest, and other users must be able to "read" (again explicitly or implicitly) the annotations, opinions, and experiences of other users in order to form a judgment about the location described, the author of the annotation, or both.

Starting from this sketch of the design space, our study set out to characterize the design requirements for systems that supports shared geospatial annotations, and the choices and tradeoffs that designers of this emerging class of applications face. In particular we sought to address the following questions:

- How do people choose to annotate spatial locations on a map?
- How do users assess other individuals' map annotations?

- How does the association of an individual with a map effect the interpretation of the annotations?
- What tools and affordances will best support both the authors and the audience for geospatial annotations?

The goal of our study was to gain insight into these questions. We examined the map annotation and interpretation behavior of 27 participants over a variety of conditions. We used the creation and act of sharing annotated maps as an analytic probe to better understand the processes used to annotate and interpret shared maps.

The paper proceeds as follows. First we lay out the design space for map annotations, followed with a brief review of the pertinent literature on this topic. We then describe the study, followed by the detailed findings. We finish with the design implications suggested by the study findings.

2 Related Work

Although there is considerable commercial interest in place-based annotations (e.g., the aforementioned Google Maps, Yahoo Maps, and Microsoft Live Maps), including recent developments focusing on shared annotations (e.g., BriteKite, Loopt, and Platial[1]), relatively little research attention has been paid to systems for sharing map-based annotations. Here we detail the research literature that does exist and briefly discuss additional literature that touches on this subject.

Studies of Social Map Annotation

A small number of studies examine social applications using maps. First, MapChat and MapMail [6, 13] are recent examples that attempt to extend the static nature of maps, using them to geographically ground real-time and asynchronous conversations, respectively. MapChat focuses on grounding conversation on place, rather than the creation of a persistent information artifact, which will be central to our study. MapMail primarily visualizes email information on a map.

Second, Ludford and others [10] studied the use of user-created, location-based reminders (LBR) and place bookmarks, as a method of capturing place metadata. They evaluated participants sharing fictitious personal bookmarks and location reminders on a prototype system. Their participants shared content created for personal purposes, rather than creating content to be explicitly shared.

Third, Priedhorsky and Terveen [14] built a geowiki, a new platform for building community maps. Their work has, to date, focused on the implementation of the geowiki, rather than on how users annotate and interpret specific information on community-created maps. A brief evaluation notes that users found the annotations and comments on the geowiki useful.

Fourth, Jones et al. [9] gathered design requirements for location-aware community systems. In interviews, they found that users were interested in having access to user contributed annotations of place (i.e comments and ratings). Users were also interested in creating their own annotations, but only if they were shared anonymously or with strict access controls. Their work does not discuss the act of creating or interpretation of these annotations.

---

1 http://britekite.com, http://www.loopt.com, http://www.platial.com

Finally, two other streams of locative research relate to social maps: the real-time sharing of location [3, 15], and annotations accessible only at specific locations [4, 7] (i.e. E-graffiti) . Studies of active location sharing systems are primarily focused on mutual awareness of location, rather than static annotations. Studies of systems such as E-graffiti have helped us understand the use of place-specific notes through real world deployments, but provide little exploration of the practices surrounding annotation.

Studies of Annotation More Broadly
A range of studies examines annotation in general. For example, Marshall [11] discusses dimensions of annotation of particular interest to our work: formal v. informal, explicit v. tacit (explicit annotations are easily understood by wide audiences, whereas tacit annotations may require specific context from the author), published v. private, and type of audience (global, institutional, workgroup, or personal). Our study explores the informal annotation of maps, where annotations are published and shared within small groups. Marshall [12] explores the relationship between personal and public annotations in hypertext, noting substantial refinement of personal annotations before they were shared online.

A number of other studies examine the technical features of and user motivations for using annotation systems. For example, Carter et al. [5] in their study of PDA-based annotations on shared multimedia content, characterize three categories of annotation: annotations for personal use, collaborative annotations, and public/social annotations. Ames and Naaman [2] describe four categories of motivation for adding tags to personal content on a photo-sharing site: personal-organization, personal-memory, social-organization, and social-communication. They found that participants who are motivated to share contextual information about photos socially often annotate with a specific known audience in mind.

Taken as a whole, these previous studies have elucidated the range of technical support required for annotation systems and the motivations for sharing, and have done so for a geospatial context. None, to our knowledge, have taken a close look at how people choose place-based information to share with others in a social context. This was the goal of our study. In the next section we describe our study, its methods, and the data collection.

3 Methods

To better understand how individuals annotate, interpret, and socialize around shared maps we designed a 4-phase study involving twenty-seven masters and PhD students (12 women and 15 men). They had lived in Ann Arbor, Michigan (USA), from 1 week to 30 years.

The four phases were designed with the following goals in mind. Phase 1: understand the uninfluenced annotation habits of students, paying particular attention to the types of annotations that they made. Phase 2: understand how users create maps after they have been encouraged to interpret the maps of others. Phase 3: understand how participants create maps to be shared with others when their identity is known. Phase 4: understand the role the shared maps could play in the social interactions of a small, known group.

Table 1. Our study consisted of four phases, in which participants created and/or viewed others' maps. Each participant was interviewed about their experiences creating and viewing the maps. (*: the same participants participated in phases 3 and 4).

Phase	Participants	Viewed Maps	Annotated Map	Interviewed	Audience
1	11		X	X	Lounge
2	6	X	X	X	Lounge
3	10*	X	X	X	Group
4	10*	X		X	Group

In phase 1, eleven participants were given markers and pens and asked to annotate a packet of 8.5"x11" (21.59 cm x 27.94 cm) paper maps (Fig. 1), showing various sections of Ann Arbor at different zoom levels. They were told the content of the maps would be posted publicly in the student lounge and shared with new students in their department, and that their name might or might not be included on their map.

In phase 2, six participants were presented serially with three or more annotated map packets from phase 1, then asked to annotate their own map packet, with the same tools and information about attribution given to phase 1 participants.

Phases 3 and 4 used the same participants: a group of 10 masters students who knew each other and had all participated in the same series of design workshops during the previous two months. In phase 3, they were shown up to three annotated map packets from phase 2, then asked to annotate their own map: a larger, 22x17" (55.88 cm x 43.18 cm) sheet with a close up of Ann Arbor's downtown and university areas on the front, and a bird's-eye view of the entire city on the back. Participants were asked to write their name on the maps, and were told that the maps would be distributed as a packet, with attribution, to the other members of this phase.

Phase 4 was a three week period immediately following phase 3 during which participants were in possession of the map packets (the packet lacked one phase 3 participant's map due to a technical issue). During the three weeks, they received two short email questionnaires about their use of the maps and interactions with other phase 3/4 members. At the end of the three weeks, participants were again interviewed about their experiences with the maps and each other, and were asked to compare two of their fellow group member's paper maps side-by-side with a duplicate created in Google Maps.

We interviewed each participant in each phase of the study. In phases where maps were created (1-3), participants were asked about decisions made during map creation. In phases where participants viewed the annotated maps of others (2-4), they were asked to discuss what they found interesting about the maps and their annotators. During analysis, we coded the 27 annotated maps, the 37 transcribed interviews, and the two rounds of email questionnaires. We paid particular attention to the process and strategies involved in annotation of the maps, interpretation of the maps, and the role maps played in one-to-one and group interactions in the Phase 3 and 4 participants.

4 Results

As noted, all study participants created maps. In addition, phase 2-4 participants both looked at and created maps. Moreover, phase 3 and 4 participants created maps for a known social group and reflected on the maps created by other group members. Thus, our findings fall into three distinct but related categories:

- ❖ Map Annotation: How did participants choose to annotate their maps? What content did they include and omit, and how did they choose to indicate elements of interest?
- ❖ Map Interpretation: How did participants make sense of other participants' maps? What aspects did they find interesting and valuable, and what aspects were viewed negatively?
- ❖ Impacts on Interpersonal Relationships: What effect did sharing maps have on the relationships among members of a known social group? Did opinions of other group members change as a result of viewing others' maps?

In this section, we report our observations related to each category.

4.1 Annotating Maps

When annotating maps (Figure 1), participants were faced with choices about what items to include on the map, what to say about each item, and how to graphically represent the annotations on the map. In making these decisions, participants considered both the usefulness of the map for their audience, and how it reflected upon them personally.

An implicit concern among participants was to include annotations that were useful to the perceived audience. In the interviews, it appeared that authors considered their experience with a location and whether it would be novel and interesting to the audience, and sought to balance what they had to offer with what they thought the reader would want to see. In addition, a few authors sought to transform some aspect of the reader's experience of the area being annotated. As a secondary concern, authors used their annotation choices to attempt to manage others' impressions of them.

Finding: Participants annotated based on personal experience

Participants mainly chose to share locations they had experience with. Some explained that by including a location, they felt they were endorsing it, and thus should be able to substantiate that choice with experience. For example:

> I wouldn't want to put something on the map that I've never been to. (p2.02)[2]

The emphasis on personal experience was evident in one participant's decision to omit a jewelry shop that she didn't view as "hers."

> There was a jewelry shop that I thought about putting on there.... I guess I feel like it's not really my place. My roommate likes to go there a lot, so I felt like it was more her place and not mine. (p1.09)

---

[2] Participant quotes are identified by phase and participant number. "p1.09" refers to the 9th interviewee in Phase 1. Since the same individuals were interviewed in Phases 3 and 4, each participant's number remains the same across the two phases. Thus "p3.01" and "p4.01" refer to two different interviews with the same person.

Finding: Participants tried to anticipate the usefulness of their annotations

The desire to be "useful" to one's audience generalized across all phases, but the construction of "usefulness" varied depending on the participants' perception of who the audience was. In phases 3 and 4, where the audience was known and familiar, participants often figured that novel content was the most likely to interest their audience. Novelty was expressed as both new locations and new information about familiar locations.

> I tried to think of things that I thought were really…that I appreciated about Our City, rather than things that everybody would put, like little secrets that I know. (p2.05)

Indeed, in phase 3 and 4, these "little secrets" were common on the maps. Participants annotated such locations as a hidden, graffiti-covered alley, a running route through an area with interesting houses, and an unlikely study location in the natural history museum. One participant created a whole category for "Food Secrets and Wine."

More quotidian annotations were included in these phases as well, but most often with the intention of being useful. One participant marked several parking garages in the somewhat congested downtown area. Another marked all the coffee shops he knew had WIFI. Some marked parks good for weekend outings, or main roads in and out of the city.

Some phase 1 and 2 participants, who only knew their map would be shared in a public space, felt that the needs and interests of their audience were unclear or elusive. As a possible consequence, more reflective individualistic annotations were included on the maps in these phases, such as commuting routes, workplaces, private personal routines and places participants aspired to visit. Though this may be related to the fact that these were among the seed suggestions in the instructions for these phases, there is also some evidence that an author's lack of knowledge about his audience may lead to a more introspective map. Here, one participant explains his choice to include a marginally-relevant location:

> I did think twice about whether I should put in the comic books and the game stores... I think most people wouldn't care where those are... I just put it down in case there was someone else who did care, who did want to see it. (p1.06)

Another participant expressed it in terms of the low consequences of disclosure:

> I though about the surgery one a little bit, because if it's public with your name on it, and people you... I don't care if it's strangers, but if it's a lot of people I know, I may not want to explain it to everyone all the time. (p1.13)

Notably, even some participants in phases 3 and 4 felt that their notion of audience was still too vague to be directive. One participant expressed his confusion about whom to annotate for this way:

> I'm looking at the maps... I might look for hospitals, maybe other people looking for schools, other people looking for restaurants. They're not all the same. (p4.05)

Thus, even though participants wanted to create content that was interesting and useful to others, they were not always able to assess whether choices they were making were going to be well received.

Finding: Participants sought to have an actionable impact on their audience

Beyond utility, some participants expressed the desire to have a certain effect on readers of their maps. One user expressed the desire to influence people towards certain behaviors:

Fig. 1. Participants annotated paper maps with information they wanted to share with their peers

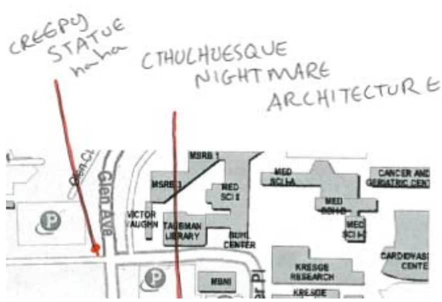


Fig. 2. Some participants labeled unusual locations in hopes to inspire certain emotions in their audience

> I like encouraging people to do socially responsible things like shopping at the farmer's market, so that was definitely going to come out in this. (p3.03)

Another participant, whose choice of locations was unusual and annotation style concise but evocative (see Fig. 2), hoped to inspire certain emotions in his audience:

> I would hope that people would recognize some of these landmarks, if you'll call them that, or be curious about certain places. (p3.06)

Finding: "Face work" played a part in participants' choice of annotations

While participants were more emphatic about the informational utility of their maps than the way the maps reflected their personality, "face work" [8] did emerge as part of some participants' filtering process and affected their annotation choices.

> I was cognizant of who I was and what kind of persona I was putting out there. (p4.10)

Novelty also played a part in some participants' expression of their personality.

> I feel that I am revealing more by putting in things that aren't obvious. (p1.11)

As outlined above, novelty was not only expressed in the choice of points, but also in the explanations that accompanied them. One user explained how this related to her presentation of self:

> The point is for people to find out something about you from your map... you have to include why you like that place, or why you should go there, instead of just pointing it out. (p2.01)

4.2 Interpreting Map Annotations

Not surprisingly, participants emphasized utility in their interpretation of others' map annotations. However, other factors came into play, such as assessing the author's similarity to themselves and credibility. Readers drew on many aspects of the annotations to assess their quality and personal relevance, including the content, style, level of detail, as well as their assessment of the author's personality and interests.

Finding: Readers focused on usefulness in judging maps

In keeping with their emphasis on creating informational value when annotating maps, participants judged others' maps most strongly on their usefulness, with a significant emphasis on novelty. In Phases 3 and 4, novel content was particularly valuable, in part because it was scarce:

> I've been to all the places that they've circled. So, except for the studying place, I didn't find it useful because I had already been to these places. (p3.10)

The quality of the explanations included in annotations was also considered important. Many participants felt that for an annotation to be useful, it needed to contain enough content for the reader to discern why the author endorsed that particular location. If the annotation contained more descriptive, informative details, it was valued more highly.

> I noticed that a couple of people actually marked the Dana Building but one person has given more details – "it's a nice place to study, it's quiet, it has natural light"... I guess what I'm saying... there's a difference between marking it and giving more detail because you can decide if you would be interested in that place or not. (p4.07)

Participants explained that they used these details to help them profile the author's perspective and compare it to their own. This helped them judge if the map would be interesting or useful to them. This was particularly prevalent in cases where the reader did not know the author of the map.

> Just the fact that there is more detail just makes me trust them more... if I circle this—Ashley's—and say, "this place is good," I don't really know how or why they are evaluating it that way, but with the detail... there's something specific there that allows me to trust it. (p3.09)

Finding: Readers tried to make sense of the author's personality

In phase 4, readers used their knowledge of the author's personality to help them evaluate the author's choice of points and decide if they were of value.

> If I know the person prior and I kind of identify them, you know, it's going to add a little bit of credence to what they chose to show, or, like, it will probably increase the likelihood too that I will actually go to any of the places that they suggest or chose to mark. (p4.09)

When readers didn't know the author, and lacked prior knowledge of his or her personality, they often used annotations of familiar interests to calibrate their assessment.

> I notice here that Blimpy Burger is marked and I approve. (p3.06)

In all the phases, when reviewing unattributed maps, participants looked for themes among the points to give them an idea of the author's character. This often resulted in an inventory of what the author included on the map, which was used to characterize the author lieu of more substantive knowledge about their personality. For example:

> They mark the Arb for great hiking, walking or frisbee, so this leads me to think that this person has engaged in those activities (p2.06)

Less frequently, participants looked to the style and form of the annotations for indications of the character of an anonymous author.

> The annotation influences your opinion of the person, I think. Like the second person had a very distinct handwriting. It was dark and legible. The third person kind of scrolled a little and they wrote very lightly. The first person seemed a little messy. (p3.07)

Finding: Readers valued aspects of informal annotations that improved information retrieval and effectively conveyed personality

While form and style were used to discern an annotation author's character, they were more often judged by how well they supported information retrieval activities such as scanning and separating points from the substrate. One participant explains:

> Color, labels followed by brief descriptions...an underlined title of each place with enough space to write it. Instead of on the map, it's on the border. I think this is a cleaner and improved way to represent these spots. (p3.06)

That said, form and style were substantially less important to our participants than the quality of the content. One participant summed up it up this way:

> If it's sloppy but the content is good, I can work around the sloppiness. (p4.10)

However, there were cases where the free-form annotation capability allowed readers to form a better sense of the author's personality. In phase 4, the comparison between paper and Google Maps revealed that there was an intangible quality lost in the digitization of the paper maps. This was expressed even when participants felt the Google Map was more legible than the paper map. Most often, participants referred to the quality as "personality."

> Just in terms of, like, being able to see somebody's handwriting and the colors they chose for, like, restaurant versus parking versus study rooms. I mean, there's just more personality, you know. Handwriting has a lot of personality in it. You can't convey that in Google Maps, or a precise Google Map. (p3.09)

The last part of the participant's comment, "a precise Google Map," refers to the fact that most annotations on the paper maps were transferred to the Google Map in the form of points or specific routes. On the paper maps, more amorphous forms of annotation were frequently used, such as lines to represent parts of streets, shapes to indicate a general area, outlines around buildings, or visual textures (such as cross-hatching) to represent a characteristic of a point or region. (See Figure 3.)

Another perceived advantage of paper maps over their Google Maps counterparts was that the paper allowed a comprehensive view of annotations and their content in a single glance. This is largely because only one complete textual annotation can be viewed at a time in Google Maps – other annotations appear as a list in a side panel, but are not directly linked to their points of reference. This ability to take in an annotation author's entire map at once was seen as valuable by participants.

> I kind of feel like the digital one is a little bit less immediately accessible because all of [p3.10]'s annotations are available immediately [on the paper map] whereas with [the Google Map] I guess I'd have to go through each one and click on it individually. (p4.04)

4.3 Effects of Map Sharing on Members of a Known Social Group

Finding: Maps yielded little new information about peers to members of a known group

While phase 3 and 4 participants knew each other, and received a copy of most other phase 3 and 4 participants' maps, they felt that the maps did not enhance sense of belonging to a group. They also felt the maps revealed little, if anything, about the character of the group; when prompted, participants mentioned only that the group had predictable points on their maps in common, like department buildings or nearby coffee houses.

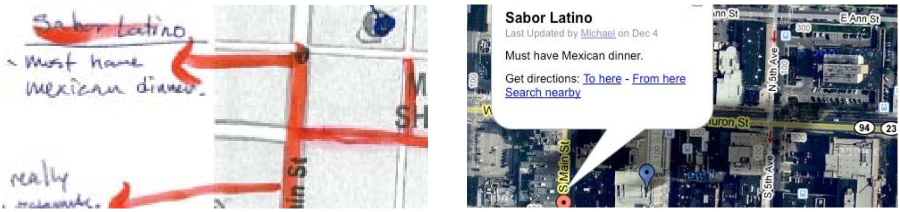


Fig. 3. In phase 4 students compared a map annotated in phase 3 with the same information annotated on a Google map. Students reported an intangible human quality that was missing from the digital Google maps.

During the three-week interval between interviews, each participant intersected with several of the others. Still, most could not recall talking with them about the content of the maps.

In a few cases, the maps revealed new information to one participant about another, but this did not inspire map-related social interaction between the individuals. Instead, participants said the maps simply reinforced their existing impression of the authors:

> [p9.06] knows a lot about the area specifically but also since he's been here for a while he knows a lot about places I wouldn't know and he's just a quirky guy, so that the places he chose to mark kind of reflect his personality in a way. (p4.10)

When the maps did diverge from a reader's impression of the author, readers enjoyed learning new things about familiar colleagues. However, several said that this did not affect their overall impression of the person nor their social behavior toward them, because that the maps were no match for their existing knowledge.

> So me seeing their maps doesn't really influence…I'm not like all of a sudden, "oh yeah now I want to hang out with [p4.01] because he goes to Bar Louie." I already want to hang out with [p4.01] because I already know his personality. (p4.10)

Many of the phase 3 and 4 participants felt that the maps would be more useful to people who were new to each other, or new to an area, such as when they were starting their graduate program.

5 Discussion

Our works sheds light on the factors that influence how people choose to annotate shared maps, and how they read those annotations to learn more about the shared physical environment represented in those maps. While we expected to see that map annotation sharing in a known social group would impact the social interactions among the members of that group, we did not see much evidence for this. Rather, we saw evidence that readers attempt to employ knowledge (or conjecture) about the annotator's personality as a means to interpret the map rather than using the map as a way to influence their view of the author's personality. In other areas, however, we believe that our findings have the potential to guide future systems that depend on shared geospatial annotations.

5.1 Implications for Design

We have described a number of observations that could impact the design of collaborative geospatial annotation systems. Here we summarize those observations and identify the tradeoffs that will face future designers in this domain.

Foster a mix of the familiar and the novel. Determining the credibility of an author is a critical step in a reader's interpretation of the value of a set of annotations. In our study, this was primarily done through calibration: the reader would probe the author's opinion of a subset of familiar places—if the author's judgment matched their own, they would be receptive to novel items included by that author. However, some authors were disinclined to annotate places they thought would be "obvious" to others. Since readers value novel content, but need familiar content to assess an author's opinion of it, application should encourage the annotation of both.

Support the feedback loop between authors and audience. Following the previous point, it is important for authors to be able to assess what content is valuable to their audience, and for the audience to be able to communicate their judgments to authors. Authors adopted strategies for content selection and styles for annotation from others' work, indicating that a dialogue is valuable for guiding authoring activities. Authors also gave considerable thought to the value that would be assessed by their audience, even when that audience was vague. In the absence of established norms and protocols, it is especially important to foster a dialog between annotators and readers, so that they may co-create a mutually-useful annotation style over time.

Understand the appropriate tradeoff between flexibility and aggregatability for a given application. In our study, participants generated annotations by creating free-form marks on paper, a method which afforded extreme flexibility that would be nearly impossible to replicate in an electronic system. At the other end of the spectrum, digital annotation systems using specific points in a highly-structured geospatial coordinate system, provide little stylistic flexibility but allow for the aggregation of many authors' inputs. We encountered shortcomings in both approaches. Hand-drawn annotations were sometimes problematic—writing was hard to read, maps became cluttered, and stylistic choices were deemed arbitrary or even detrimental. However, in some cases, the lack of constraints helped readers better assess authors' personality and credibility: thoughtfulness and thoroughness were imputed to authors based on their stylistic choices, and users expressed difficulty perceiving differences between authors in a more regimented representation. In addition, certain types of flexible annotations were valued by our participants, such as imprecisely bounded areas, routes, or locations whose precise geospatial anchor was unknown. Designers should be aware of these trade-offs for their particular applications and audiences.

Reflect authors' "personality" to aid audience members' interpretation of annotations. In Phases 3 and 4, participants used prior knowledge of authors' personalities to interpret their maps. In Phases 1 and 2, participants attempted to form an idea of the author's personality in order to assess map quality. In cases where an external relationship exists, simply providing a link between a set of annotations and an author's identity would likely suffice. However, where the relationship among authors and audience is primarily or entirely mediated by the annotation system, it becomes more important that readers see as much of the original annotation content and personal style of the author as possible.

Maintain high information density. In addition to allowing audience members to quickly assess the interests, tastes, and diligence of an author, viewing a large amount of an author's geospatial annotations in one glance allows them to perceive a coherent "story" of the author's experience of the mapped region, rather than discrete endorsements or critiques. This allows authors to more fully construct the "message" they are communicating to audience members, while also allowing audience members to calibrate the general opinions and biases of a particular author.

6 Conclusion

In recent years there has been an increasing number of systems that support authoring and sharing geospatial annotations. As this trend continues, the types of systems, users, and uses will expand ways that are difficult to foresee. The fundamental acts of annotating, sharing, and interpreting maps, however, can be expected to be a central part of any future developments. In this paper, we presented a study of map annotation, which illuminated a number of factors that influence how people decide what to annotate and how they interpret what others' have annotated. We also generated a number of implications for the design of future systems in this space.

We see many possible future directions for this work. One such direction would involve the implementation, deployment, and study of specific applications and/or tools to support map annotation and sharing. It would also be interesting to compare our findings to a study of larger annotation sharing communities such as Google Earth. Given the abbreviated feedback cycle between authors and readers represented in our study, we would be interested to learn more about how annotation practices evolve over time in other communities. We are also interested in exploring how the annotation tools themselves might affect the choices made by authors and the interpretations made by readers, specifically to explore the tradeoff between flexibility and aggregatability discussed in our design implications.

While we see the results of our study as informative in themselves, we are even more hopeful that this work opens new avenues of research for ourselves and others.

Acknowledgments. We would like to thank our study participants, Adam Steenwyk, Mouly Kumaraswamy, Sean Munson, Katherine Connors and Joel Frank. This work supported in part by the National Science Foundation (IIS 0705672).

References

1. Ackerman, M.S.: Augmenting organizational memory: a field study of Answer Garden. ACM Trans. Inf. Syst. 16(3), 203–224 (1998)
2. Ames, M., Naaman, M.: Why we tag: motivations for annotation in mobile and online media. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, San Jose (2007)

3. Barkhuus, L., Brown, B., Bell, M., Sherwood, S., Hall, M., Chalmers, M.: From awareness to repartee: sharing location within social groups. In: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems. ACM, Florence (2008)
4. Burrell, J., Gay, G.: E-graffiti: Evaluating real-world use of a context-aware system. Interacting with Computers 14, 301–312 (2002)
5. Carter, S., Churchill, E., Denoue, L., Helfman, J., Nelson, L.: Digital graffiti: public annotation of multimedia content. In: CHI 2004 extended abstracts on Human factors in computing systems. ACM, Vienna (2004)
6. Churchill, E., Goodman, E.S.: Mapchat: conversing in place. In: CHI 2008 extended abstracts on Human factors in computing systems. ACM, Florence (2008)
7. Espinoza, F., Persson, P., Sandin, A., Nystr, H., Cacciatore, E., Bylund, M.: GeoNotes: Social and Navigational Aspects of Location-Based Information Systems. In: Proceedings of the 3rd international conference on Ubiquitous Computing, Atlanta, Georgia, USA. Springer, Heidelberg (2001)
8. Goffman, E.: On Face-Work, in Interaction Ritual. Anchor Books (1963)
9. Jones, Q., Grandhi, S.A., Whittaker, S., Chivakula, K., Terveen, L.: Putting systems into place: a qualitative study of design requirements for location-aware community systems. In: Jones, Q., Grandhi, S.A., Whittaker, S., Chivakula, K., Terveen, L. (eds.) Proceedings of the 2004 ACM conference on Computer supported cooperative work. ACM, Chicago (2004)
10. Ludford, P.J., Priedhorsky, R., Reily, K., Terveen, L.: Capturing, sharing, and using local place information. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, New York (2007)
11. Marshall, C.C.: Toward an ecology of hypertext annotation. In: Proceedings of the ninth ACM conference on Hypertext and hypermedia: links, objects, time and space—structure in hypermedia systems: links, objects, time and space—structure in hypermedia systems. ACM, Pittsburgh (1998)
12. Marshall, C.C., Brush, A.J.B.: Exploring the relationship between personal and public annotations. In: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries, ACM, Tuscon (2004)
13. Nelson, L., Churchill, E.F.: Mapmail: restructuring an email client for use in distributed teams. In: CHI 2007 extended abstracts on Human factors in computing systems. ACM, New York (2007)
14. Priedhorsky, R., Terveen, L.: The computational geowiki: what, why, and how. In: Proceedings of the ACM 2008 conference on Computer supported cooperative work, ACM, San Diego (2008)
15. Want, R., Hopper, A., Falc, V., Gibbons, J.: The active badge location system. ACM Trans. Inf. Syst. 10(1), 91–102 (1992)

Effect of Peripheral Communication Pace on Attention Allocation in a Dual-Task Situation

Sofiane Gueddana[1,2] and Nicolas Roussel[1,2]

[1] LRI - Univ. Paris-Sud & CNRS, Orsay, France
[2] INRIA, Orsay, France
{gueddana,roussel}@lri.fr

Abstract. Peripheral displays allow continuous awareness of information while performing other activities. Monitoring such a display while performing a central task has a cognitive cost that depends on its perceptual salience and the distraction it causes, i.e. the amount of attention it attracts away from the user's primary action. This paper considers the particular case of peripheral displays for interpersonal communication. It reports on an experiment that studied the effect of peripheral communication pace on subjects' allocation of attention in a dual-task situation: a snapshot-based peripheral monitoring task where participants need to assess the presence of a remote person, and a central text-correcting task against the clock. Our results show that the addition of the peripheral task caused a drop in the success rate of the central task. As the pace of snapshots increased, success rate decreased on the peripheral task while on the central one, success rate remained the same but failures to reply in time occurred more frequently. These results suggest that the increase in pace of snapshots caused participants to change their strategy for the central task and allocate more attention to the peripheral one, not enough to maintain peripheral performance but also not to the point where it would affect central performance. Overall, our work suggests that peripheral communication pace subtly influences attention allocation in dual-task situations. We conclude by discussing how control over information pace could help users of communication systems to adjust their local distraction as well as the attention they draw from remote users.

Keywords: Computer-mediated communication, peripheral communication, attention allocation.

1 Introduction

Specific characteristics of interpersonal communication technologies have been studied for quite a while. Studies of video-mediated communication revealed few objective advantages of adding video to audio for focused problem solving tasks [1], for example, but they also showed the value of video for creating shared workspaces and assessing the availability of others. Media space studies particularly demonstrated this last role, emphasizing the importance of long-term and always-on connections and promoting the concept of peripheral awareness of each others' activities [2, 3]. Many of today's communication systems build on these notions of constant accessibility and

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 111–124, 2009.

peripheral awareness. A few additionally support the transformation of a peripheral communication into a primary activity. As an example, instant messaging applications not only provide constant information about people's presence and availability but also support the rapid exchange of text messages. Yet, managing users' transitions between background and foreground activities remains a key challenge of modern communication system design. In Weiser and Brown's terms, the challenge is to create calm technologies that engage both the center and the periphery of our attention and move back and forth between the two [4].

Monitoring a peripheral communication while performing another task has a cognitive cost that depends on the perceptual salience of the communication and the distraction it causes. Minimizing this cost is usually desirable and even necessary in situations like car driving where the ability to keep a communication in the background can be critical. But what appears as a cost might quickly turn into a benefit as one starts focusing on the peripheral communication and placing it at the center of attention. This paper considers the particular case of peripheral displays for interpersonal communication. We conducted a quantitative experiment to determine the effect of a snapshot-based peripheral monitoring task on subjects' attention in a dual-task situation. Our results suggest that peripheral communication pace subtly influences attention allocation in this situation. The paper is organized as follows. After introducing some related work, we describe the quantitative experiment that we conducted, summarize our results and conclude with a discussion.

2 Related Work

Attention is defined by psychologist William James as "*the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought*" [5]. James further explains: "*Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state which in French is called distraction*". As we said, a key challenge of modern communication system design is to create systems that do not monopolize user attention but can fade into the background and be promoted on demand to the foreground.

2.1 Fading into the Background: Peripheral Awareness

Communication systems designed to fade in the background avoid explicit interaction with users to prevent distraction. They instead resort to presets and automatic capture to determine the information to be communicated. Images, sounds and other automatically captured data can provide numerous awareness cues in this context, but at the cost of introducing privacy concerns. Filtering techniques have been proposed to help users mitigate these concerns by altering the data to be transmitted [6, 7, 8] or by abstracting it to communicate higher-level information [9, 10, 11].

Peripheral information displays also impose constraints on information rendering. As reported by Pousman and Stasko [12], numerous researchers have investigated ways to create systems that "*aim at presenting information in a way that is not* *distracting but aesthetically pleasing and tangible to varying degrees*". Information salience plays an important role in this context, defined as "*the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention*" [13]. Lights, haptics, scent and adapted everyday objects have been used to create low-salience ambient devices that blend into the user's environment [14, 15, 16]. Outside the specific domain of computer-mediated communication, studies have also investigated how to design peripheral displays so that they provide the most information while having the least impact on the user's performance on a primary task [17, 18, 19]. These studies particularly showed the ambivalent role of animations and the importance of physical characteristics of the display such as its size, position and orientation.

As illustrated by the previous examples, there is a substantial body of literature describing ambient communication devices or ambient information systems that might be used for communication. But creating systems that do not monopolize user attention addresses only the first part of the calm technology challenge. These systems should also be capable of promoting on demand interaction to the foreground. Unfortunately, the systems designed for background use can rarely be used for focused interaction, and conversely.

2.2 From Background to Foreground

A number of studies have investigated ways to help people maintain awareness of backgrounded tasks and facilitate their resumption [20]. Peripheral displays are known to help both interrupters to time communication and interruptees to handle interruptions [21, 20]. But the transition of background tasks to the foreground requires that the displays incorporate mechanisms for drawing users' attention. A too low salience might lead them to miss important information or opportunities for more focused interactions, compromising coordination mechanisms. It has thus been suggested that systems should support different degrees of salience [9] and could increase the one of incoming messages that may deserve immediate attention [22]. People's attentional focus is known to be vulnerable to certain kinds of stimuli [23], and numerous techniques have been proposed that take advantage of this for getting users' attention [20].

The initiation of a communication is a gradual process that involves negotiation between parties [24]. This idea of gradual and negotiated engagement is also one of the most interesting aspects promoted by early media space studies [25, 26]. As noted by Birnholtz et al., "*paying attention to someone is itself a communicative act – an implicit request for interaction*" and "*interest in interaction on the part of the initiator is expressed by paying attention to his or her target in progressively more intrusive ways*" [27]. Yet, relative little effort has been made to explore ways to actually support this gradual intrusion and the overall collaborative process of contact negotiation [28, 26], Community Bar [29, 30] and OpenMessenger [27, 31] being two notable recent exceptions.

Prior research has shown the value of peripheral displays as ways to maintain awareness of backgrounded tasks and facilitate their resumption. But monitoring a peripheral display while performing another task has a cognitive cost that system designers need to carefully take into account. The cost should be generally low, so that the display can fade into the background, but not too low to preserve its function. At times, a higher cost might be beneficial to get user's attention or in response to his or her increased interest as (s)he wants to promote the background task to the foreground. But how can we measure the resources used by a peripheral task?

2.3 Measuring Attention

The attentional cost of a particular task cannot be measured directly. Psychology studies of divided attention have thus mainly focused on dual-task situations in which participants share their limited attention between a main (or central) task and a peripheral one. When execution of at least one of the tasks deteriorates if they are performed simultaneously rather than separately, the tasks are said to interfere with each other: their combined attention requirements exceed resource capacity and execution deterioration in one can be used as a measure of the attentional cost variation of the other. Task execution deterioration is generally measured as a performance decrease, but can also be measured as an error rate or response time increase.

The idea that attention is sharable and limited was introduced by Kahneman, who also suggested that humans change their allocation policy over time [32]. According to his theory, although attention capacity is limited, it fluctuates according to arousal levels for internal causes (e.g. effort and motivation) and external ones (e.g. saliency of environment stimulation). Studies showed that motivational variables do not necessarily correlate with performance [33, 34]. Other studies have illustrated the importance of three other factors on dual-task performance: task difficulty [35, 36], practice [37, 38] and task similarity [39]. Overall, these studies suggest that the similarity, novelty and complexity of two tasks greatly contribute to the fact that they compete for the same attentional resources.

Studies of peripheral displays have often used the dual-task paradigm to measure their attentional cost by comparing single and dual task execution. Our motivation for this work was a bit different. Having a significant experience in the design of video-mediated communication systems, we were interested in attention allocation when peripherally monitoring regularly updated snapshots of a distant person. To our knowledge, the attentional cost of this task had never been formally evaluated. We were of course interested in measuring it, but we were more particularly interested in studying the effect of display richness on attention management. Dabbish and Kraut had already shown that monitoring an information-rich display imposes a substantial attentional cost and that an abstract display provides similar benefit with less distraction [40]. The information richness of regularly updated snapshots being somewhat difficult to control, we decided to focus on the pace of the update.

3 Experiment

We conducted a dual-task experiment with the following goals:
- (intermediate goal) determine the cost of adding a snapshot-based peripheral monitoring task to a central one;
- (main goal) determine the effect of varying the pace of snapshots on attention allocation.

Measuring peripheral attention requires the peripheral and central tasks to be sufficiently novel, complex and similar to interfere with each other. The nature of the central task does not really matter as long as it interferes enough with the peripheral one. To test whether our two tasks indeed interfered and to formally evaluate the attentional cost of adding the snapshot-based peripheral monitoring task to the central one, we decided to compare single and dual task execution. Assuming that increasing the pace of snapshots would make the peripheral monitoring task more difficult, we were expecting the allocation of more attention to it and possibly the deterioration of the execution of one or both tasks. To precisely determine this effect, we decided to compare two different pace conditions and to measure execution deterioration in terms success, fail and timeout rates. Our hypotheses were the following:

- the addition of the snapshot-based peripheral monitoring task to a central one would have an effect on central task execution;
- varying the pace of snapshots would have an effect on the execution of both tasks.

3.1 Tasks

Participants had to simultaneously monitor a peripheral display while performing a central task. The central task and the peripheral one appeared on two separate screens. Inspired by the one used by Maglio and Campbell in a similar experiment [17], our central task was a text correcting exercise. This task involved mainly linguistic capabilities associated to decision making. To make the cognitive load adjustable, we introduced a time limit per sentence. By limiting the time available to read and decide, we turned the task into a “race against the clock” similar to the Tetris game used by Bartram et al. [41]. The top-right image of Fig. 1 shows the text correcting display. The sentence to correct is shown highlighted. At the top of the screen, a progress bar shows the time left to answer. At the bottom, two “correct” and “incorrect” buttons can be clicked with the mouse to provide the answer. Each text correction trial of the experiment, i.e. each sentence, had one among three outcomes: participants either clicked one of the buttons or didn’t click any in time. The text used for the correcting task consisted in French articles with similar readability levels taken from a popular newspaper (Kandel and Moles measure [42] ranging from 48 to 52). Errors of two types were created by hand: syntax errors based on word inversions, and gender or number agreement errors. Zero, one or two errors were randomly introduced in each sentence of the articles presented to the participants.

For the peripheral task, we decided to simulate the peripheral awareness of distant people through an image-based communication system. Still pictures showing a desk and either a person or an empty chair were displayed for a random period of time on a separate screen, four seconds fade-in transitions being used between pictures (Fig. 1). We chose a monitoring rather than awareness task so that it would be sufficiently demanding and thus induce a dual-task performance trade-off. We also made this task more difficult in two ways: participants received instructions that emphasized the priority of the peripheral task over text correction, and changes in the remote person’s presence as well as interruptions were very frequent. During the experiment, participants were randomly interrupted and both displays suspended at frequent times, i.e every 4 to 12 seconds. To measure monitoring performance in dual-task conditions, participants were then prompted to recall the presence state of the remote person by

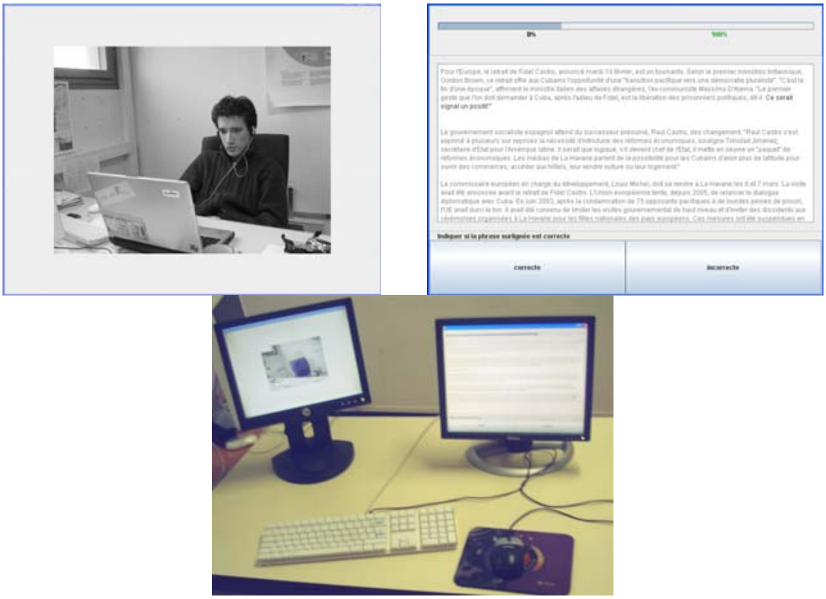


Fig. 1. Experiment tasks and apparatus

pressing the Return or Backspace key on a keyboard. To avoid introducing a bias when performing text correction alone, participants were also randomly interrupted in the single task condition, but simply had to press a key to resume it.

3.2 Design

The experiment was constructed as a one factor within subject design with a *slow update pace* condition, a *fast update pace* condition and a *single task* condition. Snapshots update pace ranged from 5 to 8 seconds in the fast condition, and from 17 to 21 seconds in the slow one. Executions of both tasks were compared across the update pace and single task conditions. Participants performed the single task condition as a control for the two dual-task conditions. The single task consisted in correcting a text document on a central screen with no peripheral task, while the dual-tasks included monitoring the presence of a remote person through snapshots displayed on a peripheral screen. Each participant performed both dual-task conditions and the single task condition.

The presentation order of conditions was balanced across two groups of participants (Fig. 2). The presentation order of articles in the text correction task was also balanced across participants. Dual-task trials were grouped in two blocks of 75 trials each, one for the fast condition and the other for the slow one. The single task condition was divided into three short blocks of 25 trials: before, between and after each dual-task condition block.

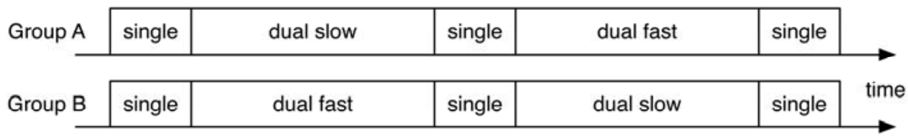


Fig. 2. Presentation order of dual-task condition blocks were counterbalanced in two groups, A and B. Single task blocks were interleaved with slow and fast blocks.

For presence monitoring, success rate (i.e. the percentage of correct recalls) was measured as a dependent variable. For text correction, each trial outcome was classified into one of three categories: correct, incorrect or failure to reply in time. Three dependent measures were then collected for each participant:
- success rate: the percentage of correct answers among all text trials;
- fail rate: the percentage of incorrect answers among all text trials;
- timeout rate: the percentage of failure to reply in time among all text trials.

3.3 Procedure and Participants

The experiment procedure included several phases: training, calibration, test and a post-test questionnaire. The task was orally briefly described, and detailed written instructions were displayed. These instructions emphasized the priority of monitoring the distant colleague over the text correction task. Participants then performed the three conditions during a training phase of 10 minutes. During this training, a visual feedback was displayed helping the participants learn the different error types. After the training phase, as text correction speed differed among participants, a speed calibration test of 15 minutes was performed. During this test, we varied the time per sentence to determine the appropriate speed for each participant to correctly perform the text correction. After calibration, the time limit was introduced and the test started. The participants performed the single task and dual-task blocks. After completing all blocks, participants answered a post-test questionnaire in two parts: participants had first to estimate their linguistic and visual abilities, and then to describe their subjective perception of the differences between the conditions and the strategies they employed in multitasking.

Twelve adult native French speakers were recruited for the experiment, ten males and two females aged from 22 to 30 (26 on average). All had normal or correct-to normal vision acuity and perimeter (more than 175°), six of them wearing lenses or glasses. The experiment apparatus consisted of two screens, a computer, a mouse and a keyboard. A distance of 20 cm separated the screens, two 1280x1024 LCD monitors (17” for the central and 15” for the peripheral), as seen at the bottom of Fig. 1. The software was implemented in Java 1.4 on a 3,4 GHz Pentium 4 computer running Microsoft Windows XP Pro. Snapshots were 512 pixels wide by 384 pixels tall, scaled down from pictures taken at a resolution of 2560x1920 pixels. The text to correct was displayed using the Lucida Grande font in 13 points.

3.4 Results

Data was analyzed to determine the effect of adding the peripheral task to the central one and the effect of varying the pace of snapshots on both tasks. We compared success rate differences in peripheral monitoring under the two dual-task conditions (slow and fast) using a repeated measures analysis of variance (ANOVA). The test shows a significant difference of the success rate in the monitoring task between the slow and fast conditions (F=22.02, p<.0001*). In the slow condition, this rate was at an average of 96,85%, while in the fast one, it fell to 92,78% (Fig. 3). We also looked at differences in the central task dependent variables under the three conditions (single, slow and fast) using ANOVA. We found a significant effect of condition on the success rate (F=9.53, p<.0001*) and timeout rate (F=12.82, p<.0001*), but not on the fail rate (Fig. 3). Concerning the success rate, the test shows a significant difference between single and dual-task conditions (F=9.35, p=.0023*), but no difference was found between the two pace conditions. For the timeout rate, we found no difference between the slow and single conditions, but the test shows a significant difference between the slow and fast conditions (F=12.8, p<.0001*).

Participants reported in the post-test questionnaires that they continued to learn multitasking and each task after the training period. We thus conducted a post-hoc analysis to evaluate learning, tiring, order and group effects on text correction. We found no significant difference in the success rate of the text correction task between the three single condition blocks. We also compared the first dual-task block with the second one for differences in text correction and monitoring success rate. The test shows significant differences between the two dual-task block for success rate of both the correction (F=5.53, p=.0188*) and the monitoring (F=5.19, p=.0227*) tasks. From the first to the second dual-task block, correction success rate rose from 69% to 74%, while monitoring success rose from 94.0% to 95.9%. Additionally, to evaluate order and group effect, we compared groups A and B in dual-task conditions for differences in success rate on both tasks. But we found no significant difference in the success rates, neither for correction nor for monitoring. We also compared groups A and B in

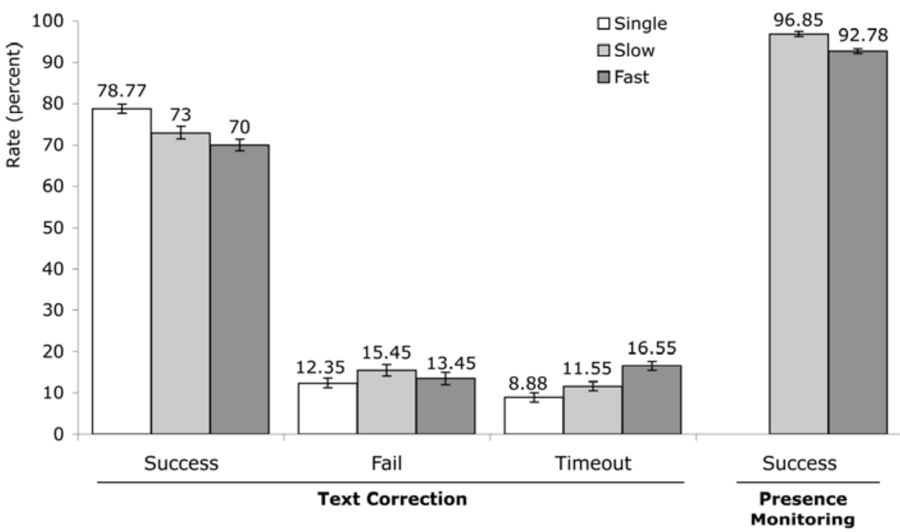


Fig. 3. Text correction and monitoring performance across the three conditions. Error bars represent standard error.

Table 1. Two by two comparison across all conditions using ANOVA test of the dependant variables: monitoring success rate, success and text correction timeout rate

Conditions	Monitoring success	Correction success	Correction timeout rate
slow vs. fast	F=22.02, p<.0001*	no difference	F=12.8, p<.0001*
single vs. fast	-	F=18.71, p<.0001*	F=24.37, p<.0001*
single vs. slow	-	F=8.46, p=.0037*	no difference
dual-task first vs. second blocks	F=5.19, p=.0227*	F=5.53, p=.0188*	F=4.42, p=.0355*
single task blocks	-	F=8.46, p=.0037*	no difference

the single task condition and did not find any significant difference in the success rate of the correction task.

Table 1 summarizes the results of our tests comparing success rate in the monitoring task, and success and timeout rates in the text correction task. The upper part shows the results of the comparison of the slow, fast and single conditions and illustrates the effect of pace variation and peripheral task addition. The lower part shows the results of the comparison of dual-task and single task blocks and illustrates the learning effect.

3.5 Results Analysis

Our results show that the addition of the peripheral monitoring task caused a drop in the success rate of the central task. Subjective data gathered from the post-test questionnaires showed that participants perceived this negative effect. While we found no significant difference in the success rate of the central task between the slow and fast conditions, the timeout rate increased with the pace of snapshots. Four participants reported through the questionnaires that they adapted their text correction strategy according to their free time and tried to schedule their glances towards the peripheral display so as to minimize performance loss on the central task. Overall, these results suggest that increasing the pace of snapshots caused participants to subtly change their text correction strategy. The increase of the timeout rate with the pace of snapshots suggests that participants allocated more attention to the peripheral monitoring task, although not to the point where it would have significantly affected central performance.

Increasing the pace of snapshots led to a significant drop in the success rate of the peripheral task. Data from the questionnaires further indicates that this performance drop was not perceived by seven out of the twelve participants. This suggests that participants did not allocate enough attention to the peripheral task in the fast condition. Combined with the fact that no significant difference was found in the success rate of the central task between the slow and fast conditions, this suggests that participants did not observe the peripheral task priority emphasized in the instructions they received; although the attention required by the peripheral task increased with the pace of snapshots, participants contained this task at the periphery.

Results of the tests on group and order show no effect on both tasks, suggesting that groups A and B are not intrinsically different. Comparison of blocks in the dual-task conditions shows a significant positive change in the success rate of the central task over time, which suggests a learning effect. Participants indeed reported through the questionnaires that they continued to learn how to perform both tasks and to multitask them after the training period. However, comparison of the blocks in the single condition doesn't show any change in the success rate of the central task, which suggests that the learning effect observed on the central task was probably related to better multitasking and not better task performance. Comparison of blocks in the dual-task conditions also shows a significant positive change in the success rate of the peripheral task over time. But in this case, it is unclear whether participants learned to better perform the monitoring task or again benefited from improved multitasking skills.

4 Discussion

> "*When our periphery is functioning well we are tuned into what is happening around us, and so also to what is going to happen, and what has just happened. (...) The periphery connects us effortlessly to a myriad of familiar details.*"
> M. Weiser and J. S. Brown, The coming age of calm technology, 1996

> "*There ain't no such thing as a free lunch*"
> P. Dos Utt, TANSTAAFL: a plan for a new economic world order, 1949

As far as snapshot-based peripheral displays are concerned, our experiment contradicts Weiser and Brown's dream of an effortless connection to the periphery and rather support Dos Utt's idea: the addition of the peripheral monitoring task caused a drop in the success rate of the central task. This corroborates in the particular case of social awareness the findings of Maglio and Campbell who also saw a negative effect on the performance of a central task when adding a peripheral one [17], and contradicts the results of Plaue and Stasko who saw no effect on their central task execution [43]. But as explained by McFarlane and Latorella, understanding human performance with interruptions is a complex problem [20]. As an example, Mark et al. recently showed that interrupted work can sometimes be performed faster, at the cost of a higher workload, more stress, higher frustration, more time pressure and effort [44]. It was not the case in our experiment. More work is certainly needed to better understand the induced costs and benefits of the various kinds of peripheral displays.

Training has long been known as a way to reduce the negative effects of interruptions [45]. As we explained, participants of our experiment reported that they continued to learn how to perform both tasks and to multitask them after the training period, the latter being confirmed by our quantitative results. This corroborates the results of Wickens and Damos that indicated that timesharing skills improve with practice, reducing the interference between the peripheral and the central task [37]. The cognitive resources involved in the peripheral task can also modulate its effects. As an example, the peripheral task used by Maglio and Campbell [17] was a ticker memorization task, and no difference was found among the displays (tickers) in memorability. In our case, the success rate of the peripheral task was indeed affected by the pace of the snapshots. A third factor that can also modulate the effects of the peripheral task is the interaction modalities it uses, which can more or less conflict with those of the central task [46, 20]. In our case, both tasks were intrinsically visual.

We believe the effects we observed of peripheral update pace on attention allocation are not limited to our experimental conditions. These effects can probably be generalized to other conditions, like awareness tasks rather than monitoring tasks. According to McCrickard et al. [18], an awareness question consists in asking participants to recall some information that had been displayed, as in our task. Plaue and Stasko explain that *motivation* makes the difference between awareness and monitoring tasks [43]. In that sense, our peripheral task was a (motivated) monitoring task. But we argue that in both motivated and non-motivated situations, the effort required to maintain peripheral attention is influenced by the update rate of the display. As we said, according to Kahneman [32], motivation is an important factor that influences attention capacity and allocation policy. In the case of (non-motivated) awareness, the peripheral task might simply be ignored by participants in which case no noticeable interference on the central task would probably be found.

Results from our experiment show that increasing the pace of snapshots increased the attentional cost of the peripheral task. Reducing the pace should reduce distraction and thus help keep the task in the background. However, it is unclear whether further increasing the pace would end up promoting the peripheral task to the foreground, or whether users would manage to contain it at the periphery, at the price of lowering performance. Future studies should investigate this. In any case, from a broader perspective, precise control over the update rate of a peripheral display might be a good way of both ensuring that the associated task remains in the background and, at times, initiate a gradual attention shift.

5 Conclusion and Future Work

Monitoring a peripheral display while performing another task has a cognitive cost. This paper reported on a quantitative experiment that we conducted to determine the effect of a snapshot-based peripheral monitoring task on subjects' attention in a dual-task situation. Our results show that the addition of the peripheral task caused a drop in the success rate of the central task. They also suggest that the increase in pace of snapshots caused participants to change their strategy for the central task and allocate more attention to the peripheral one, not enough to maintain peripheral performance but also not to the point where it would affect central performance. Overall, our work suggests that peripheral communication pace subtly influences attention allocation in dual-task situations.

While Romero et al. suggest that communication systems designers should strive for a better balance between distraction, awareness, and screen resources [30], we believe that users should be given the opportunity to negotiate this balance together. Deliberately raising the cognitive cost of a communication can be interpreted as an increased interest in it. By varying the pace of a communication, a user might hope gaining remote people's attention more easily and inciting them to engage a little further. As pointed out by Tang [28], current communication systems leave very little room for this kind of negotiation. Our work suggests that systems should allow the initiator to decide how important and salient a communication is, and not only the recipient. Pousman and Stasko also suggested that designers might start building systems supporting a range of notification levels and not just one [12]. We plan to create new communication systems prototypes to further explore these design spaces.

Acknowledgements

This work has been supported by France Télécom R&D as part of the DISCODOM project (2005-2008). We wish to thank our experiment participants and Wendy Mackay, Olivier Chapuis and the anonymous reviewers for their feedback.

References

1. Whittaker, S.: Rethinking video as a technology for interpersonal communications: theory and design implications. International Journal of Human-Computer Studies 42(5), 501–529 (1995)
2. Dourish, P., Adler, A., Bellotti, V., Henderson, A.: Your place or mine? Learning from long-term use of audio-video communication. Computer Supported Cooperative Work 5(1), 33–62 (1996)
3. Dourish, P., Bly, S.: Portholes: supporting awareness in a distributed work group. In: Proceedings of CHI 1992, pp. 541–547. ACM Press, New York (1992)
4. Weiser, M., Brown, J.S.: Designing calm technology. PowerGrid Journal 1(01) (1996)
5. James, W.: The principles of Psychology. Henry Holt (1890)
6. Hudson, S.E., Smith, I.: Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems. In: Proceedings of CSCW 1996, pp. 248–257. ACM Press, New York (1996)
7. Zhao, Q., Stasko, J.T.: Evaluating Image Filtering Based Techniques in Media Space Applications. In: Proceedings of CSCW 1998, pp. 11–18. ACM Press, New York (1998)
8. Boyle, M., Edwards, C., Greenberg, S.: The effects of filtered video on awareness and privacy. In: Proceedings of CSCW 2000, pp. 1–10. ACM Press, New York (2000)
9. Greenberg, S., Kuzuoka, H.: Using digital but physical surrogates to mediate awareness, communication and privacy in media spaces. Personal and Ubiquitous Computing 3(4), 182–198 (1999)
10. Fogarty, J., Hudson, S.E., Atkeson, C.G., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J.C., Yang, J., Yang, J.: Predicting human interruptibility with sensors. Transactions on Computer-Human Interaction 12(1), 119–146 (2005)
11. Avrahami, D., Hudson, S.E.: Responsiveness in instant messaging: predictive models supporting inter-personal communication. In: Proceedings of CHI 2006, pp. 731–740. ACM Press, New York (2006)
12. Pousman, Z., Stasko, J.T.: A taxonomy of ambient information systems: four patterns of design. In: Proceedings of AVI 2006, pp. 67–74. ACM Press, New York (2006)
13. Itti, L.: Visual salience. Scholarpedia 2(9), 3327 (2007)
14. Strong, R., Gaver, B.: Feather, scent and shaker: supporting simple intimacy. In: Proceedings of CSCW 1996, pp. 29–30. ACM Press, New York (1996)
15. Brave, S., Ishii, H., Dahley, A.: Tangible interfaces for remote collaboration and communication. In: Proceedings of CSCW 1998, pp. 169–178. ACM Press, New York (1998)
16. Dey, A.K., de Guzman, E.: From awareness to connectedness: the design and deployment of presence displays. In: Proceedings of CHI 2006, pp. 899–908. ACM Press, New York (2006)
17. Maglio, P.P., Campbell, C.S.: Tradeoffs in displaying peripheral information. In: Proceedings of CHI 2000, pp. 241–248. ACM Press, New York (2000)

18. McCrickard, D.S., Catrambone, R., Stasko, J.T.: Evaluating animation in the periphery as a mechanism for maintaining awareness. In: Proceedings of Interact 2001, pp. 148–156. IOS Press, Amsterdam (2001)
19. Campbell, C.S., Maglio, P.P.: Segmentation of display space interferes with multitasking. In: Proceedings of Interact 2003, pp. 575–582. IOS Press, Amsterdam (2003)
20. McFarlane, D.C., Latorella, K.A.: The scope and importance of human interruption in human-computer interaction design. Human-Computer Interaction 17(1), 1–61 (2002)
21. Dabbish, L., Kraut, R.: Controlling interruptions: awareness displays and social motivation for coordination. In: Proceedings of CSCW 2004, pp. 182–191. ACM Press, New York (2004)
22. Avrahami, D., Hudson, S.E.: QnA: augmenting an instant messaging client to balance user responsiveness and performance. In: Proceedings of CSCW 2004, pp. 515–518. ACM Press, New York (2004)
23. Muller, H.J., Rabbitt, P.M.A.: Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. Journal of Experimental Psychology: Human Perception and Performance 15(2), 315–330 (1989)
24. Clark, H.H.: Using Language. Cambridge University Press, Cambridge (1996)
25. Gaver, W., Moran, T., MacLean, A., Lövstrand, L., Dourish, P., Carter, K., Buxton, W.: Realizing a Video Environment: EuroPARC's RAVE System. In: Proceedings of CHI 1992, pp. 27–35. ACM Press, New York (1992)
26. Roussel, N., Gueddana, S.: Beyond "Beyond being there": towards multiscale communication systems. In: Proceedings of Multimedia 2007, pp. 238–246. ACM Press, New York (2007)
27. Birnholtz, J.P., Gutwin, C., Hawkey, K.: Privacy in the open: how attention mediates awareness and privacy in open-plan offices. In: Proceedings of Group 2007, pp. 51–60. ACM Press, New York (2007)
28. Tang, J.C.: Approaching and leave-taking: negotiating contact in computer-mediated communication. Transactions on Computer-Human Interaction 14(1), a. 5 (2007)
29. McEwan, G., Greenberg, S.: Supporting social worlds with the community bar. In: Proceedings of Group 2005, pp. 21–30. ACM Press, New York (2005)
30. Romero, N., McEwan, G., Greenberg, S.: A field study of Community Bar (mis)-matches between theory and practice. In: Proceedings of Group 2007, pp. 89–98. ACM Press, New York (2007)
31. Birnholtz, J.P., Gutwin, C., Ramos, G., Watson, M.: OpenMessenger: gradual initiation of interaction for distributed workgroups. In: Proceedings of CHI 2008, pp. 1661–1664. ACM Press, New York (2008)
32. Kahneman, D.: Attention and effort. Prentice-Hall, Englewood Cliffs (1973)
33. Wickens, C.D.: Processing resources in attention. In: Varieties of Attention, pp. 63–102. Academic Press, London (1984)
34. Yerkes, R., Dodson, J.D.: The relation of strength of stimulus to rapidity of habit formation. Journal of Comparative Neurological Psychology 18, 469–482 (1908)
35. Sullivan, L.: Selective attention and secondary message analysis: A reconsideration of broadbent's filter model of selective attention. Quarterly Journal of Experimental Psychology 28(2), 167–178 (1976)
36. Duncan, J.: Divided attention: The whole is more than the sum of its parts. Journal of Experimental Psychology 5(2), 216–228 (1979)
37. Damos, D.L., Wickens, C.D.: The identification and transfer of timesharing skills. Acta Psychologica 46(1), 15–39 (1980)

38. Spelke, E., Hirst, W., Neisser, U.: Skills of divided attention. Cognition 4(3), 215–230 (1976)
39. Treisman, A., Davies, A.: Divided attention to ear and eye. In: Kornblum, S. (ed.) Attention and performance IV, pp. 101–117. Academic Press, London (1973)
40. Dabbish, L., Kraut, R.: Coordinating communication: awareness displays and interruption. In: Proceedings of CHI 2003, pp. 786–787. ACM Press, New York (2003)
41. Bartram, L., Ware, C., Calvert, T.: Moticons: detection, distraction and task. International Journal of Human-Computer Studies 58(5), 515–545 (2003)
42. Moles, A., Kandel, L.: Application de l'indice de Flesch à la langue française. Cahiers d'Etudes de Radio-Télévision (19), pp. 252–274 (1958)
43. Plaue, C., Stasko, J.T.: Animation in a peripheral display: distraction, appeal, and information conveyance in varying display configurations. In: Proceedings of GI 2007, pp. 135–142. ACM, New York (2007)
44. Mark, G., Gudith, D., Klocke, U.: The cost of interrupted work: more speed and stress. In: Proceedings of CHI 2008, pp. 107–110. ACM Press, New York (2008)
45. Hess, S.M., Detweiler, M.C.: Training to reduce the disruptive effects of interruptions. In: Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting, pp. 1173–1177 (1994)
46. Storch, N.A.: Does the user interface make interruptions disruptive?: a study of interface style and form of interruption. In: CHI 1992 extended abstracts, p. 14. ACM Press, New York (1992)

Is the Writing on the Wall for Tabletops?

Nadia Pantidi, Yvonne Rogers, and Hugh Robinson

Computing Department
Open University
Milton Keynes
MK7 6AA, UK

Abstract. We describe an ethnographic study that explores how low tech and new tech surfaces support participation and collaboration during a workshop breakout session. The low tech surfaces were post-it notes and large sheets of paper. The new tech surfaces were writeable walls and a multi-touch tabletop. Four groups used the different surfaces during three phases: i) brief presentation of position papers and discussion of themes, ii) the creation of a group presentation and iii) a report back session. Participation and collaboration varied depending on the physical, technological and social factors at play when using the different surfaces. We discuss why this is the case, noting how new shareable surfaces may need to be constrained to invite participation in ways that are simply taken for granted because of their familiarity when using low tech materials.

Keywords: Collaboration, equitable participation, interactive tabletops, workshops.

1 Introduction

Workshops are commonplace in education and business, providing an occasional forum for groups of people to discuss and share ideas. Typical goals of a workshop include establishing a new area, identifying key themes, and addressing new concerns. The extent to which a workshop is successful in achieving its aims, however, depends on a number of factors. These include ensuring everyone has an opportunity to participate and express their views [21]. One method that is commonly employed is to break out into smaller groups for certain sessions of the workshop, with the specific intention of providing more opportunities for participants to contribute in a smaller setting. The sub-groups typically work on a specific task, such as identifying a set of requirements for a new research area and then reporting back to the larger group to share their ideas.

A common practice used to facilitate participation within a breakout group is to provide a variety of low tech materials that group members can externalize and share their ideas on (e.g., post-it notes, sheets of papers and flipcharts). Group members can write their own ideas, opinions or questions down first on individual post-its and then compile them into a meaningful group structure using a shared flipchart. Recently, ‘new tech’ surfaces have been experimented with, such as interactive tabletops and special writeable walls. An assumption is that because of their shareable properties (e.g., they can be interacted with simultaneously by those standing or sitting around

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 125–137, 2009.

them) they will promote more equitable participation and greater awareness of what others are saying and doing. However, it is unclear if this is the case for meetings, where the goal is for groups of individuals to collectively generate and explore ideas together, assemble and agree on them and subsequently present the outcomes to a wider group.

Our research is concerned with how different kinds of surfaces support the creation and sharing of ideas in breakout groups, where equitable participation is a desirable outcome. Specifically, we describe an ethnographic study that explored how different low and new tech surfaces were used by breakout groups to work together, initially to discuss ideas, then generate themes and finally to present these back to the wider meeting. Our study revealed marked differences between the groups in terms of the participants' contribution and the way they progressed with the task. We discuss these findings in terms of how the different technologies mediate collaboration, in terms of how participants approach and appropriate the different surfaces, the extent of participation, the nature of the turn-taking that took place and how the technologies can be exploited by gatekeepers in a group.

2 Background

Low tech materials, such as paper and post-it notes, are frequently used to support collaborative activities, especially early design phases [2]. They are also used for collaborative planning in medical settings [e.g., 10, 13], the military [19] and control centres [16]. Their advantages are well known and have been written about extensively; for example, paper is cheap, lightweight, intuitive to use, robust and can be easily manipulated and shared [3, 32]. Physical artifacts, such as annotated notes, are also highly tangible. They may be used as external thinking props to explain a principle, an idea or a plan to others in the group. The persistence and ability to manipulate physical artifacts may also result in a diversity of options being explored in a group setting.

Software tools have been developed for use in structured meetings to facilitate and record the discussions that occur such, as those that take place in the planning of NASA's mission control. They have been used to integrate notes, visualise various forms of pre-stored data and to annotate plans and ideas. Examples include Debategraph [7] and Compendium [33] which capture and represent information and argumentation. Interactive whiteboards have also been used to support focused meetings around a single issue [25]; ongoing, continuous work across a host of domains [22] and small group collaboration in informal meetings [30].

Several studies investigating how interactive displays are used by people in collaborative settings have examined the context in which they are initially approached. The presence of social and visual cues that invite people to use them has been found to be critical [1, 5, 6, 11]. How comfortable people feel knowing that their actions and their effects on a shared display are highly visible to others in a group setting may also affect their willingness to participate. Such self-consciousness can deter people from taking part in a group activity [4]. For example, people needed constant encouragement and demonstration to interact with a shared whiteboard system called Plasma Poster [5]. User studies have also shown how actual use of interactive displays can be quite different from their intended use [e.g., 4, 15, 18, 24]. When the shared display is vertical, as is typically the case with whiteboards, people work around them in somewhat unnatural and uncomfortable ways [9]. This can make it difficult for participants to take over control or for them to hand it over to others. The effect can be sub-optimal communication of ideas and activity progression. One strategy people use is to monitor the current speaker or user and explicitly orient themselves in order to find when it is possible to take the floor [31, 26]. Similarly, the speaker or user can help by making his/her withdrawal visibly noticeable.

In contrast to vertical displays, horizontal tabletop displays have been designed so that more than one person can use them simultaneously [27]. Examples include Microsoft's Surface [20] and MERL's DiamondTouch [8]. An assumption is that they can support more flexible forms of collaboration because they provide more opportunities for interaction and the switching of control between group members. People seem more willing to interact with multi-touch tabletops in the presence of others despite their actions being highly visible and observable by others in the group [11]. This may be because these kinds of shared horizontal surfaces lure people to touch them without feeling intimidated or embarrassed by the consequences of their actions. The familiar and lightweight action of touching a surface may also make it easier for people to take part in a social/public setting. User studies have shown how groups of people, new to tabletops, find it easy and enjoyable when sharing and assembling of sets of digital images for a variety of collaborative tasks [12, 28]. However, studies of whether this new form of shared interaction promotes more equitable participation and less awkward ways of taking control in small groups have been mixed. Marshall et al. [17] and Rogers et al. [28] have found that multi-touch tabletops supported more equitable participation in terms of group member's digital interactions but that their verbal contributions remained uneven. Dominant people continued to speak the most while shy and non-native speakers spoke the least.

Ideally, these new kinds of shareable technologies should be designed to allow groups comfortably and easily to access, create, interact with and move digital content in an equitable and free-flowing manner. However, the extent to which these goals can be met depends on a number of factors, including how obvious it is to the group members to know what to do at an interface and how to take turns to progress with a collaborative task. Our research is concerned with how different technologies (tabletop, wall, flipcharts, post-its) invite people to generate, add, manipulate or structure ideas from a shared display. In particular, we are interested in whether groups use these new tech surfaces in equitable and collaborative ways and how this compares with the appropriation of more familiar low tech materials.

3 The Study

To examine how low tech and new tech surfaces affect group participation, especially the creation, sharing and organization of ideas we conducted an ethnographic study of an academic two-day workshop, where the organizers had decided to provide four breakout groups with different kinds of technologies. The workshop took place in a technologically augmented space, Qspace, an adaptable space with an emphasis on creativity. To begin, we describe our methodological approach, the setting and the way the breakout session was setup in terms of the groups and the surfaces.

3.1 The Method

The methodological approach employed in our study was ethnographic, involving observations of naturally occurring activities and semi-structured interviews. The workshop was planned and run by organizers who we were acquainted with but our involvement with the event was purely as participant observation. Hence, we had no influence over the design of the breakout groups or the materials used, but observed it through the eyes of those who inhabited it. Following the ethnographic approach, all activities observed were considered as 'strange'; no preliminary hypotheses were formed beforehand and no particular feature of use or interaction was given a priori significance. The collected data consisted of field notes, photos, documents and audio recordings from participant observation and semi-structured interviews. We provide incidents, activities and practices within their context, to emphasise that their meaning is properly comprehended within the appropriate social context.

3.2 The Setting

Qspace is a Centre of Excellence in Teaching and Learning (CETL) space that was created as part of Higher Education Funding Council for England (HEFCE) joint initiative between two universities. It was envisioned to be a technologically rich but not technology-driven learning space that was quite different from the constraints of a traditional lecture hall or seminar room. It provides both physical and technological resources that can be used in a variety of configurations with the purpose of supporting innovative creative processes. The space has been designed to create a relaxed atmosphere with all white moving walls, in which multi-coloured LEDs, curtains, bean bags, numerous projectors and PLASMA screens have been placed. Some of the walls have been customized as writeable surfaces. For the workshop, a DiamondTouch tabletop was placed into the space by the workshop organisers. Qspace, therefore, is a rather untypical space for an academic workshop, especially for paper presentations. On the other hand, the breakout sessions could be designed flexibly by using different configurations of technology and physical space.

3.3 Groups and Materials

During initial planning meetings between the managers/facilitators of QSpace and the workshop organisers an explicit suggestion was made to introduce a more 'creative' activity such as building conceptual representations with low tech materials (clay, polyethylene, cardboard), as this would match the QSpace 'owners' criteria of use[1]. It is also an activity that has been found to be successful in the past for that space. The workshop organisers incorporated this suggestion by providing groups with different surfaces to work on during a breakout session. From their perspective, they were

---

[1] Before the facilitators allow any activity to take place in the space, they were going through a list of criteria that the activity should match up to some degree. These criteria ensure that the activity is well suited for the space and complies with Qspace's evaluation pre-requirements in terms of its funding.

Table 1. Types of surface used in each breakout group

	Group A	Group B	Group C	Group D
Surface	Post-its	Large sheets of paper	Writeable wall	Tabletop

interested in whether and how the different surfaces would affect collaboration within the groups.

For the breakout session that was observed during the study, four groups were created (A, B, C and D), each consisting of 9-10 participants. The organisers assigned one person in each group to initially present the group's position papers in a summarized version as a starting point for a discussion. The next phases involved them creating a presentation of the main discussion themes as a concept map that they would later present back to the other groups. Each group was given either low or new tech surfaces to create their presentation on (see Table 1). Group A was provided with different sized post-its and coloured markers; Group B was provide with large sheets of paper and coloured markers; Group C was given coloured markers for writing on the special wall surface and Group D was provided with the tabletop that displayed a simple concept map tool for typing in their ideas and arranging into a particular structure. In terms of our classification of technology, Groups A and B were considered as low tech and Groups C and D as new tech.

4 Findings

Observations and interviews with the participants revealed differences between the groups in terms of how they approached and appropriated the different surfaces and the turn-taking that took place. A main finding was that the extent and kind of collaboration and participation varied across the four groups. Roles, such as scribe, were adopted by or assigned to individuals and were determined to some extent by the kind of surface used. Below we describe the findings in terms of, firstly, how participants approached and appropriated the different surfaces; secondly, the turn-taking that took place; thirdly, the extent to which equitable participation took place and fourthly, whether gatekeeping emerged in a group and the role the technology played in enabling it to materialize.

4.1 Approaching and Appropriating the Different Surfaces

In the low tech groups, the post-it notes were handed out early on to the participants. As expected, Group A wrote their own ideas down on the big post-its. The group shared the big post-its and used them as if they were small ones (figure 1a) rather than using them as a poster surface and adding small post-its notes to it. One of the members (A1) commented: "*From the [participant's] presentation [of the position papers] we recognized the general problem area that we were to address. We decided that we should choose some aspects to address more specifically and we did sort of an informal brainstorming (…) For that, each one of us wrote on a post-it [a big one] issue or idea about the area (…)*" .

In Group B, after only a few minutes of discussion, one of the participants (B1) suggested they start working on their presentation while still discussing the ideas. He also requested from the organizers that they be given post-its rather than begin writing their ideas on the large sheets of paper. This suggests that he wanted them to write their own ideas down first and then move them onto a shared display – rather than try to write them straight onto the paper. It also meant they could move the ideas around which would not be possible if written directly onto the paper. As the session progressed, B1 placed one of the large pieces of paper on the floor in the middle of the group and invited the others to start writing on the post-its important issues that arose from the discussion.

Both Groups A and B took a photo of their shared map of post-it ideas and presented it using one of the data projectors to the rest of the groups in the report back session. Hence, the second phase was integrated with the first, meaning that the groups were discussing and creating their ideas at the same time. Prior familiarity with the paper-based surfaces resulted in them having no problems writing their ideas on them and then assembling them into a shared structure. All of them readily understood that the post-its were intended for writing a comment, idea or question on and afterwards making it public by reading it or sticking it on the flip chart paper or the wall.

In Group C one participant began by projecting slides of the main points of the position papers on the wall. The group then discussed their ideas. Only towards the end of the session did one or two of them approach the writeable wall and write up their ideas. Hence, the assembling and explication of ideas was left to the end of the session rather than being a shared integrated activity throughout – as was the case in the low tech groups. One of the participants commented: "*(…) it was as if we were mesmerized by the screen and the presentation and didn't want to move [on to the next stage]*". The participants who wrote on the wall also wrote lists of points rather than create a concept map. One participant noted that their initial ideas were too small and illegible from far away and so she rewrote these to be neater and bigger on an adjacent wall so that a larger audience could read them. Thus the transition from the generation and sharing of ideas in the sub-group to the reporting back to the whole workshop was much more cumbersome than in the low tech groups. Part of the problem may have been because the group was provided with the same surface for the generation and the reporting back of ideas. It was only when they realized this that they wrote them up in a form that could be viewed from a wider audience.

For Group D, a simple interactive concept map was developed for use in conjunction with the tabletop surface. The participants were instructed on how use it by the person who had designed the concept map software (D1). He explained how to create nodes, type text and build up a concept map. It was observed that the group chose not to sit around the tabletop at the beginning (see figure 1b); instead, they approached the tabletop later on only after they were explicitly asked by D1 to do so. Hence, similar to Group C, and in contrast with the familiar post-its and paper sheets, the participants appeared to be more reluctant in the beginning to use the tabletop concept map. It required scaffolding and encouragement by the software developer. D1 then took a screen shot of the resultant concept map and projected it onto the wall during the subsequent feedback session.

Fig. 1a. Writing on the big post-it notes

Fig. 1b. Group D sitting next to the tabletop

4.2 Turn-Taking in Conversation and Writing Up of Ideas

Turn-taking was evident in all groups during the discussion phase, although the extent of contribution differed between the groups. In each group, the occasional speaker paused after making his/her point, providing an opportunity for another participant to comment or make another point. Some breakdowns, such as interrupting the speaker and simultaneous speaking were observed but were not seen as out of the ordinary.

When it came to writing up the ideas generated in the discussion for the subsequent presentation to the whole workshop there was some evidence of turn-taking in Group A (post-its) but less so in the other groups. Members of Group A tried to place the post-its on the wall but they did not stick on the surface. They then resorted to writing their ideas up on the wall, where they all took turns to write up what was on the post-its by using pen markers. In contrast, Group C split into three sub-groups that worked on separate themes and where only one person from each sub-group took charge of writing their notes on the wall (figure 2b). There was some negotiation between the members of each sub-group about the written content but the members did not take turns in the writing. A similar situation was observed in Group B where one person took charge of making a concept map during the whole session and where no turn-taking occurred (figure 2a). A pattern of parallel work was also observed in Group D where four members worked largely by themselves to type in text and create nodes (figure 2c). Even though three of them stepped back to enable the other group members to have a go, only one out of the five remaining members took the opportunity to do so. It appears that in contrast to Group A, these three groups adopted a way of working that is more commonly observed in "whiteboard interaction scenarios"[2].

4.3 Extent of Contribution

The extent each member contributed to the task varied across groups and phases of the task. In Group A, all members of the group contributed when discussing and

[2] By "whiteboard interaction scenario" is meant a situation where one person stands in front of the whiteboard and writes on it while the rest of the group makes only verbal contributions. In such cases, it is possible that some group members may not contribute at all.

Fig. 2a. B1 in charge of creating the concept map

Fig. 2b. The wall used as a big whiteboard

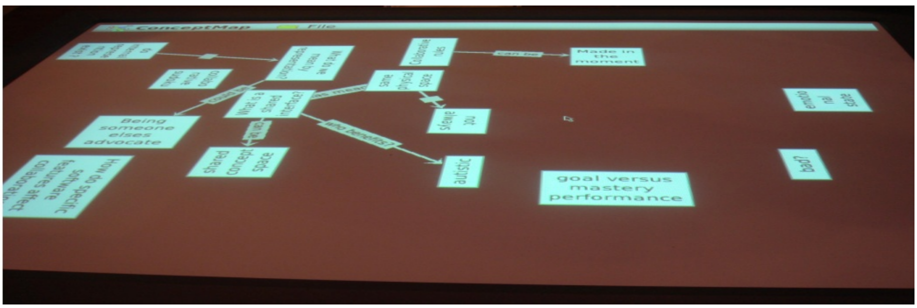


Fig. 2c. The tabletop concept map: the nodes that are in the centre are the ones created by D1 as starting point while the directionality of the nodes typed indicates parallel individual work

writing on the post-its and then when writing these up on the walls. In Group B each participant contributed to the task by writing at least one post-it but not all took part in placing the post-its and assembling the concept map. Moreover, it was observed that not all members participated equally in the discussion. This observation is also supported by a record of verbal contributions that one member of the group sketched during the session, where each participant was represented by a node and was checked every time he/she contributed to the discussion. The majority of the participants were checked once or twice and only two had multiple checks for verbal contributions. Only two participants were observed to take a leading role: B1 and B2. B1 can be described as the "doer" and B2 as the "talker" based on their contributions.

In Group C, although one participant was in charge of writing and no physical turn-taking took place, our observations indicate that there was constant verbal negotiation between the members of the sub-groups about the content of the final result. The participant holding the marker neither wrote exclusively his/her ideas nor was a simple listener to the suggestions of others. In some instances the group was even negotiating about details such as the appropriate phrasing and not just the content of the text. In Group D, all members contributed during the discussion. When the move to the tabletop took place, however, only five out of the nine people followed; the other four continued with the discussion during the rest of the session. Of the 4 participants who moved to the tabletop, only two remained for any length (one being D1). One of them when asked later, why, provided a range of reasons: "*I had to move from one area to another and we [meaning the four participants] were discussing… I don't even remember what we were discussing… what I remember is that I felt intimidated and that it [adding text to the surface] is more permanent than the discussion <pause> I also felt more conscious that I was the most junior one*". A possible factor that appears to have affected the extent to which members of this group made a contribution was their status in the group and how comfortable they felt about making public their ideas in front of others, who knew more. This did not appear to happen in the other groups – presumably because they were less intimidating to use in the presence of others.

4.4 Gatekeeping

The inequality in participation across the groups can also be explained in part by gatekeeping, where *"actors have control over key sources or avenues of opportunity. Gatekeepers exercise control at and during key phases"* [12]. This was most evident in Groups B and D, but not in Groups A and C. B1 and D1 took control of the materials that restricted others from participating at various stages. B1 was always in charge of the post-its and the concept map creation (figure 2a). He took over the manipulation of these almost from the beginning of the session by handing out the post-its, suggesting how to work, taking the poster paper and placing it on the floor and positioning himself next to it. His assertive behavior constrained the form of collaboration that could take place for the rest of the session. He implicitly assigned roles to the other participants that were limited to verbal and not physical contribution at least in terms of the concept map task.

D1 was in a different position than B1, as he had programmed the concept map software. From this point of view, it was clear that he possessed knowledge on the tool that no other participant had and in that sense he was an obvious gatekeeper. However, the rest of the group who were familiar with tabletops could have overcome this obstacle with some help from him. This did not happen and other more implicit signs contributed to his being established as a gatekeeper and thus inhibiting other participants from interacting with the tabletop. At some point during the discussion phase, D1 left the group circle and moved to the tabletop where he started typing text and creating nodes for the concept map (figure 3a). As D1 explained later in an interview he wanted to make some examples that would be of use when he explained to the group how to use the software. However, this action, viewed from the other participants, implied his right and ownership over the tabletop that they did not share. Similarly, during the concept map task, D1's presence and guidance was continuous even when he had supposedly stepped down; he was leaning towards the tabletop commenting, assisting and making sure everything was working properly (figure 3b). Moreover, when everyone from the group had finished working he added some final touches by re-organising parts of the concept map.

It was also noted how the gatekeepers in Groups B and D maintained that role in the report back session to the rest of the workshop. Hence, their role of gatekeeper was able to be preserved, transferred and extended beyond the physical or digital manipulation of the tools.

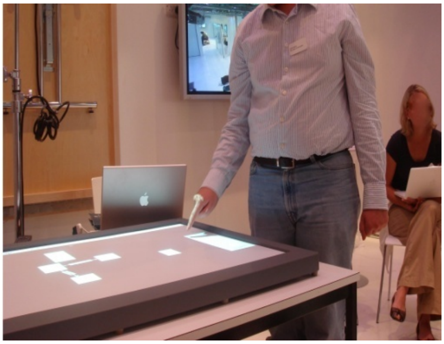


Fig. 3a. D1 typing on the tabletop while the rest of the group is discussing

Fig. 3b. D1 leaning towards the tabletop

5 Discussion

The findings from our ethnographic study reveal how the type of surface that is provided to support the generation and sharing of ideas can affect the nature of the collaboration and interactions that take place during a breakout group. While all the groups were able to complete the task using a particular technology, it was the ubiquitous, low tech post-it note that turned out to be the most democratic medium; each member of the group was able to contribute their ideas through using it, which they then took turns to combine into a shared representation as written points on the wall. The two novel technologies – the writeable wall and the tabletop – were used in less equitable ways. In particular, they were not used until the end of the breakout session and then only by a few members of the group. One reason for this is that the actions of writing on a wall or entering text into existing software node structures are unfamiliar to the participants and which they feel awkward initially approaching.

Although writing is a very familiar activity, writing on a wall is not. It is also unnatural for a group of 8-10 people to stand around together at a wall and write together. Hence, it is not surprising that Group C broke up into smaller groups and where each allocated a scribe to write up their ideas on the wall. However, the result was that three streams of ideas generated by the sub-groups were not integrated in the way they were in the low tech groups. Similarly, although the participants in Group D were familiar with multi-touch surfaces, they had not created a concept map together by entering their ideas into a software tool. As it was not possible for all group members to take part at the same time at the tabletop (it was constrained for only 4 people to use) it was, likewise, not surprising that they broke up into two sub-groups. However, the sub-groups did not switch roles despite those at the tabletop inviting them to do so, mid-way through. This resistance to switching places when using such technologies is redolent of the fixing of roles that occur when groups work around whiteboards [24]. Typically, one person is in charge and the rest contribute verbally and not in equal terms, such as was found in Whittaker and Schwarz's [34] study; the advantage of using a physical wall compared to a PC-based software wall was attributed to it being collaboratively constructed whereas the data was entered into the software tool and thus 'owned' exclusively by the project manager.

In these settings, certain members of a group can unwittingly or explicitly turn into gatekeepers. As found in Group D, one individual took control throughout in terms of the creation of the ideas, their organization and then the subsequent presentation of them to the rest of the workshop. However, the same also happened in one of the low tech groups, where B1 took control of the situation and effectively became a leader. Hence, this suggests that individuals may exploit the affordances of any kind of technology to take charge and control regardless of whether it is new or low tech.

The extent to which low or new tech tools promote collaboration in group settings, therefore, would appear to be a function of the perceived or implied ownership and whether gatekeeping can be prevented. While previous research has shown how shared surfaces, namely tabletops can increase group awareness and encourage more equitable participation in design tasks and collaborative games [e.g., 14], it is less clear whether such benefits can accrue in more open-ended meetings, such as breakout sessions, where the participants are not necessarily familiar with the technology, the outputs of the task or each other.

Another factor that can affect participation levels is the size of the group. In our study, the workshop organizers had decided to split the workshop into 4 breakout groups, comprising 9-10 people. With the exception of Group A, the large size made it difficult for all group members to use the surface they had been allocated. Smaller groups at the tabletop or the writeable wall might have enabled them to remain as one unit and maintain awareness of each other rather than having to further divide into sub-groups. While keeping the size of the groups small in breakout sessions is considered optimal [21], the workshop organizers chose not to do so here, partly because of the constraints and the availability of the technology in the QSpace. It would be interesting to see whether more equitable participation and less gatekeeping would happen with smaller sized breakout groups.

An important consideration that needs further research is the permanent and public nature of the interactions that result when using different technologies. Whereas a post-it note written on can be easily discarded, a wall cannot. It needs a special kind of cloth and liquid solution to erase it, which often is not ready to hand. This can affect how easily or comfortable someone is when writing on a wall. Likewise, a node filled in on a software tool is perceived as being permanent, especially if the participation does not know how to delete it. The act of writing is also private on a post-it note and the person can decide whether to make it public or not; whereas at the tabletop it is public from the beginning.

In sum, while the writing is not on the walls for tabletops (or writeable walls for that matter) it is important to consider what their added value is in more open-ended meetings where there may be considerable social awkwardness and uncertainty as to what a group member can and cannot do. Having a facilitator in a group may help group member's overcome these barriers although he/she may, too, find it difficult to resist turning into a gatekeeper.

Acknowledgments. We thank all the participants, the facilitators of the space and the organizers of the workshop for making that study possible.

References

1. Agamanolis, S.: Designing displays for Human Connectedness. In: O'Hara, K., Perry, M., Churchill, E., Russell, D. (eds.) Public and situated displays: Social and interactional aspects of shared display technologies, pp. 309–335. Kluwer Academic, Dordrecht (2003)
2. Bailey, B.P., Konstan, J.A., Carlis, J.V.: Demais: Designing multimedia applications with interactive storyboards. In: Proceedings of ACM Multimedia, pp. 241–250. ACM Press, New York (2001a)
3. Barthelmess, P., Kaiser, E., Lunsford, R., McGee, D., Cohen, P., Oviatt, S.: Human-centered collaborative interaction. In: Proceedings of the 1st ACM international Workshop on Human-Centered Multimedia, HCM 2006, Santa Barbara, California, USA, October 27, 2006, pp. 1–8. ACM, New York (2006)
4. Brignull, H., Izadi, S., Fitzpatrick, G., Rogers, Y., Rodden, T.: The introduction of a shared interactive surface into a communal space. In: Proceedings of CSCW 2004: Conference on Computer Supported Cooperative Work, ACM, New York (2004)
5. Brignull, H., Rogers, Y.: Enticing people to interact with large public displays in public spaces. In: Proceedings of INTERACT 2003: Ninth IFIP TC13 International Conference on Human-Computer, IOS Press, Zurich (2003)
6. Churchill, E.F., Nelson, L., Denoue, L., Girgensohn, A.: The Plasma Poster Network: Posting multimedia content in public places. In: Proceedings of INTERACT 2003: Ninth IFIP TC13 International Conference on Human-Computer. IOS Press, Zurich (2003)
7. Debategraph, `http://debategraph.org`
8. Dietz, P., Leigh, D.: DiamondTouch: A multi-user touch technology. In: Proceedings of the 14th Annual ACM Symposium on User interface Software and Technology. UIST 2001, pp. 219–226. ACM, New York (2001)
9. Eden, H., Hornecker, E., Scharff, E.: Multilevel design and role play: experiences in assessing support for neighbourhood participation in design. In: Proceedings of DIS 2002, pp. 387–392. ACM, London (2002)
10. Gorman, P., Ash, J., Lavelle, M., Lyman, J., Delcambre, L., Maier, D.: Bundles in the Wild: Managing Information to Solve Problems and Maintain Situation Awareness. Library Trends 49(2), 266–289 (2000)
11. Grasso, A., Muehlenbrock, M., Roulland, F., Snowdon, D.: Supporting communities of practice with large screen displays. In: O'Hara, K., Perry, M., Churchill, E., Russell, D. (eds.) Public and situated displays: Social and interactional aspects of shared display technologies, pp. 261–283. Kluwer Academic, Dordrecht (2003)
12. Hammersley, M., Atkinson, P.: Ethnography: Principles in practice, 3rd edn. Routledge, London (2007)
13. Heath, C., Luff, P.: Technology in Action. In: Pea, R., Brown, J.S., Heath, C. (eds.) Learning in Doing: Social, Cognitive and Computational Perspectives, Cambridge University Press, Cambridge (2000)
14. Hornecker, E., Marshall, P., Dalton, N., Rogers, Y.: Collaboration and Interference: Awareness with Mice or Touch Input. In: Proceedings of ACM CSCW Conference 2008, pp. 167–176. ACM Press, New York (2008)
15. Huang, E.M., Mynatt, E.D., Trimble, J.P.: When design just isn't enough: The unanticipated challenges of the real world for large collaborative displays. Personal and Ubiquitous Computing, Special Issue on Ubiquitous Computing in the Real World 11(7), 537–547 (2007)
16. Mackay, W.E.: Is paper safer? The role of flight strips in air traffic control. ACM Transactions on Computer- Human Interaction 6(4), 311–340 (1999)

17. Marshall, P., Hornecker, E., Morris, R., Dalton, N., Rogers, Y.: When the fingers do the talking: A Study of Group Participation With Varying Constraints to a Tabletop Interface. In: Proceedings of IEEE Tabletops and Interactive Surfaces, pp. 37–44. IEEE, Los Alamitos (2008)
18. McDonald, D.W., McCarthy, J.F., Soroczak, S., Nguyen, D.H., Rashid, A.M.: Proactive displays: Supporting awareness in fluid social environments. ACM Transactions on Computer- Human Interaction, 14(4), Article 16 (2008)
19. McGee, D.R., Cohen, P.R., Wu, L.: Something from nothing: Augmenting a paper-based work practice with multimodal interaction. In: Proceedings of the Conference on Designing Augmented Reality Environments, Helsingor, Denmark, April 12-14, 2000, pp. 71–80. ACM Press, New York (2000)
20. Microsoft Surface, http://www.microsoft.com/SURFACE/index.html
21. Muller, G.: Workshop How to To (2008), http://www.gaudisite.nl/workshophowtopaper.pdf
22. Mynatt, E.D., Igarashi, T., Edwards, W.K., LaMarca, A.: Flatland: New dimensions in office whiteboards. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the CHI Is the Limit, Pittsburgh, Pennsylvania, United States, May 15-20, 1999, pp. 346–353. ACM, New York (1999)
23. O'Hara, K., Glancy, M., Robertshaw, S.: Understanding collaborative play in an urban screen game. In: Proceedings of CSCW 2008: The 2008 ACM Conference on Computer Supported Cooperative Work. ACM, San Diego (2008)
24. Pantidi, N., Robinson, H.M., Rogers, Y.: Can technology-rich spaces support multiple uses? In: Proceedings of HCI2008: The 22nd annual conference of Interaction, a specialist group of the British Computer Society. BCS, Liverpool (2008)
25. Pedersen, E.R., McCall, K., Moran, T.P., Halasz, F.G.: Tivoli: An electronic whiteboard for informal workgroup meetings. In: Proceedings of the INTERACT 1993 and CHI 1993 Conference on Human Factors in Computing Systems, Amsterdam, The Netherlands, April 24 - 29, 1993, pp. 391–398. ACM, New York (1993)
26. Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., Oulasvirta, A., Saarikko, P.: It's mine, don't touch!: Interactions at a large multi-touch display in a city centre. In: Proceedings of CHI 2008: Conference on Human Factors in Computing Systems. ACM, Florence (2008)
27. Preece, J., Rogers, Y., Sharp, H.: Interaction Design: Beyond Human-Computer Interaction, 2nd edn. John Wiley & Sons, New York (2007)
28. Rogers, Y., Lim, Y., Hazlewood, W.R., Marshall, P.: Equal Opportunities: Do Shareable Interfaces Promote More Group Participation than Single User Displays? To Appear in Human Computer Interaction (2008)
29. Rogers, Y., Lindley, S.: Collaborating around vertical and horizontal displays: which way is best? Interacting With Computers 16, 1133–1152 (2004)
30. Russell, D.M., Sue, A.: Large interactive public displays: Use patterns, support patterns, community patterns. In: O'Hara, K., Perry, M., Churchill, E., Russell, D. (eds.) Public and situated displays: Social and interactional aspects of shared display technologies, pp. 3–18. Kluwer Academic, Dordrecht (2003)
31. Sacks, H., Schegloff, E.A., Jefferson, G.: A simplest systematics for the organization of turn taking in conversation. Language 50(4), 696–735 (1974)
32. Sellen, A.J., Harper, R.H.R.: The Myth of the Paperless Office. MIT Press, Cambridge (2002)
33. Sierhuis, M.: Collaboratively Modeling Mission Control at NASA (2006), http://news.kmi.openac.uk/rostra/news.php?r=55&t=2&id=20#
34. Whittaker, S., Schwarz, H.: Meetings of the Board: The Impact of Scheduling Medium on Long Term Group Coordination in Software Development. Computer Supported Cooperative Work 8(3), 175–205 (1999)

Investigating the Effect of Hyperlink Information Scent on Users' Interaction with a Web Site

Nikolaos Tselios, Christos Katsanos, and Nikolaos Avouris

Human-Computer Interaction Group, Department of Electrical and Computer Engineering,
University of Patras, GR-265 00 Rio Patras, Greece
{nitse,ckatsanos}@ece.upatras.gr, avouris@upatras.gr

Abstract. In the study presented in this paper we investigate how variations of information scent of hyperlinks of a webpage influence users' behavior in terms of attention-focusing, confidence, effectiveness and efficiency, while exploring a website. In the reported study, 19 participants completed eight different tasks associated with eight simplified websites. Analysis of the results showed that even small differences in the target-link's information scent can substantially affect users' performance, distribution of attention and confidence. The study contributes to the related literature by quantifying the impact of even small differences in the target-link's scent on users' success ratio, time for first click, confidence and distribution of attention. In addition, a scent threshold value was identified, below which all the measured variables were substantially affected, and thus the link could be characterized as of "weak scent".

Keywords: Information scent, eye tracking study, Latent Semantic Analysis.

1 Introduction

Information scent has been defined as the users' assessment of the value of following a particular hyperlink in a webpage, based on its perceived semantic association with their goal [1], [2], [3], [7]. Users have lower success rates and require more time to complete their tasks when they are presented with navigation options of weak information scent compared to high information scent [2], [3], [7].

Thus, systematic study of the impact of scent variations on users' performance is important. Such quantitative studies are also needed to identify the threshold value over which one can characterize a link as of high scent, which is often determined in an arbitrary way currently. Furthermore, eye tracking user studies, such as [4] and the one reported in this paper, can provide insights on the effect of information scent on users' distribution of attention while selecting links in a webpage.

In the eye-tracking user study presented in this paper, we investigated the impact of slight variations of the target-link's information scent on users' effectiveness, efficiency, confidence and attention-focusing while selecting a link in a webpage. The paper is organized as follows: First, the method of the study is presented, followed by the analysis of results. Finally, we discuss the conclusions, implications and future directions of the presented research.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 138–142, 2009.

2 Method of the Study

Eight webpages were built that contained navigation menus of actual websites, related to specific tasks (e.g., find a destination, buy a specific object, etc.)[1]. The textual descriptions of the tasks for these pages were produced in a way that created eight levels of the correct-link's information scent, measured using an estimation of semantic similarity between the correct-link's text and the description of the task.

This estimation was provided by Latent Semantic Analysis (LSA, [6]), a technique that relies on the statistical elaboration of corpora to estimate the semantic similarity between two texts. Overall, the LSA index of the correct-links of the eight pages ranged from 0.7 to 0. The rest of the hyperlinks had LSA values below 0.15 (Table1). Two experts in information architecture provided judgments of the information scent of all links, as a means of preliminary evaluation of the validity of LSA in the context of the given tasks. Miller and Remington [8] involved three judges to assess the information scent of hyperlinks in their experimental webpages. Thus, two judges were deemed enough to evaluate the validity of LSA in our case.

Subsequently, 19 participants, 5 female, with a mean age of 24, all proficient in English (the language of the webpages), were asked to perform the eight menu selection tasks using a typical Web browser. Their behavior was monitored. All navigation menus were presented in a single-column format and included eight link-options. The order of presentation of the websites and the order of links in the homepages were randomized to avoid serial order effects. First, users were presented with a goal-description screen. Next, they were presented with the associated menu and were asked to select a link as they would normally do. Afterwards, they subjectively rated how confident they felt for their selection on a 1-10 scale. An unobtrusive 17'' Tobii T60 eye tracker with minimum fixation duration set to 100ms was used to record users' eye movements during this process.

Four measures of users' behavior were gathered: a) *Effectiveness* = correct first-click percentage, b) *Efficiency* = time for first-click, c) *Confidence* = users'

Table 1. Users' metrics of link-selection behavior collected

Task[1]	**Correct-link Scent**	**2nd-best link Scent**	**Success (%)**	**Click time (sec)**	**Confidence (1-10)**	**Number of observations**		**Observations Duration (sec)**	
						Correct link	***All links***	***Correct link***	***All links***
1	0.70	0.09	89%	12	8.9	2.5	13.8	2.3	11.0
2	0.59	0.14	74%	21	8.1	3.6	18.1	5.2	19.7
3	0.50	0.10	68%	19	7.9	1.9	15.9	2.6	16.5
4	0.40	0.01	68%	17	6.8	3.0	19.3	2.4	15.8
5	0.30	0.11	68%	14	7.5	3.4	17.3	3.0	14.6
6	0.21	0.06	37%	28	6.3	3.9	24.2	3.7	27.0
7	0.11	0.08	21%	22	6.3	3.4	19.6	4.3	19.5
8	-0.01	0.08	11%	31	6.1	4.2	26.6	4.7	30.0

[1] Tasks and menus used can be found at http://hci.ece.upatras.gr/Tselios_et_al_INT2009.

confidence for first-click, d) *Efficiency of visual search* measured in terms of number of observations on each link and duration of observations on each link.

3 Analysis and Results

The values of users' effectiveness, efficiency, confidence and visual search efficiency for each task, as measured by the selected dependent variables are presented in Table 1. Scent values of the correct link and the link with the highest scent value among the rest are also shown.

Correlation analysis indicated a very high degree of correlation between the scent value of the correct link and the observed participants' success ratio (Fig. 1a, r=0.945, p<.001), average time for click (Fig. 1b, r=-0.725, p<0.05), subjective average confidence (Fig. 1c, r=0.934, p<0.01), and number of observations on all links (r=-0.839, p<0.05). A trend was also observed between the scent value of the correct link and the average number (r=-0.688, ns) and duration (r=-0.436, ns) of observations on the correct link. Furthermore, there seems to be a critical scent value for the correct link between 0.21 and 0.30. Below this value, the measured variables appear to be substantially affected and therefore, the link can be characterized as of weak scent.

Fig. 1d presents heatmaps of participants' total duration of fixations in each webpage. In higher scent pages, attention was mainly focused in the area containing the

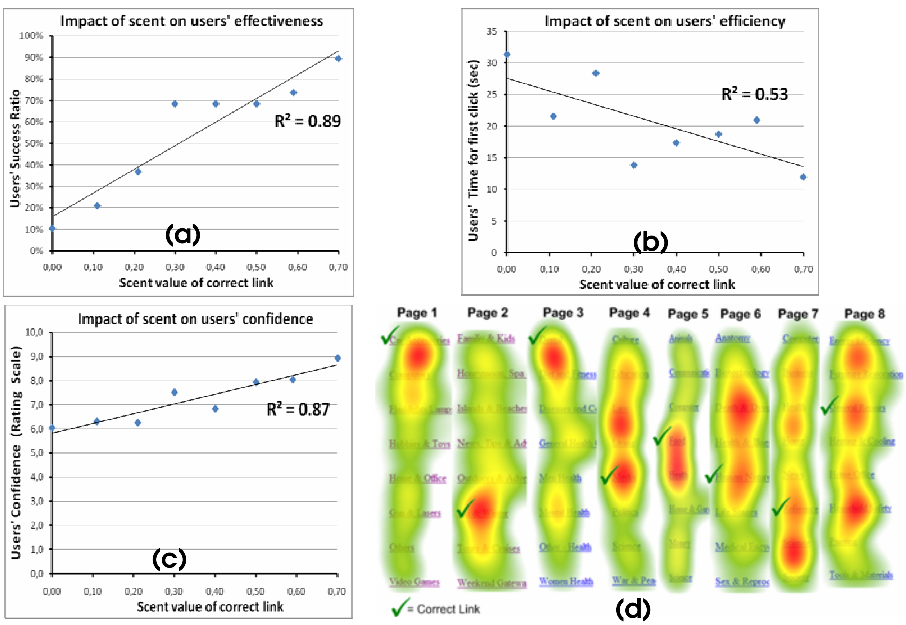


Fig. 1. Scatter plots of the scent values of the correct link for each task and the observed participants' success ratio (a), average click time (b), and average confidence (c). Heatmaps of participants' total duration of fixations for each webpage (d).

correct link, indicating a focused and efficient search. In webpages 4-5 there is a slight distribution of attention across some other links indicating a potential lack of confidence regarding the correct choice, which can be also observed in the users' average self-reported confidence (see Table 1). In webpages 6 to 8, in which the scent value of the correct link dropped substantially, the attention was distributed across the majority of links indicating increased confusion. In these three tasks, users' success ratio was on average 23%, whereas in the rest five it was 73%.

4 Conclusions

The results of an eye-tracking study are reported in which the impact of eight, slightly different levels of information scent was measured on various aspects of users' link-selection behavior. When scent increased, users were gradually getting more effective and efficient, had significantly more focused attention-allocation patterns (also reported in [4]) and reported higher levels of selection confidence. It was observed that even small differences in the target-link's scent could substantially affect the collected measures of users' behavior. For instance, link selection time in task 1 was 12 seconds, while in task 2 was 21 seconds, with a target-link's scent value of 0.70 and 0.59 respectively. Pirolli [7] also found high correlation (r=0.949) between the observed and theoretically predicted (generated by Monte Carlo simulations using an ACT-Scent model) frequency of users' link-selections, very close to our results.

In addition, it was found that the LSA threshold above which a link could be characterized as of *high scent* is between 0.21 and 0.30, also empirically reported in [1]. Such identification of threshold levels could significantly extend the utility of automated evaluation tools, such as ACWW [1] and ISEtool [5].

The findings of this study may be extended and further validated by extensive testing with a richer set of tasks that take into account other influencing factors (e.g. layout, number of available links), and other methods to compute information scent (such as human rating, PMI-IR, GLSA). These constitute future research aims.

References

1. Blackmon, M., Kitajima, M., Polson, P.: Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In: Proceedings of CHI 2005, pp. 31–40. ACM Press, New York (2005)
2. Card, S., Pirolli, P., Wege, M., Morrison, J., Reeder, R.W., Schraedley, P., Boshart, J.: Information scent as a driver of Web behavior graphs: results of a protocol analysis method for Web usability. In: Proceedings of the CHI 2001, pp. 498–505. ACM Press, New York (2001)
3. Chi, E., Rosien, A., Supattanasiri, G., Williams, A., Royer, C., Chow, C., Robles, E., Dalal, B., Chen, J., Cousins, S.: The Bloodhound Project: Automating Discovery of Web Usability Issues using the InfoScent Simulator. In: Proceedings CHI 2003, pp. 505–512. ACM Press, New York (2003)

4. Habuchi, Y., Kitajima, M., Takeuchi, H.: Comparison of eye movements in searching for easy-to-find and hard-to-find information in a hierarchically organized information structure. In: Proceedings of ETRA 2008, Savannah, Georgia, pp. 131–134. ACM Press, New York (2008)
5. Katsanos, C., Tselios, N., Avouris, N.: InfoScent evaluator: a semi-automated tool to evaluate semantic appropriateness of hyperlinks in a web site. In: Proceedings of OZCHI 2006, Sydney, Australia, pp. 373–376. ACM Press, New York (2006)
6. Landauer, T.K., Dumais, S.: A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review 104(2), 211–240 (1997)
7. Pirolli, P.: Information Foraging Theory: Adaptive Interaction with Information. Oxford University Press, USA (2007)
8. Miller, C.S., Remington, R.W.: Modeling information navigation: implications for information architecture. Human-Computer Interaction 19(3), 225–271 (2004)

Interpersonal Privacy Management in Distributed Collaboration: Situational Characteristics and Interpretive Influences

Sameer Patil[1], Alfred Kobsa[1], Ajita John[2], Lynne S. Brotman[2], and Doree Seligmann[2]

[1] Department of Informatics, University of California, Irvine, Irvine CA 92697, USA
{patil,kobsa}@uci.edu
[2] Collaborative Applications Research, Avaya Labs, Basking Ridge NJ 07920, USA
{ajita,lynne,doree}@avaya.com

Abstract. To understand how collaborators reconcile the often conflicting needs of awareness and privacy, we studied a large software development project in a multinational corporation involving individuals at sites in the U.S. and India. We present a theoretical framework describing privacy management practices and their determinants that emerged from field visits, interviews, and questionnaire responses. The framework identifies five relevant situational characteristics: issue(s) under consideration, physical place(s) involved in interaction(s), temporal aspects, affordances and limitations presented by technology, and nature of relationships among parties. Each actor, in turn, interprets the situation based on several simultaneous influences: self, team, work site, organization, and cultural environment. This interpretation guides privacy management action(s). Past actions form a feedback loop refining and/or reinforcing the interpretive influences. The framework suggests that effective support for privacy management will require that designers follow a socio-technical approach incorporating a wider scope of situational and interpretive differences.

Keywords: Privacy, Awareness, Distributed collaboration, Privacy management.

1 Introduction

High-speed networking has enabled organizations to form teams of geographically dispersed knowledge workers. Using such teams offers several advantages: opportunities to employ talent from across the world, possible expansion to distant markets, and savings in operating expenses and travel costs. At the same time, the physical [1, 2] and/or temporal [3] separation in such teams poses several challenges when compared with conventional co-located collaboration. A crucial factor that impacts distributed collaborative activities is the impoverishment of interpersonal interactions (professional as well as social) and, consequently, reduced awareness of the activities of others [2, 4]. Awareness information, however, is instrumental for structuring and coordinating one's own work in light of interdependencies [5]. As a remedy, various interpersonal awareness and interaction systems (IAIS) have been built to serve the

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 143–156, 2009.

needs of distributed collaborators (see e.g., [5–7]). There is a growing trend in organizations to deploy IAIS to support work activities and interactions of their employees [8–10]. The estimated economic impact of IAIS usage may be to the tune of several billion dollars. IBM alone, for example, estimates more than $110 million in annual savings from the use of IAIS [11].

Designers of these systems face the significant challenge of having to attend simultaneously to users' awareness as well as privacy needs [12, 13]. Insufficient attention to either need could undermine the usage of a system. When users cannot effortlessly reconcile privacy desires with awareness needs, they may not utilize the system's full potential [7, 14, 15]. Moreover, insufficient attention to privacy aspects may evoke strong user backlash [16] that can lead to minimal use, or even abandonment, of the system. If this were to happen, organizations stand to lose their investments in IAIS. Moreover, the companies that design and build these systems, as well as their clients, face the prospect of longer-term damage to trust and credibility [17, 18]. In order to build and deploy IAIS that are sensitive to the privacy expectations and behaviors of collaborators, we must first understand how knowledge workers engaged in distributed collaboration currently manage their privacy. This paper reports the findings from a study carried out with this goal.

2 Related Work

Prior field studies of geographically distributed teams have been conducted with the primary objective of comparing their practices with those of conventional co-located teams. Therefore, these efforts have attempted to measure efficiency (e.g., [2]), uncover cultural barriers (e.g., [19]), or identify effective resource allocation (e.g., [20]). Our study sets itself apart with its focus on privacy management practices.

Initial findings related to privacy in collaboration were primarily noted "on the side" in studies aimed at evaluating awareness systems. For instance, Dourish [21] characterized privacy controls along a "social-technical continuum." On the social end of this continuum, social pressures and norms are relied upon to prevent system abuse, while on its technical end, technology prevents attempted misuse. Social controls are likely to work well only within small and relatively well-knit communities [21, 22]. Recently, studies of awareness systems have targeted privacy as the primary object of investigation and unveiled a number of factors that affect users' privacy judgments. These include the user's relationship with the information recipient, the purpose and usage of requested information, the context, and the sensitivity of the content [15, 17, 18, 23–25]. In contrast, our study does not limit itself to specific systems or scenarios; we examined privacy management as it occurs in distributed collaboration. We treated IAIS as "part of the circumstance within which those concerns are formulated and interpreted" [26].

Prior research has also generated models to conceptualize privacy in networked multimedia environments. Adams [17] proposed an empirically derived privacy model in which a user's privacy perceptions are based on the sensitivity, receiver, and usage of information. However, her subjects were not seasoned team workers but university students/staff and conference attendees. Palen and Dourish [26] presented a conceptual analysis that describes privacy as a process regulating the boundaries of disclosure, identity and temporality. This process is both dynamic (i.e., shaped by personal and collective experiences and expectations) and dialectic (i.e., under continuous boundary negotiation). Lederer et al. [27] further listed system properties, actor relations, and information types as relevant interrelated dimensions that are important for the privacy management process. The framework we describe in this paper supports and extends these insights by grounding them in an empirical study of geographically distributed collaborators.

3 Study Setting

We studied a software development project (henceforth Project X) at a large multinational telecommunications company. At various stages, Project X included anywhere between 80 to 130 contributors spread across at least five different locations: four in the U.S. (in three different time zones) and one in India. The software developed by Project X comprised eighteen interdependent modules integrated into a single release. Project X releases were arranged in cycles of three to four months. Each new release incorporated new features and fixed bugs from the previous releases. During each cycle, the management team of Project X tracked the progress of the ongoing cycle and planned out the details of the next cycle.

A variety of contributors worked on Project X: managers at different hierarchical levels, software architects, systems engineers, software developers, testers, source code management (SCM) support staff, and internal and external customers. The number of people at each location varied from a handful to more than thirty. It should be noted that team[1] membership was independent of geographical location. In other words, it was not uncommon for members of a team to be distributed across different sites. In addition, a limited amount of telecommuting was common, but only in the U.S. Several forms of telecommuting were practiced: telecommuting for part of the day, telecommuting for the entire day, telecommuting on an as-needed basis etc. In a couple of cases, project members worked mostly from their “virtual office” at home, coming in to the workplace only as needed.

Project X required frequent intra-module as well as inter-module collaboration in all phases: design, architecture, coding, testing, integration, and maintenance. Collaboration was also required when the employees responsible for a module changed; the new individual(s) needed to be briefed. Some collaborative activities spanned the entire project scope (for example, when management planned project activities, tracked progress, made adjustments, and set strategic goals using estimates and forecasts). The SCM unit provided services to Project X for managing its source code repositories. Representatives from other software projects that would use the Project X platform were also consulted on requirements and feedback. For collaboration, Project X members used not only face-to-face meetings but also a variety of technical communication, coordination and awareness tools. These included email, Instant Messaging (IM), shared calendars, Microsoft Project®, telephones (desk phones as well as mobile phones), the SCM system, shared document repositories, and various Internet-based conferencing tools (such as Virtual Network Computing (VNC) or Microsoft NetMeeting®).

[1] We loosely define a team to be a group of individuals supervised by the same manager.

While most developers, testers and lower-level managers worked solely on Project X, other contributors were engaged in multiple projects. The fraction of time they spent on Project X ranged from 3-4% to more than 50% and was partly dependent on the current stage in the release cycle. For instance, in some stages a systems engineer would spend nearly all of his or her work time on Project X, while during the rest of the cycle, his or her effort would be an order of magnitude lower.

4 Methodology

Our research objective was to investigate how geographically distributed collaborators reconcile the often conflicting needs of awareness and privacy. To avoid biasing the participants, the advertised goal of the study was to investigate collaborative work practices. We did not impose a definition of privacy on the study participants. Rather, we asked them to explain what "privacy" meant to them in the context of their work and work practices. Given the lack of a universal definition for privacy, our intention was to avoid biasing the participants or confining them to a specific view of privacy. Instead, we sought to uncover the various contextual meanings of privacy for the participants, and the associated behaviors and practices aimed at satisfying privacy needs. An initial hour-long conversation with the Head of Project X helped us gain a basic understanding of the project and formulate a plan for conducting the study. Thereafter, we used the following methods:

- **Non-participant Observation.** Our exploration started with non-participant observation of the meetings of the Project X management team. These meetings, which involved managers from all project sites and all hierarchical levels, were used to formulate the detailed plan for the next release cycle of the project. These observations improved our understanding of the organization and activities of the project, which in turn guided the development of the other aspects of the study that are described below.
- **Site Visits.** During the first phase of the study, the researcher was based at a Project X site on the East Coast of the U.S., and visited three of the other sites, in mid-U.S., on the U.S. West Coast, and in India, each for about a week. At every location, interviews were conducted with project members (see below), and site-specific factors such as architecture, layout, work practices and culture were documented. We also interviewed project contributors from the remaining U.S. site – also located on the East Coast – while they were visiting the first-mentioned site.
- **Semi-structured Interviews.** At each visited site, we conducted semi-structured interviews lasting about 90 minutes. The questions were divided into three main themes: work practices, awareness and privacy needs, and desired enhancements to collaborative tools. In total, we interviewed fifty-two project members across the five sites. Interviewees were chosen in such a way that the different job functions at the various sites were covered to the extent possible.
- **Online Questionnaire.** Based on key insights from the above activities, we formulated an online questionnaire. It aimed at probing deeper into some of the aspects we learned from the earlier activities and at attaining broader coverage across Project X by reaching those whom we had not been able to interview. Additionally, we used questions from the literature to measure privacy attitudes and practices in the

domain of consumer privacy [28, 29], and collected demographic information. The questionnaire was distributed to all individuals involved with Project X at that time (roughly 125). We obtained 90 valid responses (response rate of 74%) which included responses from 30 of our original 52 interviewees.

Transcripts of the interviews were analyzed using the grounded theory approach [30]. Categories that emerged from open coding were further refined into ten higher-level categories by selective coding. Then, axial coding was employed to identify the relationships between these higher-level characteristics resulting in a framework that illustrates how privacy management operates in the collaborative work context.

5 Privacy Management Framework

Fig. 1 shows the framework that emerged from our analysis. It presents privacy management practices in collaborative work as dependent not just on a number of situational characteristics but also on a hierarchy of personal interpretive influences that an individual applies to the situation. As can also be seen in **Fig. 1**, the interpretation applied by an individual to the situation at hand leads to privacy management action(s), or lack thereof. Moreover, these actions themselves form a feedback loop that contributes to the reinforcement and/or evolution of the interpretive influences over time. Privacy is known to be a context-dependent and highly personal concept [26, 31]. The framework captures the former aspect in terms of the situational characteristics, and the latter through the interpretive influences. (The interpretive influences are related to the "identity" boundary whereas the situational characteristics encompass the "disclosure" and "temporal" boundaries described by Palen and Dourish [26].)

5.1 Situational Characteristics

Our analysis revealed five key situational characteristics that interviewees deemed important when reconciling awareness and privacy needs.

- **Issues.** The details of the issue(s) at hand were instrumental in judging aspects such as confidentiality, urgency, audience, or communication medium. These judgments, in turn, affected privacy management actions.

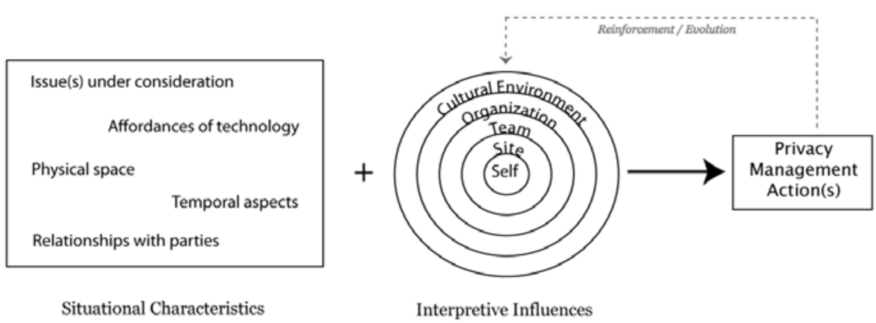


Fig. 1. Privacy management described in terms of interpretive influences applied to situational characteristics

- **Relationships.** This characteristic refers to the nature of relationships – formal as well as informal – that existed among the various parties involved in a particular situation.
- **Temporality.** Two temporal aspects impacted a situation. The first was the time of day in one's own time zone as well as in those of one's collaborators. The second was the temporal *extension* of the present action(s) into the future in the form of archives, logs, records, or people's memories.
- **Technology.** As described in Section 3, collaborative activities utilized a host of technologies. The affordances as well as the limitations of a system constrained which actions it could support, and in what manner [32].
- **Space.** This characteristic refers to the physical space in which work was carried out. It includes the design and layout of workspaces and work sites, and also of other locations from which individuals worked (such as homes, conference rooms, offices of others, cars, and hotels).

The situational characteristics Issues, Relationships and Technology map to and refine, respectively, the concepts System Properties, Actor Relations, and Information Types proposed by Lederer et al. [27]. Similarly, Issues, Relationships and Temporality subsume Information Sensitivity, Receiver and Usage from Adams' model [17]. Even though we discussed each of the characteristics separately above, a multiple of these often came into play in any given situation. In other words, all of them were subject to interpretation simultaneously. This is illustrated by the supporting examples in Section 6.

5.2 Interpretive Influences

While the characteristics described above set the stage, the privacy management action(s)[2] of each individual further depended on his or her interpretation of the situation. We identified five major influences that guided this interpretation.

- **Self.** Individuals drew upon their personal disposition and characteristics when interpreting a situation.
- **Team.** The practices, norms or policies of one's team were also crucial in deciding how situations were interpreted. We observed that the impact of this influence was dependent on factors such as the length of time the team members had worked together, the degree of work coupling, and the management style of the team leader.
- **Site.** This influence refers to practices and local factors that were unique to a given site. For example, the typical practice at the U.S. West Coast site was to arrive at the workplace later in the morning than at the other U.S. sites.
- **Organization.** The multi-national corporation was the umbrella uniting the different sites. It influenced interpretation by providing policies and norms, a shared sense of identity, as well as a shared technical infrastructure for carrying out work activities.
- **Cultural Environment.** The cultural environment external to the organization in which one was embedded also influenced how situations were interpreted. The large differences between the privacy preferences and practices at the India site

[2] The framework also treats inaction, i.e., deciding not to act, as an action.

compared with those at the U.S. sites is one of the salient findings of the study (described in [33]). These differences were partially attributed to the impact of the cultural environment.

We also noted that the above influences could be arranged in a hierarchy beginning with the most inward influence (self), and growing progressively outward toward the larger environment one is embedded in. The interrelationship between the influences also needs to be emphasized. For instances, differences between sites can be attributed not just to local factors (such as the history of the site, the interactions among local colleagues, the weather etc.) but also to organizational factors (such as policies or infrastructural variations) as well as cultural influences. In order to isolate the contribution of an individual influence, it may be necessary to make comparisons. For example, the cultural environment is unlikely to be a major contributor to differences among sites within the U.S.

6 Supporting Examples

In this section, we present four frequently encountered situations in which collaborators are faced with reconciling awareness and privacy needs: making communication[3] choices, handling interruptions, working from home, and dealing with urgent matters. The following subsections describe how the framework presented above explains the privacy management actions that we encountered in these situations (the labels in italics indicate the applicable situational characteristics and interpretive influences). Although we discuss the four separately, it should be noted they are often interrelated. For instance, communication choices may need to be made when handling interruptions, or an interruption may need to be handled to deal with an urgent situation.

6.1 Making Communication Choices

Our interviewees indicated that privacy concerns impacted their communication choices, i.e., what was communicated to whom and how. Several privacy management practices occurred in the context of communication, such as self-censorship, medium switching, location switching, etc. When engaging in these practices, interviewees reported taking into account the situational characteristics outlined in our framework. Thus, privacy management actions depended on factors such as the importance, sensitivity, and confidentiality of the matter being communicated (*issues*), hierarchical as well as social relationships with the audience (*relationships*), temporal considerations such as whether or not the communication could be archived and accessed at future times (*temporality*), the richness of expression afforded by a communication technology (*technology*), and the presence of others around oneself who may come to know about the communication (*space*).

[3] "Communication" refers not just to the contents of verbal or written conversations, but also to other more implicit interaction aspects, such as IM status, calendar entries, or code submission time stamps, which can serve to communicate awareness information about actions, availability, etc.

Commensurate with our framework (see **Fig. 1**), the link between a particular communication situation and the corresponding privacy management actions is established by the interpretive influences. For instance, some individuals preferred IM over email for short messages (*self, issues*). Interviewees also reported that their communication choices were influenced by the norms in the team, and the management style of their manager. For example, some participants reported backchannel IM conversations with other team members in order to present a "uniform voice" during meetings with others, while others reported being available by mobile phone at all times because of managerial expectations (*team*). Site-specific influences were also observed: due to cubicle environments at some of the sites, private phone conversations necessitated reserving conference rooms or stepping outside the building (*site*). General organization-wide influences such as the shared technical infrastructure, and corresponding communication norms (e.g., on sharing one's calendar or accessing the calendars of other employees), also shaped how privacy was managed in communication (*organization*). Moreover, external factors unrelated to work, such as family, commuting conditions, and the cultural background, influenced how situations were interpreted (*cultural environment*).

Interviewees expressed vastly different expectations of privacy for written vs. non-written communication (*technology*). Written communication (which included email) was composed with care, and was often self-censored (*issues, relationships, self, organization*). The fact that it could be saved, or be forwarded beyond the original recipient, was often taken into account (*temporality*). The choices of whom to copy, or leave out, were made deliberately (*issue, relationships*). Interviewees also applied their interpretations to the variety of non-communicative functions of written communication, such as its role as an individual memory aid, knowledge management archive, organizational record, and instrument of accountability. On the other hand, non-written communication (such as phone calls and face-to-face meetings) was more informal and impromptu (*relationship, technology, team, site*). In terms of privacy, it was sometimes used for discussions deemed too sensitive for the written medium and/or when the individual wished to avoid a written trail (*issue, technology, temporality*). IM fell somewhere in between [34]. Although IM is written communication, a majority of our interviewees treated it as ephemeral and informal. With the exception of a few, who exercised the same caution with IM as with email, the interviewees did not report archiving IMs (in fact, some even claimed not to know that it was possible to save IMs), and assumed that other employees did not save IMs either (*technology, temporality, organization*).

6.2 Handling Interruptions

One characterization of privacy is "freedom to be left alone" [35], i.e. control over access to oneself. Interruptions, which are common in knowledge work [36], have a direct impact in this regard [37]. Our interviewees mentioned several kinds of interruptions, planned as well as unplanned. These included: scheduled meetings (*temporality*), incoming communication (e.g., email, IM, phone) (*technology*), colleagues dropping by one's office (*relationships*), urgent issues that required immediate attention (*issues, temporality*), and lunch breaks (*temporality*). We noted several privacy management

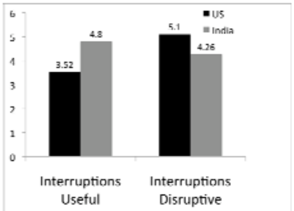


Fig. 2. Contrasting attitudes towards interruptions at the U.S. and India sites

practices for dealing with interruptions: closing the office door (*space*), scheduling "busy" blocks on one's calendar (*temporality, technology*), turning off IM or setting the IM status to "busy" (*technology*), multitasking during conference calls (*temporality, technology*), eating lunch at one's desk (*temporality, space*), working from home (*space*), and working during hours when few others are present (*temporality*).

Interruptions exhibited disruptive as well as useful characteristics. On the one hand, interviewees complained that interruptions took attention away from their current tasks (*issues*), required extra time and effort for refocusing on the original task (*issues, temporality*), and split time into short blocks resulting in the "filler" blocks being spent unproductively (*temporality*). On the other hand, interviewees recognized the value of interruptions for learning about issues that required immediate attention (*issues*), coordinating with colleagues (*issues, relationships, temporality*), taking a break (*temporality*), multi-tasking (*issues, temporality*), interacting with friends and colleagues at the workplace (regarding work as well as non-work matters) (*relationships*), and attending to domestic errands (*issues, space*).

In handling interruptions, privacy management involved applying the interpretive influences to resolve the tension between the disruptive influences of interruptions with their potential usefulness. (Note that merely "mentally processing" an interruption is also disruptive even if the interruption is not dealt with immediately.) For instance, we found that the job function affected preferences regarding how the tension ought to be resolved (*self*). For example, managers recognized greater value in interruptions since managerial duties require them to be available to resolve the issues brought forth by their subordinates [38]. In contrast, developers desired long, uninterrupted time blocks to concentrate on their programming tasks. Based on responses to the online questionnaire, we also uncovered differences between those at the U.S. sites with those in India (*space, site, cultural environment*) [33]. As **Fig. 2** shows, those in India agreed more than those in the U.S. that interruptions are useful (India: 4.80, US: 3.52, $p < 0.0001$). In contrast, workers at the U.S. sites found interruptions disruptive to a larger extent (India: 4.26 US: 5.10 $p < 0.015$).

6.3 Balancing Work and Home

As mentioned in Section 3, it was not uncommon for Project X members to work from home, at least in the U.S.[4] The form of telecommuting ranged from occasionally

[4] Telecommuting was practically non-existent in India due to the lack of adequate domestic technical infrastructure, and demographic and cultural differences (*cultural environment*).

checking email from home to having a permanent home office (*temporality, space*). Most of our interviewees fell somewhere between these two extremes (see Section 3). Moreover, more than half of our interviewees worked from home during late evenings and/or early mornings (*temporality*). This was necessitated by the need to "get work done" in long uninterrupted time blocks (*issues, temporality*), to "catch up" with unfinished tasks (*issues, temporality*), to handle urgent matters (*issues, temporality*), to interact with people in other time zones (*issues, relationships, temporality, technology*), and to build new knowledge and skills. Similarly, several interviewees admitted attending to domestic tasks at the workplace (*issues, temporality, space*). Examples include errands, doctor's appointments, personal email/IM/Web activities, family phone calls, or childcare tasks[5]. Traditionally, the two primary spheres of an individual's life – domestic and professional – have been markedly distinct. For Project X workers, however, the possibilities of getting work done even when away from the office blurred the boundary between home and work, and consequently impacted the traditional notion of the 9-to-5 work day (*temporality, space*). A second factor impacting the notion was the time zone differences with one's collaborators (*temporality*). As a result, what was considered to be "typical" work hours varied from location to location (*site*) and person to person (*self*).

Privacy expectations at home and at work are different. Accordingly, we noted several privacy management practices to balance work and home based on one's interpretation of the telecommuting situation. Outside of the standard business hours, almost all interviewees desired and exercised more control over their availability to others (*issues, relationships*) and over the tasks they worked on (*issues*). For example, during non-business hours some interviewees chose not to sign-in to IM (*technology*), or chose to work on tasks that did not require interaction with others (*issues*). In many cases, such preferences (*self*) applied regardless of one's physical location during these times, i.e., regardless of whether one was working at home or from one's workplace (*space*). For instance, frequent telecommuters reported that during standard business hours, they strived to make themselves available to their collaborators to the same extent that they would at the workplace (*temporality, space, team*). Most interviewees also reported designating a separate room or work space at home (*space*), and trying to separate work and home activities by using separate computers and phones (*issues, technology*).

6.4 Dealing with Urgent Matters

While all the situations discussed above pertain to normal routines, dealing with urgent matters (*issues, temporality*) required deviating from typical privacy expectations and practices. Judgments of urgency were based not only on the matter at hand but also on the parties involved (*relationships*). For instance, a request from one's boss was often assigned higher urgency. Discussions of urgency came up frequently in our interviews; participants stated that "normal" privacy expectations and practices did not apply in urgent situations. In such cases, they engaged in privacy management practices different from their normal preferences (*self*): reorganizing their calendar to

[5] Interestingly, the India site employed a person who was charged solely with attending to the domestic errands of the knowledge workers (e.g., paying bills at different places in town) (*cultural environment*).

devote time to “put out the fire” (*issues, temporality*), limiting their availability for other tasks (issues, temporality), making themselves reachable on their mobile phones (*temporality, space, technology*), answering phone calls at odd hours (*temporality, space, technology*), and ensuring their availability even on vacation (*temporality, space*). Others were expected to make similar adjustments to their privacy expectations. Thus, when urgent matters arose, our interviewees interrupted others (*issues*), preferred a phone call over an email (*issues, technology, temporality*), called the mobile phones of their collaborators (*issues, technology, temporality*), or contacted higher-level managers with whom they would not normally communicate (*issues, relationships*).

In general, the views of all stakeholders (*team, site, organization*) on what should be regarded as urgent were aligned, but not always. The interviewees reported a handful of instances of differing interpretations of urgency. Since urgency led to changed privacy expectations and practices, mismatches in urgency evaluations could pose problems (such as receiving a phone call in the middle of the night for an issue that one does not deem equally urgent as the caller).

7 Implications and Future Work

It is often argued that (a) employees of an organization ought not desire privacy at the workplace, in the interest of organizational needs for transparency and accountability of employee actions, and (b) employees would not harbor privacy concerns from colleagues due to familiarity and shared organizational bonds. In fact, some of our interviewees also expressed similar views initially because they began by characterizing privacy in terms of monitoring by the organization, preventing the disclosure of personal information unrelated to work, or hiding inappropriate activities from the organization and colleagues. Perhaps this resulted from media coverage in other contexts that also spur privacy concerns (e.g., government surveillance, commercial data mining, or personally identifiable data usage by entities such as vendors, companies, universities, hospitals, and governments). In contrast, as the above discussion illustrates, privacy considerations in collaborative work are focused on interpersonal relationships and interactions[6]. Our findings demonstrate that knowledge workers do engage in privacy management, and underscore the need for IAIS that elevate the effectiveness and efficiency with which privacy desires can be reconciled with conflicting collaborative needs.

Users of IAIS use multiple interaction channels; they work in spaces shared with other workers; they multitask; they interact with people with whom they have diverse kinds of relationships; they move between work and home spheres; they need to prioritize their tasks and availability. Our study shows that these situational characteristics are *all* considered (and usually several at the same time) when managing privacy. Therefore, in order to enhance the consideration of privacy needs, IAIS designers will need to take situational characteristics into account. Improvements to IAIS need not be complex nor purely technical. For example, technical enhancements could be

[6] The interviewees and survey respondents attached low significance to the issue of “employer monitoring” of their activities in an institutional sense. Many were, however, significantly more concerned about “managerial monitoring.”

augmented by simple additions such as privacy screens for monitors, or dual monitor setups that allow one to channel interruptions to a secondary screen. It may even be beneficial to co-design organizational policies regarding interpersonal privacy expectations and practices along with the technological developments.

In earlier work [15, 39], we established "impression management" [40] as an underlying cause of privacy concerns. We uncovered impression management practices in the present study as well. A number of our interviewees expressed the desire to "appear professional," "present an appropriate image," "avoid being perceived negatively," and "tailor the message to the audience." We plan a deeper exploration of how to translate impression management considerations into design suggestions for improving privacy management.

8 Conclusion

We reported on a study aimed at understanding the interpersonal privacy management practices of geographically distributed collaborators. A framework that describes these practices emerged from a grounded theory analysis of interviews with members of a geographically distributed project, coupled with qualitative field observations at the project sites. The framework identifies five key situational characteristics (issues, relationships, temporality, technology, and space) which collaborators interpret based on five key influences (self, team, site, organization, and cultural environment). These interpretations guide privacy management practices, which in turn provide feedback for the reinforcement or evolution of the interpretive influences. These privacy management practices often require weighing multiple situational characteristics under various simultaneous interpretive influences. We therefore believe that in order to facilitate more effective and efficient privacy management, collaborative systems ought to allow a broader consideration of situational characteristics and a greater differentiation of interpretive influences than is currently the case. Our findings highlight the deep embedding of collaborative technical systems as part of the situational context. To build privacy management solutions, it is therefore necessary to follow a socio-technical design approach that encompasses aspects such as the physical space in which IAIS are embedded, and the diversity of relationships between the parties whose interactions they mediate.

Acknowledgments. We thank the study participants for their time. We also acknowledge Mihir Mahajan, Xinru Page and Joel Ossher for their comments. This research has been supported by NSF Grant Nos. 0205724 and 0808783.

References

1. Olson, G.M., Olson, J.S.: Distance Matters. Human-Computer Interaction 15(2/3), 139–178 (2000)
2. Herbsleb, J.D., Mockus, A., Finholt, T.A., Grinter, R.E.: An Empirical Study of Global Software Development: Distance and Speed. In: Proc. of ICSE 2001, Toronto, Canada, pp. 81–90. IEEE Press, Los Alamitos (2001)

3. Pankoke-Babatz, U., Syri, A.: Collaborative Workspace for Time Deferred Electronic Co-operation. In: Proc. of GROUP 2003, Sanibel Island, FL, USA, pp. 187–196. ACM Press, New York (1997)
4. Kraut, R., Egido, C., Galegher, J.: Patterns of Contact and Communication in Scientific Research Collaboration. In: Proc. of CSCW 1988, Portland, OR, USA, pp. 1–12. ACM Press, New York (1988)
5. Dourish, P., Bly, S.: Portholes: Supporting Awareness in a Distributed Work Group. In: Proc. of CHI 1992, Monterey, CA, USA, pp. 541–547. ACM Press, New York (1992)
6. Cheng, L.T., Hupfer, S., Ross, S., Patterson, J.: Jazzing up Eclipse with Collaborative Tools. In: Proc. of Eclipse 2003, Anaheim, CA, USA, pp. 45–49. ACM Press, New York (2003)
7. Herbsleb, J.D., Atkins, D.L., Boyer, D.G., Handel, M., Finholt, T.A.: Introducing Instant Messaging And Chat In The Workplace. In: Proc. of CHI 2002,, Minneapolis, MN, USA, pp. 171–178. ACM Press, New York (2002)
8. DiMicco, J., Millen, D.R., Geyer, W., Dugan, C., Brownholtz, B., Muller, M.: Motivations for Social Networking at Work. In: Proc. of CSCW 2008, San Diego, CA, USA, pp. 711–720. ACM Press, New York (2008)
9. FaceTime Communications: The Collaborative Internet: Usage Trends, End User Attitudes and IT Impact (2008)
10. Muller, M.J., Raven, M.E., Kogan, S., Millen, D.R., Carey, K.: Introducing Chat into Business Organizations: Toward an Instant Messaging Maturity Model. In: Proc. of GROUP 2003, Sanibel Island, FL, USA, pp. 50–57. ACM Press, New York (2003)
11. IBM: IBM Reaps Business Benefits and Major Cost Savings From Unified Communications and Collaboration (2008)
12. Hudson, S.E., Smith, I.: Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems. In: Proc. of CSCW 1996, Boston, MA, USA, pp. 248–257. ACM Press, New York (1996)
13. Patil, S., Kobsa, A.: Privacy Considerations in Awareness Systems: Designing with Privacy in Mind. In: Markopoulos, P., de Ruyter, B., Mackay, W. (eds.) Awareness Systems: Advances in Theory, Methodology and Design. Springer, Heidelberg (2009)
14. Lee, A., Girgensohn, A., Schlueter, K.: NYNEX Portholes: Initial User Reactions and Redesign Implications. In: Proc. of GROUP 1997, Phoenix, AZ, pp. 385–394. ACM Press, New York (1997)
15. Patil, S., Kobsa, A.: Instant Messaging and Privacy. In: Proc. of HCI 2004, Leeds, UK, pp. 85–88 (2004), `http://www.ics.uci.edu/~kobsa/papers/2004-HCI-kobsa.pdf`
16. Calore, M.: Privacy Fears Shock Facebook. Wired News (September 2006), `http://www.wired.com/science/discoveries/news/2006/09/71739`
17. Adams, A.: Users' Perception of Privacy in Multimedia Communication. In: Proc. of CHI 1999: Extended Abstracts, Pittsburgh, PA, USA, pp. 53–54. ACM Press, New York (1999)
18. Adams, A., Sasse, M.A.: Privacy Issues in Ubiquitous Multimedia Environments: Wake Sleeping Dogs, or Let Them Lie? In: Proc. of Interact 1999, Edinburgh, UK, pp. 214–221 (1999)
19. Damian, D.E., Zowghi, D.: An Insight into the Interplay between Culture, Conflict and Distance in Globally Distributed Requirements Negotiations. In: Proc. of HICSS 2003, vol. 1, p. 19c (2003)
20. Battin, R., Crocker, R., Kreidler, J., Subramanian, K.: Leveraging Resources in Global Software Development. IEEE Software 18(2), 70–77 (2001)

21. Dourish, P.: Culture and Control in a Media Space. In: Proc. of ECSCW 1993, Norwell, MA, USA, pp. 125–137. Kluwer Academic Publishers, Dordrecht (1993)
22. Ackerman, M.S., Starr, B., Hindus, D., Mainwaring, S.D.: Hanging on the 'Wire: A Field Study of an Audio-only Media Space. ACM Trans. Information Systems 4(1), 39–66 (1997)
23. Lederer, S., Mankoff, J., Dey, A.K.: Who Wants to Know What When? Privacy Preference Determinants in Ubiquitous Computing. In: Proc. of CHI 2003, Fort Lauderdale, FL, USA, pp. 724–725. ACM Press, New York (2003)
24. Consolvo, S., Smith, I.E., Matthews, T., LaMarca, A., Tabert, J., Powledge, P.: Location Disclosure to Social Relations: Why, When, & What People Want To Share. In: Proc. of CHI 2005, Portland, OR, USA, pp. 81–90. ACM Press, New York (2005)
25. Olson, J.S., Grudin, J., Horvitz, E.: A Study of Preferences for Sharing and Privacy. In: Proc. of CHI 2005, Portland, OR, USA, pp. 1985–1988. ACM Press, New York (2005)
26. Palen, L., Dourish, P.: Unpacking "Privacy" for a Networked World. In: Proc. of CHI 2003, Fort Lauderdale, FL, USA, pp. 129–136. ACM Press, New York (2003)
27. Lederer, S., Mankoff, J., Dey, A.K.: Towards a Deconstruction of the Privacy Space. In: Ubicomp 2003 Workshop on Ubicomp Communities: Privacy as Boundary Negotiation (2003)
28. IBM: IBM Multi-National Consumer Privacy Survey (October 1999)
29. Westin, A.F.: Harris-Equifax Consumer Privacy Survey 1991 (1991)
30. Glaser, B.G., Strauss, A.L.: The Discovery of Grounded Theory: Strategies for Qualitative Research, 8th edn. Aldine Transaction, Chicago (1967)
31. Acquisti, A., Grossklags, J.: Privacy and Rationality. In: Strandburg, K., Raicu, D.S. (eds.) Privacy and Technologies of Identity: A Cross-Disciplinary Conversation, pp. 15–29. Sprinter, New York (2006)
32. Norman, D.A.: The Design of Everyday Things, 3rd edn. MIT Press, Cambridge (1998)
33. Patil, S., Kobsa, A., John, A., Brotman, L.S., Seligmann, D.: Comparing Privacy Attitudes of Knowledge Workers in India and the U.S (forthcoming, 2009)
34. Voida, A., Newstetter, W.C., Mynatt, E.D.: When Conventions Collide: The Tensions of Instant Messaging Attributed. In: Proc. of CHI 2002, Minneapolis, MN, USA, pp. 187–194. ACM Press, New York (2002)
35. Warren, S.D., Brandeis, L.D.: The Right to Privacy. Harvard Law Review 4(5), 193–220 (1890)
36. González, V.M., Mark, G.: Constant, Constant, Multi-tasking Craziness": Managing Multiple Working Spheres. In: Proc. of CHI 2004, Vienna, Austria, pp. 113–120. ACM Press, New York (2004)
37. Altman, I.: The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding. Brooks/Cole, Monterey (1975)
38. Hudson, J.M., Christensen, J., Kellogg, W.A., Erickson, T.: I'd be overwhelmed, but it's just one more thing to do: Availability and Interruption in Research Management. In: Proc. of CHI 2002, Minneapolis, MN, USA, pp. 97–104. ACM Press, New York (2002)
39. Kobsa, A., Patil, S., Meyer, B.: Privacy in Instant Messaging; An Impression Management Model (forthcoming, 2009)
40. Goffman, E.: The Presentation of Self in Everyday Life. Doubleday, Garden City (1959)

Assessing the “Quality of Collaboration” in Technology-Mediated Design Situations with Several Dimensions

Jean-Marie Burkhardt[1], Françoise Détienne[2,3], Anne-Marie Hébert[2], and Laurence Perron[4]

[1] Paris Descartes University, Ergonomics–Behavior-Interaction Lab.
45 rue des Saints-Pères 75270 Paris Cedex 06, France
[2] LTCI- UMR 5141 - CNRS - Telecom Paris Tech
46 rue Barrault 75634 Paris Cedex 13, France
[3] INRIA-Paris, France
[4] France Telecom R&D / Orange Labs, Lannion, France
jean-marie.burkhardt@univ-paris5.fr,
{francoise.detienne,hebert}@telecom-paristech.fr,
laurence.perron@orange-ftgroup.com

Abstract. Our objective is to measure and compare the quality of collaboration in technology-mediated design activities. Our position is to consider collaboration as multidimensional. We present a method to assess quality of collaboration composed of seven dimensions concerning communication processes such as grounding, coordination processes, task-related processes, symmetry of individual contributions as well as motivational processes.

Keywords: multi-user interaction /cooperation, collaboration, design, methodology, cognitive ergonomics, CSCL.

1 Introduction

Methods for assessing the quality of mediated-collaboration are to be as central in user-centred design as the methods used for assessing the ergonomic quality of interface, i.e. the usability of UI [1]. In spite of a growing number of methods to evaluate groupware technologies and group work [e.g. 2], no measurement method of this facet of activities in collaborative design situations has been proposed as far as we know. In the close field of Computer-Supported Collaborative Learning (CSCL), the analysis of the process of collaboration is a central topic of research. The Spada rating scheme [3,4] is certainly the most representative of recent effort made in this field to assess nine dimensions of collaboration and its quality.

In this paper, we propose an evaluation method to cover these several dimensions of the quality of collaboration in technology-mediated design situations together with explicit qualitative criterion whenever possible. The term “quality” has two distinct meanings. Firstly, it refers to the “essence” of an object, i.e., a specific feature of the object. Quality in this sense does not suppose any positive or negative value. Secondly, the word “quality” refers to the “good” versus “bad” values of an object’s properties, based on a set of (more or less explicit) criteria. Quality also implies that norms and references are provided to support the comparison of objects according to

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 157–160, 2009.

measured values. We use the notion of "quality" to underline that (a) our method focuses on those dimensions of activity related to collaboration quality and (b) that we aim to elicit references (often implicit in the literature) and relevant standards regarding collaborative activities. In the next parts, we briefly report on the rationale of this multidimensional approach on the basis of theoretical arguments and of results from empirical studies (see [5] for an extended version). Then, we present our method to assess the quality of collaboration, and report on our current work and perspectives.

2 Assessing Quality of Collaboration: Multidimensionality and Explicit References

Previous work in cognitive ergonomics of design and in CSCL has shown collaboration as multidimensional. For example, empirical studies on collaboration in design teams (for a state of the art, see [6]), in various application domains (e.g., software design, architectural design), have highlighted the following distinctive collaborative processes most important for successful design: communication processes such as grounding [7], task-related processes (e.g. exchanges of knowledge relevant for the task at hand; argumentation processes), coordination processes, and motivational processes. The latter encompass interpersonal relationship and motivation, which are less covered in studies on design, although this more subjective dimension might affect strongly the actual way of collaboration. Finally, how symmetric are individual contributions through all these dimensions provides a complementary aspect of quality of collaboration. These processes can be taken as a referential of good collaboration with respect to design. Evidence from studies can be used to set potential referential values regarding collaboration and efficacy. As an illustration, empirical studies of collaborative design (e.g. [8, 9, 10] found that grounding, although time-consuming, was most important to ensure good design: for instance, in [10] when teams bypassed grounding (referred to as "analysis"), this led them to premature evaluation of design ideas. Other studies show that at distance, characteristics of communication media, such as no visibility, or no simultaneity [7], may affect grounding and awareness. As another illustration, recent research on collaboration processes in design [11] considers the balance between the roles of participants according to communication, group management and task management as a good indicator of collaboration. This aspect is similar to the notion of reciprocal interaction [3] and symmetry in the interaction [12] in CSCL.

3 Assessing the Quality of Collaboration in Technology-Mediated Design: Our Method

Our method is initially (and thus partly) based on Spada's method [3]. This method shows some limits (e.g. indicators to assess collaboration are underspecified, subjective rating), however its multidimensional characteristic is a good basis to further develop a method to assess the quality of collaboration in technology-mediated design. The method has been modified so as to take into account empirical results on collaborative design. We have also modified the scoring method to improve the

Table 1. Dimensions and indicators of our method

Dimensions	Definition	Indicators
1. Fluidity of collaboration	It assesses the management of verbal communication (verbal turns), of actions (tool use) and of attention orientation	– Fluidity of verbal turns – Fluidity of tools use (stylet, menu) – Coherency of attention orientation
2. Sustaining mutual understanding	It assesses the grounding processes concerning the design artifact (problem, solutions), the designers' actions and the state of the Augmented Reality desktop (e.g. activated functions).	– Mutual understanding of the state of design problem/solutions – Mutual understanding of the actions in progress and next actions – Mutual understanding of the state of the system (active functions, open documents)
3. Information exchanges for problem solving	It assesses design ideas pooling, refinement of design ideas and coherency of ideas.	– Generation of design ideas (problem, solutions, past cases, constraints) – Refinement of design ideas – Coherency and follow up of ideas
4. Argumentation and reaching consensus	It assesses whether or not there is argumentation and decision taken on common consensus.	– Criticisms and argumentation – Checking solutions adequacy with design constraints – Common decision taking
5. Task and time management	It assesses the planning (e.g. task allocation) and time management.	– Work planning – Task division – Distribution and management of tasks interdependencies – Time management
6. Cooperative orientation	It assesses the balance of contribution of the actors in design, planning, and in verbal and graphical actions.	– Symmetry of verbal contributions – Symmetry of use of graphical tools – Symmetry in task management – Symmetry in design choices
7. Individual task orientation	It assesses, for each contributor, its motivation (marks of interest in the collaboration), implication (actions) and involvement (attention orientation).	– Showing up motivation and encouraging others motivation – Constancy of effort put in the task – Attention orientation in relation with the design task

assessment procedure[13]. The judges have to give additional explicit answers (yes, no, yes/no) to paired questions with positive or negative valence targeting specific indicators of each dimension (Table 1). We added an explicit scoring algorithm based on the number of positive and negative answers to questions for each dimension.

4 Conclusion and Perspectives

The test of this new version showed a strong reliability based on inter-raters correlations (see [5]). We plan now to compare in a systematic way between the results obtained with this rating method and those obtained with more time-consuming coding methods. We will also explore to which extent this method can be used by judges from the design domain, e.g., architectural design.

Acknowledgments. This research has been partly funded by OrangeLabs. We thank Michael Baker for his input and feedbacks on our method.

References

1. Scapin, D.L., Bastien, J.M.C.: Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behavior and Information Technology 17(4/5), 220–231 (1997)
2. Pinelle, D., Gutwin, C., Greenberg, S.: Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the nechanics of collaboration. ACM Transactions on Computer-Human Interaction 10(4), 281–311 (2003)
3. Spada, H., Meier, A., Rummel, N., Hauser, S.: A new method to assess the quality of collaborative process in CSCL. In: Koschmann, T., Suthers, D., Chan, T.W. (eds.) Computer Supported Collaborative Learning 2005: The Next 10 Years!, pp. 622–631. Lawrence Erlbaum Associates, Mahwah (2005)
4. Meier, A., Spada, H., Rummel, N.: A rating scheme for assessing the quality of computer-supported collaboration processes. Computer-supported Collaborative Learning 2(1), 63–86 (2007)
5. Burkhardt, J.-M., Détienne, F., Hebert, A.-M., Perron, L., Safin, S., Leclercq, P.: An approach to assess the quality of collaboration in technology-mediated design situations. In: European Conference on Cognitive Ergonomics 2009, Otaniemi, Finland, September 30-October 2 (2009)
6. Détienne, F.: Collaborative design: managing task interdependencies and multiple perspectives. Interacting With Computers 18(1), 1–20 (2006)
7. Clark, H.H., Brennan, S.E.: Grounding in communication. In: Resnick, L., Levine, J.-M., Teasley, S.D. (eds.) Perspectives on socially shared cognition, pp. 127–149. American Psychological Association Pres, Washington (1991)
8. D'Astous, P., Détienne, F., Visser, W., Robillard, P.N.: Changing our view on design evaluation meetings methodology: a study of software technical review meetings. Design Studies 25, 625–655 (2004)
9. Olson, G.M., Olson, J.S., Carter, M.R., Storrosten, M.: Small Group Design Meetings: An Analysis of Collaboration. Human-Computer Interaction 7, 347–374 (1992)
10. Stempfle, J., Badke-Schaub, P.: Thinking in design teams - an analysis of team communication. Design Studies 23, 473–496 (2002)
11. Barcellini, F., Détienne, F., Burkhardt, J.M., Sack, W.: A socio-cognitive analysis of online design discussions in an Open Source Software community. Interacting With Computers 20(1), 141–165 (2008)
12. Baker, M.: Forms of cooperation in dyadic problem-solving. Revue d'Intelligence Artificielle 16(4-5), 587–620 (2002)
13. Détienne, F., Burkhardt, J.-M., Hébert, A.-M., Perron, L.: Assessing the quality of collaboration in design: bridging cognitive ergonomics and CSCL approaches. In: Workshop CSCW and Human Factors, CSCW 2008, San Diego, USA, November 9 (2008)

A Multi-touch Tool for Co-creation

Geke D.S. Ludden and Tom Broens

Novay, Brouwerijstraat 1, P.O. Box 589, 7500 AN, Enschede, The Netherlands
{Geke.Ludden,Tom.Broens}@telin.nl

Abstract. Multi-touch technology provides an attractive way for knowledge workers to collaborate. Co-creation is an important collaboration process in which collecting resources, creating results and distributing these results is essential. We propose a wall-based multi-touch system (called CoCreate) in which these steps are made easy due to the notion of connected private spaces and a shared co-create space. We present our ongoing work, expert evaluation of interaction scenarios and future plans.

Keywords: multi-touch, wall, interaction scenario, CoCreate.

1 Introduction

Multi-touch technology gains a lot of attention in research and industry. Bill Buxton gives a nice summary of work in this area [1]. Multi-touch technology is used for phones (e.g. Iphone), computer input interfaces (e.g. track pads) and surfaces (e.g. tables and walls). Multi-touch technology allows multiple users to interact with a surface simultaneously and therefore seems suitable to support the co-creation processes of knowledge workers. Despite this potential, the number of tools using multi-touch technology for co-creation is still limited. In this paper we investigate the feasibility of a wall-based multi-touch tool for co-creation.

Co-creation is all about creating and sharing. In general we distinguish three phases in a co-creation process: (i) collecting public and private resources of participants (ii) create a shared result and (iii) distribute this result to the participants. For an optimal process all phases should be supported.

Much of the existing co-creation tools have a table form factor [e.g., 2, 3, 4]. Rogers and Lindley [5] investigated differences in collaborating using vertical and horizontal displays. They concluded that a table form factor facilitated collaboration better. However, the setup they used may have negatively influenced their findings for collaborating using a vertical display. We doubt whether a horizontal display is the most suitable for co-creation. Analogous to stand-up (or 'scrum') meetings that are claimed to be more efficient [6], a wall-based system stimulating standing sessions might result in a quicker and more creative process compared to sitting sessions at a table. Furthermore, intrinsically a wall has the same orientation for all participants, which eases the creation process. Additionally, in a remote table-based setting, such as Digitable [2] creating a feeling of connectedness by showing a camera feed of the remote participants, which is the de-facto standard, is more awkward due to the face-camera orientation.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 161–164, 2009.

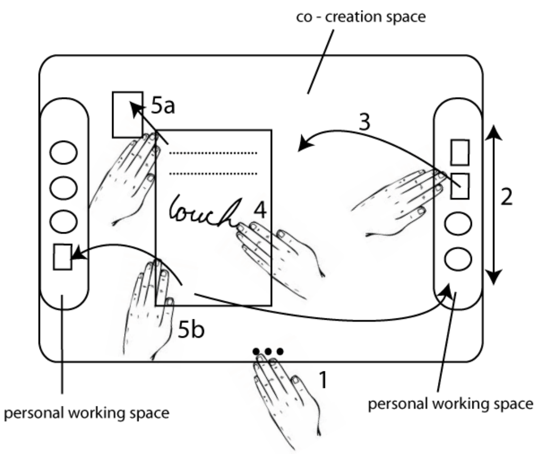
Hence, we propose a wall-based CoCreate system in which collecting resources and sharing the result is made easy due to the notion of connected private spaces and a shared co-create space. In the remainder of this paper we discuss the design of the CoCreate tool, and an initial feasibility evaluation, conclusions and our future plans.

2 Design

We created interaction scenarios for co-creation using a multi-touch wall at one location ('co-located') and at a distance ('at a distance'). For these scenarios, we envision a multi-touch wall with a display area of 185 x 104 cm.

2.1 'Co-located' Co-creation Using Multi-touch

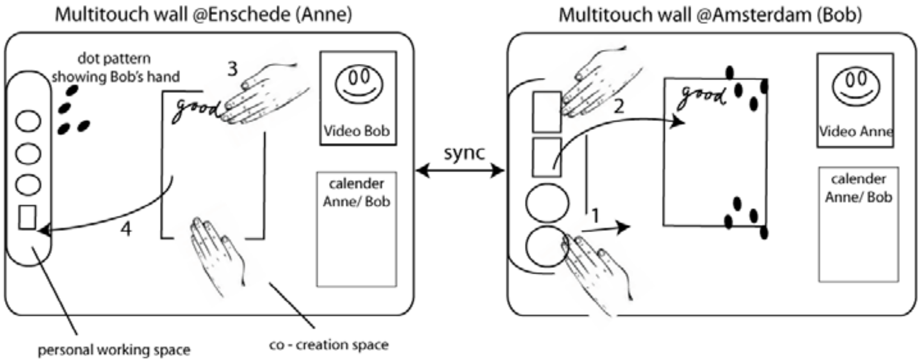
Generally, the 'co-located' CoCreate tool comprises both a co-creation space and personal working spaces for current users. The personal working space contains links to, for example, Flickr or Delicious accounts, the user's desktop, and his or her frequently used folders. The following figure illustrates how this tool will typically be used.



(1) Open CoCreate from the application dock. Personal working spaces are opened for all users. (2) Browse through personal working space. (3) Drag article into the co-creation space. The article enlarges. (4) Write a comment on the article. (5a) Put aside the article, article minimizes. (5b) Drag the article into personal spaces.

2.2 Co-creation Using Multi-touch 'at a distance'

The CoCreate tool 'at a distance' is aimed at supporting remote knowledge workers to create together. Please note that multi-touch technology is not essential in this scenario. However, this scenario provides an additional way to use a multi-touch wall. The tool typically shows a video of a remote person, calendars of all current users, a personal working space and a co-creation space. The following figure provides an example of how such a tool could be used.



(1) Enlarging personal working space. (2) Drag document to co-creation space, an enlarged document appears on both walls. While Bob is manipulating items on his wall, Anne sees a dot pattern representing the location of Bob's hand on her wall. The two can talk about the document using the video connection. (3) Anne makes a note on the document, now Bob sees a dot pattern representing Anne's hand. (4) Anne drags the document into her personal space.

3 Evaluation

The feasibility of our CoCreate tools was evaluated by a group of experts. To do so, the scenarios as described above were posted on a website together with two comics illustrating situations in which multi-touch displays could help people to work together more efficiently. Visitors of our website were asked to indicate to what end they agreed to three statements about the specified interaction choosing from five options: 'strongly agree', 'agree', 'neither agree nor disagree', 'disagree' and 'strongly disagree'. The statements were: (Q1)'I like the features of this concept'; (Q2)'I think this concept can help people to work together', (Q3)'I think the interaction expressed in this scenario will work well'. Furthermore, we asked for additional remarks about the scenarios.

Seven experts that were selected based on their expertise on designing for interaction. (6 male, 1 female, aged 22-31) evaluated the interaction scenarios. Generally, most respondents evaluated both scenarios positively. For both scenarios, as a response to Q1 and Q2, all respondents selected either 'agree' or 'agree completely'. Responses on Q3 were somewhat less positive, 43% of the responses for the 'co-located' scenario and 29% of the responses for the 'at a distance' scenario were 'neither agree nor disagree'. Other responses were either 'agree' or 'completely agree'. Although the responses on scales generally were positive, the remarks that respondents made in addition to their responses on scales indicate that some of them had doubts about the added value of using a multi-touch wall over traditional working methods. Also, it was mentioned that using CoCreate would require new ways of organizing meetings. One respondent mentioned privacy issues coming from the fact that all users can see each others' personal spaces. (This of course only applies to the 'co-located' scenario.) Some respondents mentioned that they would see more use to simpler applications on a multi-touch wall. For the 'at a distance' scenario, respondents did not agree about the value of seeing the other persons' actions as dots representing a hand. One respondent suggested using the image of a shadow of a hand instead of dots. These are all typically issues we would like to test in a pilot set-up. At the same time, we would like to understand the impact of this concept on day-to-day work. Therefore, a working prototype of the proposed concept will be implemented in a real work environment in industry.

4 Conclusions and Future Work

It must be noted that for our ‘at a distance’ scenario, a multi-touch wall is not the only way to solve the issue of sharing interfaces. Other projection tools such as Lucid Lab’s Virtual Desktop [7] also allow co-creation. However, for the ‘co-located’ scenario, multi-touch is essential to allow multiple persons to work at the same interface. To implement and test the ideas laid out above, we are currently developing both hardware and software. A multi-touch wall implementing CoCreate functionality will be created in cooperation with Soco Amsterdam [8]. The wall will rely on Frustrated Total Internal Reflection (FTIR) technology and will feature a display area of 185 x 104 cm. At the same time, we are working on improvements of the scenarios, more specifically we are working out the interaction with personal spaces especially in situation where more than two people use the screen.

References

1. Bill Buxton, `http://www.billbuxton.com/multitouchOverview.html`
2. Pauchet, F., Coldefy, L., Lefebvre, S., Louis Dit Picard, A., Bouguet, L., et al.: Mutual awareness in collocated and distant collaborative tasks using shared interfaces. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. Part. I. LNCS, vol. 4662, pp. 59–73. Springer, Heidelberg (2007)
3. Shen, C., Vernier, F.D., Forlines, C., Ringel, M.: DiamondSpin: an extensible toolkit for around-the-table interaction. In: ACM Conference on Human Factors in Computing Systems (CHI), pp. 167–174. ACM Press, New York (2004)
4. Hunter, S., Maes, P.: Wordplay: a table-top interface for collaborative brainstorming and decision making. Submitted to IEEE TableTop Collaborative Surfaces (2008)
5. Rogers, Y., Lindley, S.: Collaborating around vertical and horizontal large interactive displays. Interacting with Computers 16, 1133–1152 (2004)
6. Vizdos, M.: Opening communication within a scrum team. Methods & Tools 16, 31–35 (2008), `http://www.methodsandtools.com/PDF/mt200803.pdf`
7. Elsen, C., Leclercq, P.: “SketSha” – The sketch power to support collaborative design. In: Luo, Y. (ed.) CDVE 2008. LNCS, vol. 5220, pp. 20–27. Springer, Heidelberg (2008)
8. SocoAmsterdam, `http://www.socoamsterdam.nl`

GColl: A Flexible Videoconferencing Environment for Group-to-Group Interaction

Petr Slovák[1,2], Pavel Troubil[1,2], and Petr Holub[2]

[1] Faculty of Informatics, Masaryk University, Botanicka 68a, 60200 Brno, Czech Republic
[2] CESNET, z.s.p.o, Zikova 4, 16000 Prague, Czech Republic
{slovak,pavel,hopet}@ics.muni.cz

Abstract. In this paper, we present GColl, a group-to-group videoconferencing environment concept, which aims to provide a natural communication channel even for ad-hoc groups or other teams that require frequent changes in the number of participants or videoconferencing locations. GColl supports mutual gaze as well as partial gaze awareness for all participants while still retaining very modest technical requirements: a camera and an echo-canceling microphone at each site; and a notebook with two USB cameras for each user. A working prototype is available for download.

Keywords: Videoconferencing, CMC, Gaze Awareness, Mutual Gaze, Partial Gaze Awareness, Mixed Presence, Presence Disparity.

1 Introduction

In the last few years, videoconferencing environments and tools came once again into the focus of human-computer interaction researchers. While videoconferencing between individuals has become quite common (e.g., Skype), attention is now turned to videoconferencing environments that would support collaboration among remote *groups* of people. In this setting, problems known from previous research (e.g., preservation of gaze awareness and other non-verbal cues) become even more salient and creation of specialized systems is therefore necessary.

Such systems might be used by distributed teams both in academic and commercial sphere. Although some of these teams might be stable enough to settle for a fixed system in a dedicated room, others require the possibility of frequent changes in the number of attendants and videoconferencing locations. Moreover, some might even be created only for very short-term tasks, thus meeting just a few times altogether. Throughout the rest of this paper, we will denote groups with these needs as *ad-hoc groups*.

To support ad-hoc groups, our goal was to create a videoconferencing environment, which would be a compromise between the need for preserving the non-verbal cues and the requirement of low-cost and flexibility. Our resulting system, GColl, supports *mutual gaze* (knowledge, that someone else is looking at you) as well as *partial gaze awareness* (knowing general direction someone else is looking) for all participants. At the same time, it requires only a camera and an echo-canceling microphone at each remote site, and a common laptop with two USB cameras for each user. Users can join or leave GColl without any changes to the physical layout at any site; this can be done even during a videoconference meeting.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 165–168, 2009.

Related Work. Concept of gaze awareness has been studied in great detail in the literature due to its importance for effective communication as well as other group activities (e.g., in [7, 3]). In most videoconferencing environments, gaze information is not conveyed easily due to the usual discrepancy between the camera position and the visualization of the other person's eyes. Several multi-person videoconferencing systems were invented to mediate some or all forms of gaze awareness (e.g., GAZE2 [8], MAJIC [5], and Hydra [6]). The most recent one, Multiview [4], even achieved a full spatial faithfulness and reported very positive user evaluations. None of them are, however, directly suitable for ad-hoc groups due to either the lack of support for group-to-group interaction ([8, 5, 6]), or problems with mobility and flexibility ([4]).

In all of these systems, a change of one site (e.g., an increase in the number of participants) calls for non-trivial changes of the physical layout at remote sites. Also, a change of site location is not easy due to the need of transferring a complex structure of cameras and viewing screens. Once set up, these designs provide the users with excellent communication environment; however, ad-hoc groups might be forced to spend a lot of energy on re-installment of the system.

GColl mediates a slightly restricted variant of mutual gaze and partial gaze awareness than, e.g., Multiview. On the other hand, we believe the restriction to be very well counter-weighted by the highest mobility and flexibility of GColl among all similar systems known to the authors.

2 Environment Description

2.1 Basic Structure

Support of mutual gaze was one of our main priorities during the design phase of GColl. To fulfill this requirement, the videoconferencing system has to supply each user with her own visualization device as otherwise those users, who are sharing a common device, are in a sense sharing the "same eyes"—either all of them experience a feeling of eye-contact with a remote participant, or none does.

We could not use designs based on several projectors or monitors as this would severely limit flexibility and mobility of GColl. Instead, we opted for assumption that each user has a personal notebook or other computer at her disposal, which, as we believe, is not too restrictive for the anticipated class of possible users.

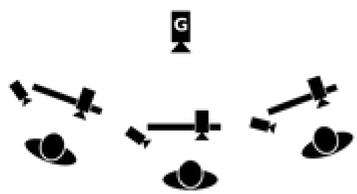


Fig. 1. A scheme of site installation

A basic structure diagram of a GColl installment at a site with three users is depicted in Figure 1. An image of the whole group is recorded by a single camera *G*. An echo-canceling microphone is used to capture audio at this site. Additionally, each user uses her notebook, two common web cameras and a headset with a microphone: the *focus camera* is attached to the top of the notebook screen itself while the *side camera* is placed on the desk next to the notebook. If available, a shared screen can be used for slides or any other material the group needs.

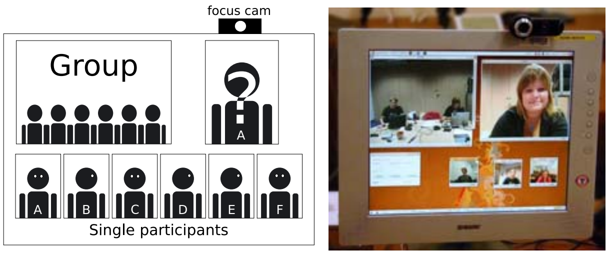


Fig. 2. A scheme of the visualization layout and a photo of its prototype implementation (with two remote sites)

A scheme of GColl visualization screen and a photo of our prototype implementation is shown in Figure 2. The screen is divided into three parts: the top left image provides a “whole room” view of the remote group (or splits into subframes if more than two groups are communicating); images of other participants (remote as well as co-located) are shown on the bottom of the screen and after clicking on any of these small images, the chosen stream is then shown in higher quality in the top right (*focus*) frame. A click on the *whole group* frame itself unselects the chosen stream (and restores it to the original size) once it is not needed.

2.2 Functionality

The *focus* frame allows each participant to easily select some other user, on whom she wants to currently focus, and see him in more detail. The system informs other users whether or not they are “in focus” of each of the other participants by selecting the appropriate videostreams to be shown in the bottom (or top right) frames of their visualization screen: those users, who are not selected, receive the *side camera* image; the selected user receives the stream from the *focus camera*. For example, the user whose screen is depicted on Figure 2 knows that she has been selected by the participants A, C and F, and that she is herself focusing on the participant A.

This setup allows us to mediate a sensation of mutual gaze. It is based on findings by Chen [2] that show asymmetry in human sensitivity to eye-contact—that is, people would still perceive eye-contact if the other person’s gaze is directed less than 5° below the camera. By having a fixed place for the *focus* window in GColl, it is easy to attach the *focus camera* appropriately close to the visualized eyes. Thus, if you are selected by another participant and that participant looks at you in his *focus* frame, you will perceive eye-contact as can be seen in Figure 2.

3 Conclusions and Future Work

In this paper, we have presented the GColl system for group-to-group collaboration with mutual gaze support. GColl is designed to be usable even for ad-hoc communicating groups due to its flexibility and modest technical requirements. Fully functional (yet still prototype) implementation is available for download at [1].

Currently, we are analyzing data gathered in a usability study comprising more than 90 people and the preliminary results seem to be quite promising. A real world case study is being run simultaneously, involving working teams from several universities and a national research organization.

To bring GColl closer to face-to-face communication, we are now exploring the possibilities of using USB cameras streams for eye-tracking. If we can overcome an inaccuracy of this gaze-estimation method, it might be used for user selection instead of the current mouse/keyboard interface. We believe that this method could be more natural and therefore ease users' interaction with GColl. Once we have this modification ready, we plan to conduct another user study to verify our hypothesis.

GColl currently supports group-to-group communication, but has no inherent shared workplace or any other means of task support. We presume that such extensions of the design are possible (e.g., a shared tabletop at each site) and would enhance the collaboration experience greatly.

Acknowledgments. We would like to thank Milan Kabát for his help with the implementation of the GColl prototype. This project has been supported by a research project CESNET FR 254/2007.

References

1. SourceForge.net: GColl, `http://www.sourceforge.net/projects/gcoll`
2. Chen, M.: Leveraging the asymmetric sensitivity of eye contact for videoconference. In: CHI 2002, pp. 49–56. ACM, New York (2002)
3. Monk, A., Gale, C.: A Look Is Worth a Thousand Words: Full Gaze Awareness in Video-Mediated Conversation. Discourse Processes 33, 257–278 (2002)
4. Nguyen, D.T., Canny, J.: Multiview: improving trust in group video conferencing through spatial faithfulness. In: CHI 2007, pp. 1465–1474. ACM, New York (2007)
5. Okada, K.-I., Maeda, F., Ichikawaa, Y., Matsushita, Y.: Multiparty videoconferencing at virtual social distance: MAJIC design. In: Furuta, R., Neuwirth, C. (eds.) CSCW 1994: Proceedings of the 1994 ACM conference on Computer supported cooperative work, pp. 385–393. ACM, New York (1994)
6. Sellen, A., Buxton, B., Arnott, J.: Using spatial cues to improve videoconferencing. In: CHI 1992, pp. 651–652. ACM, New York (1992)
7. Vertegaal, R., van der Veer, G., Vons, H.: Effects of Gaze on Multiparty Mediated Communication. In: Moller, T., Ware, C. (eds.) Graphics Interface, pp. 95–102. Morgan Kaufmann, San Francisco (2000)
8. Vertegaal, R., Weevers, I., Sohn, C., Cheung, C.: Gaze-2: conveying eye contact in group video conferencing using eye-controlled camera direction. In: Cockton, G., Korhonen, P. (eds.) CHI 2003, pp. 521–528. ACM, New York (2003)

Space as a Resource in Creative Design Practices

Dhaval Vyas[1], Gerrit van der Veer[2], Dirk Heylen[1], and Anton Nijholt[1]

[1] Human Media Interaction, University of Twente, The Netherlands
[2] School of Computer Science, Open University, The Netherlands
d.m.vyas@ewi.utwente.nl

Abstract. Based on longitudinal ethnographic fieldwork in two industrial design departments and two design companies, we explore the role of spatial arrangements for supporting creative design practices within different design studios. From our results, we show that designers explicitly make use of the physical space for 1) communicating and inspiring design ideas, 2) exploring design solutions and 3) managing design projects. We believe that these design practices could bring insightful implications for developing ubiquitous technologies to support the design profession.

Keywords: Workspace, design practice, ethnography, creativity.

1 Introduction

The spatial aspects of design studios play an important role in supporting different design activities, techniques and use of material and digital tools. Within a typical design studio, desks, office walls, clipboards, drawing boards and so on are full of informative, inspirational and creative design-related artefacts such as handmade sketches, drawings, posters, post-it notes, physical models, prototypes and the like. This kind of information can be seen as external representations of the ongoing design activities. The physical space within design studios is not just a carrier for design-related information but importantly it is a site of methodic design practices. The space offers resources for social, organizational and individual activity [3] that designers routinely encounter and use on an everyday basis.

Conceptual frameworks such as Distributed Cognition [2] and Organizational Memory [1] suggest that information about a particular work setting or a work culture is distributed between different external representations and people situated there. In a domain such as design, where the whole design process progresses through different multimodal information (from handmade sketches to final prototypes), the study of spatial arrangement becomes an important issue.

We studied two industrial design departments and two design companies over a period of 8 months. We used an ethnomethodologically-informed ethnographic approach [4] to understand designers' practices in their natural settings and to get an insight into the approaches and methods they use in going about their everyday work. We used three methods: 1) contextual interviews, 2) naturalistic observations, and 3) recorded design sessions of design students and professionals. In this paper, we organize our results focusing on the spatial aspects and show that designers explicitly make use of the physical space for the purpose of communicating and inspiring design

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 169–172, 2009.

ideas, exploring design solutions and managing design projects. We believe that in order to better support designers' work and to develop ubiquitous collaborative technologies, we need to understand how spatial aspects support cooperative practices of designers' and what role they play in supporting creativity in the everyday work.

2 Use of Space in Cooperative Design

In our study we observed that physical space within design studios supports 'explorative' and 'coordinative' activities amongst co-designers. Focusing on the creative use of physical space, we provide our results in three themes: communicating and inspiring design ideas, exploring design solutions and managing design projects.

Communicating and inspiring design ideas. Designers keep sketches, drawings, pictures, design models and other relevant things on their working space and surfaces such as clipboards, drawing boards and office walls in a way that constantly informs and inspires their design work (figure 1). The multimodality and ability to support and convey information through multiple senses facilitate rich communications between co-workers. Additionally, the design artefacts that are attached to particular spatial objects are indicative of different phases of the design process, the current state, future planning, and so on. One designer commented, "*depending on the phase of the project, I arrange my surroundings. It's important for me to have these artefacts around so that I can register where I am at in the project*". So, these design artefacts are also reminders.

The construction of these evocative workplaces is also intended to develop new ideas, inspiration and creative thinking. One designer commented, "*Normally within a project I need a strong foundation to start with. So, when I am communicating my ideas I need to have several different aspects about my design. Because when the foundation is strong it helps in convincing people. These visual objects around me show my foundational work and work as strong building blocks.*"

Fig. 1. A section of a design studio where design students have utilized different surfaces to support design work

Fig. 2. An example where a group of design students are brainstorming on an ephemeral workspace

Another important aspect of utilizing space to support design activities is to experiment with the physical space in order to develop a creative ecology within design studios. As can be seen in figure 2, designers have created an ephemeral workspace where they can brainstorm and collect or develop new design ideas for their ongoing projects. In some cases the students would intentionally move around, change location of their working, create new collaborative spaces, play with different things in the studio, and so on.

Exploring design solutions. Another reason to utilize space is to explore design solutions by elaborating and/or dividing design challenges in order to generate detailed descriptions and related aspects. This in turn would help in resolving a particular design situation. The way physical space allows the representation of design tasks can affect designers' reasoning abilities and performance. As one designer suggested, "*I normally try to visualize all the material and data that I collected from my user studies and try to find out patterns and explore design opportunities from this data. I then make my own sketches and models and keep all these in a way that can help me find out new ideas*". Several examples of this were seen in both of the design studios. Designers keep, for example, pictures from ethnographic or other field studies on their office walls and on their desks (figure 3). The aim here is not just to solve a design problem but to collect greater and useful insights of a given situation so that a solution can be envisioned.

Fig. 3. An example where an office wall is used to elaborate a design problem, by arranging and connecting images from field studies and design sketches

Fig. 4. An example where a group of designers have devised physical space to keep project-specific information

Managing design projects. Design being a collaborative process requires organization and management of the work of co-designers. The spatial aspects within design studios also play a role in supporting organization and management of design projects. Figure 4 shows an example where a design team has used wooden panels, and movable drawing boards to store information related to detailed project descriptions, project plan, data generated from brainstorming sessions, design concepts, work division within a team and to-do lists. Each individual piece of design-related information has a strong even explicit link to some aspect of the project at hand. The ecology of these pieces of information creates an information rich environment needed to stimulate creativity and to develop novel ideas.

Within an ongoing design project, designers deal with a plethora of design materials, and being aware of different ‘happenings’ is an important issue. We observed that the way designers keep project-related design materials on different spatial objects within their studios improves the visibility and provides an overview of the work being carried out. Understanding how design artefacts within a work environment are organized, configured, manipulated and handled supports the awareness of co-workers’ activities and, hence, contributes to the coordination of work. Design iterations, methods, and conventions can be easily extracted when design artefacts and related materials are kept in public visibility using physical space. The visibility of design activities is also manifested in and through the use of these artefacts. At the same time such a creative space could provide opportunities to reflect on the ongoing project and to allow designers to change, combine or divert aspects of their design process.

3 Conclusions and Future Work

What has been presented here is an account of how space is used to support creative design practices by designers in the academic as well as the professional domains. As we observed, the physical space within a design studio offer resources that range from clearly defining design problems, exploring new possibilities and easing communicative difficulties to developing a communication language with co-workers.

Creativity is a critical aspect of design and needs to be supported though technological means. We believe that an account of real-world design practices such as this could be very fruitful if we are to design collaborative technologies. Within an ongoing research project, we are investigating other aspects related to collaborative design practices (e.g. [5]). We plan to combine the results of these ethnographic studies and develop implications for designing collaborative technologies. We also plan to develop a system to be tested in real-world design practices.

Acknowledgements. This work is supported by the EU Project AMIDA (FP6-0033812).

References

1. Ackerman, M.S., Halverson, C.: Organizational Memory: Processes, Boundary Objects, and Trajectories. In: Proc. of the 32nd Annual Hawaii international Conference on System Sciences. HICSS, vol. 1. IEEE Computer Society, Washington (1999)
2. Hutchins, E.: Cognition in the wild. MIT Press, Cambridge (1995)
3. Perry, M., O’Hara, K.: Display-Based Activity in the Workplace. In: Rauterberg, M., et al. (eds.) Proc. of INTERACT 2003. IFIP, pp. 591–598. IOS Press, Amsterdam (2003)
4. Randall, D., Harper, H., Rouncefield, M.: Fieldwork for Design – Theory and Practice. CSCW series. Springer, London (2007)
5. Vyas, D., Heylen, D., Nijholt, A.: Physicality and Cooperative Design. In: Popescu-Belis, A., Stiefelhagen, R. (eds.) MLMI 2008. LNCS, vol. 5237, pp. 325–337. Springer, Heidelberg (2008)

five: Enhancing 3D Wall Displays with a 2D High-Resolution Overlay

Daniel Steffen[1], Achim Ebert[1], Matthias Deller[1], and Peter Dannenmann[2]

[1] Deutsches Forschungszentrum für Künstliche Intelligenz GmbH,
Trippstadter Str. 122, 67663 Kaiserlautern, Germany
{Daniel.Steffen,Achim.Ebert,Matthias.Deller}@dfki.de
[2] Fachhochschule Wiesbaden, Fachbereich Ingenieurwissenschaften,
Am Brückweg 26, 65428 Rüsselsheim, Germany
dannenmann@mndu.fh-wiesbaden.de

Abstract. Projection-based stereoscopic wall displays allow users to immerse themselves into virtual scenes, such as architecture simulations or games. However, the usually low resolution (dpi) of such displays and slight alignment offsets between the two projectors result in a loss of detail and bad readability of textual information.

We propose addressing the problem by overlaying a third projector, so that its image is visible to both eyes. We eliminate offset artifacts by extracting 2D contents from the scene and rendering it using this dedicated "2D projector". In addition, we locally increase resolution by focusing the 2D projector onto a smaller region. This allows us to reduce the size of overlaid 2D annotations, thereby reducing interference with the 3D scene. In our paper, we describe the design of our display system called *five*. Following, we present two detailed user studies that compare *five* with an overview-and-detail and a pan-and-zoom interface.

Keywords: Focus-plus-context screens, stereoscopy, peripheral vision, experimental evaluation.

1 Introduction

Projection-based stereoscopic wall displays allow users to immerse themselves into virtual scenes. Application scenarios include construction in mechanical or civil engineering, process control, and interactive games. In remote surgery, stereoscopic presentation allows exact hand-eye coordination. In architecture, 3D visualization helps convey complex spatial relationships without requiring experience with the interpretation of blueprints. Displaying contents stereoscopically can also improve users' perception.

Many 3D scenarios involve context information. Providing this information to the users in the form of annotations, such as text overlays, allows users to perform visual sensemaking tasks [18]. Unfortunately, stereo projection systems can only display detail and especially text in a limited way, because the wide viewing angle required for immersion comes at the expense of detail. In addition, subtle alignment offsets between the two projectors result in double images, further impacting on the

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 173–186, 2009.

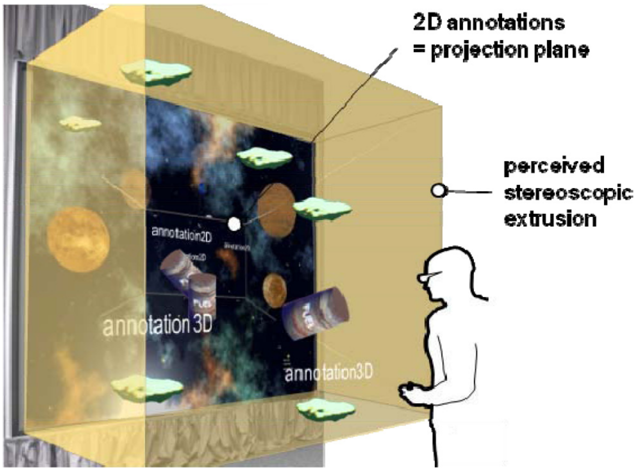


Fig. 1. The five display system complements wide-angle stereoscopic projection with a more focused, 2D hi-res projector. This allows overlaying readable annotations, such as text.

readability of 2D contents, such as text (Fig. 3a). As a result, displaying a 3D scene on a stereoscopic projection wall limits the number of annotations that can be displayed at reasonable readability. At the same time, the annotations have to be comparably large, a fact which can lead to substantial interference with the 3D contents it is overlaid onto.

2 The *five* Display System

We propose solving the problem by using a hybrid display system that combines stereoscopic projection with an additional third projector dedicated to rendering 2D annotations. Fig. 1 illustrates a model of the display system we built. Fig. 2 shows a schematic view of it. Since the system combines 2D and 3D, we named it *five*. It is inspired by *focus-plus-context screens* [4], but extends the concept into *mixed dimensionalities*, i.e., the context area is not only of a different resolution, but also of a different dimensionality. The resulting system offers benefits for tasks that require stereoscopic presentation and high resolution. Various application scenarios can be found, in which the data to be visualized is inherently 3D as well as 2D. Construction datasets (like in car manufacturing, architectural design, urban planning, etc.) are challenging examples that contain three-dimensional geometries and also user manuals, measurements, etc. Thus, only an interweaved visualization can properly transport the needed contextual information.

As shown in Fig. 2, the first two projectors create a wide-angle stereoscopic image, i.e. one projector renders the scene from the perspective of the user's left eye, the other one from the right eye. Filters over each projector polarize the light emerging from them; users wear matching filter glasses; this assures that each of the user's eyes sees only the image intended for it. We are using a standard PowerWall™ system for that purpose.

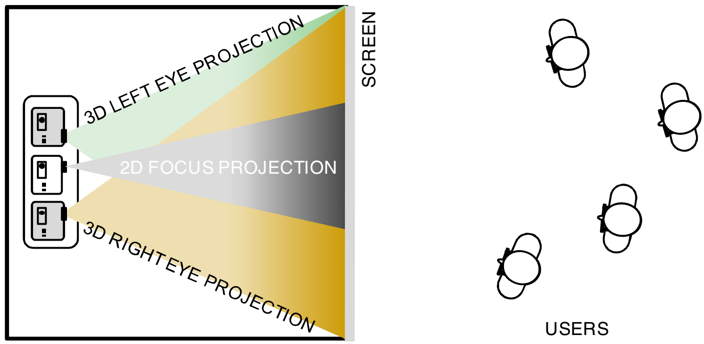


Fig. 2. Top view of the *five* prototype setup

The third projector does not use a polarization filter, which makes its projection visible to both eyes, and thus 2D. We focus this projection to a smaller region, resulting in an area of increased resolution. We typically point this *focus* projector to the display center, as shown in Fig. 2. This allows displaying information in high detail in the area of interest and does not affect three-dimensional context space by reason of its small size. By focusing the projector, light intensity per surface increases. We use a software-based filter to adapt the brightness of the focus projector to the same level as the one of the *context* projectors.

Compared to an all-high-resolution stereoscopic display, the resulting display has two limitations. Firstly, hi-res annotations can only be displayed in that area since the focus-projector covers only a sub-area of the wall display. Secondly, since both eyes see the same focus image, overlays offer no stereo disparity. Hi-res annotations are therefore flat and are always perceived as being located in the plane of the projection screen.

Before we display a scene on *five*, we partially alleviate these limitations by pre-processing the scene in the following five steps. (1) We translate the 3D scene such that as many annotations as possible fall onto the focus area. In Fig. 1, for example, we center a vanishing point over the focus area, as it holds the majority of objects in the scene. We then translate the scene in z-direction in order to position annotations as close to the projection plane as possible. This minimizes local z offsets, when we switch the annotation layer to 2D. We illustrate this process at the example of a document environment in the next section. (2) We extract the 2D annotations from the 3D scene. Since projection is based on the additive color model, we mask the space that held the annotation in the 3D model with a black area. Projecting onto the black mask with the focus projector will then result in the original brightness. (3) Where beneficial, we move additional annotations from the context area into the focus area. We link these annotations with the objects they describe, typically using lines turning them into *external labels*, as described in [7]. In cases where adding the line would introduce more clutter than can be saved by having it in the hi-res area, we do not translate that annotation but keep it low-res and stereo-projected. (4) We scale annotations down to reduce interfere with the 3D scene (Fig. 3c) or to be able to show additional annotations or additional details of an annotation (Fig. 3d).

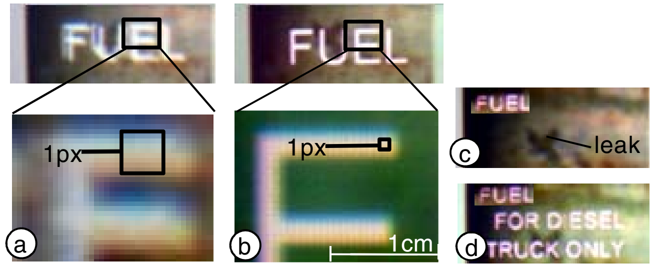


Fig. 3. (a) On a stereo projection wall: low resolution and alignment offsets impact readability of 2D text (b) Overlaying the text using a separate dedicated "2D" projector eliminates these artifacts. The underlying 3D scene is still rendered using the stereo projectors. (c) The increased resolution allows reducing the size of the 2D annotation, here revealing a leak in the fuel container. This minimizes interference with the underlying 3D scene or (d) can be used to overlay additional, here potentially vital 2D contents.

Alternatively, we render annotations at their original size to maximize readability of the overlay (Fig. 3b) or we mix all three objectives. (5) We render the scene.

In summary, delegating 2D annotations to a dedicated projector leads to a display that is more readable and/or more informative. The additional detail can be used to (1) reduce artifacts substantially, resulting in more readable overlays, as shown in Fig. 3b, (2) minimize the occlusion of 3D contents, as shown in Fig. 3c, (3) display additional annotations or to display annotations in additional detail, as shown in Fig. **3**d, or a combination of these.

Even though high-res annotations are limited to a certain display area and z-depth, adding the third projector significantly increases the amount of information that can be displayed and that can be directly associated with 3D on-screen objects. When regular stereoscopic projection runs out of annotation space, additional annotations can only be made available by flipping back and forth between multiple views or by switching back and forth visually between the display wall and an external display. As we found in the two studies mentioned earlier, retrieving such external annotations costs time and imposes cognitive load on users. Overlaying a focus projector, in contrast, increases the information density in situ, thereby minimizing switching effort and cognitive load.

3 Demo Application

To demonstrate our prototype and to illustrate the pre-processing process of 3D scenes, we implemented a visual document management application. It allows users to retrieve documents based on spatial memory. By mounting the focus projector in portrait mode we maximize the benefit for document viewing in this application.

Fig. 4 shows a walkthrough. In clockwise order: (a) In the default view, each document is shown as a thumbnail. Documents are visually grouped by shifting them horizontally and vertically. Proximity to the viewer, i.e. promotion in z, indicates relevance either absolute or with respect to a query entered by the user. The user selects a thumbnail to see that document in full, readable detail. (b) Resolution and

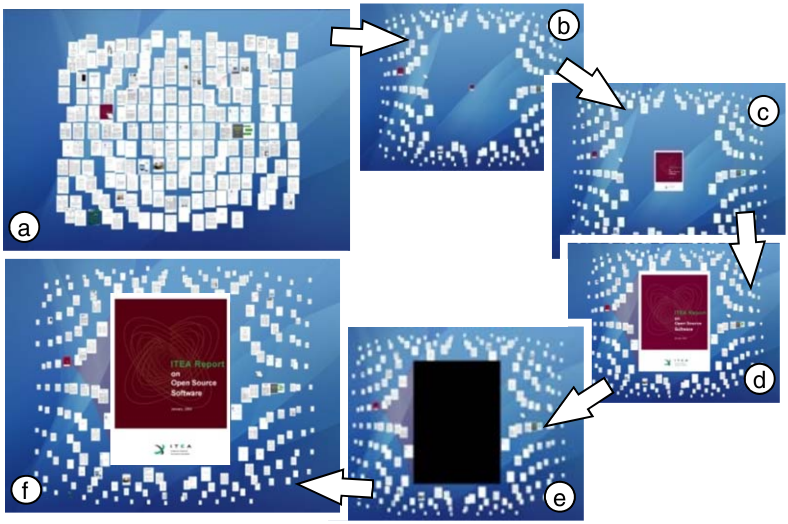


Fig. 4. (a) To view a document, (b) the document space is distorted to make space and the selected thumbnail moves to the center, (c-d) the thumbnail grows, and (e-f) is finally replaced by a readable high-resolution version of the page

sharpness are insufficient for displaying the selected stereoscopic projection in situ, but the embedded 2D focus projector offers sufficient resolution. The context space distorts to free up the focus area and the selected thumbnail is *copied* into the focus region, (c) where it is smoothly scaled (d) until it fills the entire focus area. At the same time, the document moves towards the user until it reaches the projection plane (stereo-disparity zero). At this point, the entire display is still projected stereoscopically. (e) The focus part of the 3D projects is faded to black and (f) the 2D focus area is faded in, turning the document high-resolution.

To visually clarify which thumbnail was moved, the system can alternatively link the thumbnail copy to its original using a *rubber band*, as shown in Fig. 4f (see *drag-and-pop* [5]).

4 The Five Prototype

Our prototype is based on a PowerWall™, which serves as the 3D context display. Two JVC DLA-SX21S projectors with a 1400 x 1050 resolution rear-project the stereoscopic image pairs onto a 2.9m x 2.3m semi-translucent screen through circular polarizing filters of opposite handedness. A Windows PC controls the stereoscopic visualization. It features two NVIDIA GeForce 8800 GTX graphics cards, 2 GByte main memory as well as two Intel Core 2 Duo CPUs with 2.4 GHz each.

The focus display is implemented using an Epson EMP-TW 1000 projector, located below the two JVC projectors. It offers a resolution of 1920 x 1080 pixels. We focused it on an area of 1.03m x 0.58m. As a result, a context pixel corresponds in size to approximately 5x5 focus pixels (Fig. 3).

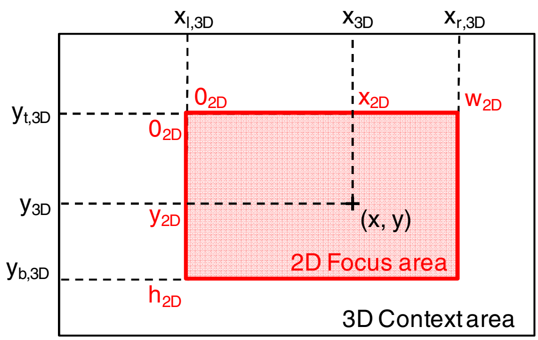


Fig. 5. Calculating the position of the 2D focus area within the focal plane of the 3D context

five runs on top of our OpenGL and C++ based framework called *AnyScreen* on MS Windows. It controls all three projectors. Moreover, it offers four stereoscopic (side-by-side, horizontally and vertically interlaced, and anaglyph stereoscopy) and one non-stereoscopic rendering mode for each render window. In the traditional stereo configuration, we set up the two stereo projectors in side-by-side mode. We add the focus projector as a third screen located in the center of the stereoscopic projection (Fig. 2).

In order to integrate the 2D focus projection seamlessly into the 3D virtual environment, the position of the focus region is configured using the following calibration process: Using the nomenclature from Fig. 5, the calibration process determines the coordinates (x_{3D}, y_{3D}, z_{3D}) in the 3D context that correspond to the coordinates (x_{2D}, y_{2D}) in the 2D focus region. Using this correspondence, we compute the 3D context coordinates for every object in the 2D focus region.

Since the 2D focus region is located on the projection plane, its z value is predefined. We have set up our 3D coordinate system in a way that the viewing direction is always along the negative z-Axis and the projection plane is orthogonal to the viewing direction. Thus, in this system the projection plane is described by $z = const$. Therefore, the z_{3D} coordinate is easily determined for every object that is lying in the 2D focus region.

For calculating the x_{3D} and y_{3D} coordinates, the 3D coordinates $(x_{l,3D}, y_{t,3D})$ and $(x_{r,3D}, y_{b,3D})$ that correspond to the upper left-hand and lower right-hand corners of the 2D focus region have to be determined. Once this information is available, the corresponding x and y 3D-coordinates for any pair (x_{2D}, y_{2D}) of 2D-coordinates can be computed as:

$$\begin{pmatrix} x_{3D} \\ y_{3D} \end{pmatrix} = \begin{pmatrix} x_{l,3D} \\ y_{t,3D} \end{pmatrix} + \begin{pmatrix} x_{2D} w_{2D}^{-1} \left(x_{r,3D} - x_{l,3D} \right) \\ y_{2D} h_{2D}^{-1} \left(y_{b,3D} - y_{t,3D} \right) \end{pmatrix}$$

To determine the coordinates $(x_{l,3D}, y_{t,3D})$ and $(x_{r,3D}, y_{b,3D})$ of the corners of the 2D focus area calibration, a calibration dialog is displayed on the focus and the context projectors.

5 Related Work

The purpose of the *five* system is to display additional information in limited space. It therefore addresses the same problem as corresponding software techniques (zooming, overviews, and fisheyes) and in particular focus-plus-context screens.

5.1 Zooming, Overviews, and Fisheyes

Zoomable interfaces [15] allow users to navigate information sequentially by panning and zooming. Semantic zooming makes it possible to assign different appearances to an object at different zoom factors. While zooming into a map, for example, additional detail may be revealed. Other variants include, *constant density zooming* [22] and *non-linear* panning and zooming [11].

Overview-and-detail interfaces consist of two or more display windows, one of which typically shows the entire information space, while one or more detail views provide a close-up view [12, 16]. These interfaces have been applied to a variety of applications, including program code editors [7] and image collections [16]. Positioning the detail view from within the overview allows efficient navigation [6]. In addition, the overview window helps users to keep track of their current position in the information space [16] so that they have the feeling that they are in control [21]. On the flipside, overviews require additional screen space and users are forced to switch between views to reorient themselves, which costs time [3].

Focus-plus-context techniques, in contrast, avoid switching between windows by combining detail area and periphery into a single view (e.g., the *bifocal display* by Apperley and Spence [1]). Examples for focus-plus-context techniques are *fisheye views* [9, 10], the *perspective wall* [14], and *document lens* [17]. The main drawback of focus-plus-context approaches is that they introduce visual distortion, which can interfere with the user's task [4, 13].

5.2 Focus-Plus-Context Screens

This shortcoming is addressed by *focus-plus-context screens* [4], large low-resolution screens with an embedded high-resolution screen. Information is displayed across both display units, so that scale is preserved, while resolution varies. As a result, distortion is avoided. The original screen device was developed by Baudisch et al [4]. In a quantitative evaluation [3], participants performed map-navigation and chip-design tasks faster when using a focus-plus-context screen than when using a zooming or overview interface with the same number of pixels.

Several researchers have adapted and enhanced the idea of focus-plus-context screens. Ashdown created a more flexible system by using projectors for both focus and context [2]. Sanneblad and Holmquist made the focus area moveable by combining a high-resolution tablet with a large projected screen [19]. Flider and Bailey physically separated the focus screen from the context screen [8]. However, none of the current research has addressed the usage of the focus-plus-context screen in stereoscopic environments.

In order to verify our design, we conducted two user studies. The first one evaluates *five* in the context of a static document, the second in the context of a dynamic, real-time task.

6 User Study 1: Visual Search in a Static Scene

In this study, participants performed a visual search task on a static document: participants located five "buildings" of a given height within a map of simple, abstract buildings of different heights. Participants used a pan-and-zoom interface, an overview-and-detail interface, and an interface based on *five*.

Our main hypothesis was that participants would complete the task faster when using the *five* interface than when using the pan-and-zoom interface (which required manual navigation) and the overview-and-detail interface (which required users to reorient themselves when visually switching between views). We expected these results to be reflected in terms of subjective satisfaction as well. Given that participants could verify findings before committing to them, we expected negligible error rates across interfaces.

Participants. Twelve participants volunteered. Participants were students of the local university or technical/non-technical staff of our department. Participants ranged in age from 23 to 40 years, average 30.92. All participants had normal or corrected-to-normal vision.

Task. During each trial, participants were presented with a birds-eye view of an abstract 75x50 cell "map" containing 90 tall "buildings", as shown in Fig. 6a. Participants were also given a height value. Their task was to locate the five buildings that were of the specified height on the 90-building map.

Participants could *estimate* building heights, e.g. based on the 3D cues provided by the stereo projection system. However, to be certain, participants had to look up the

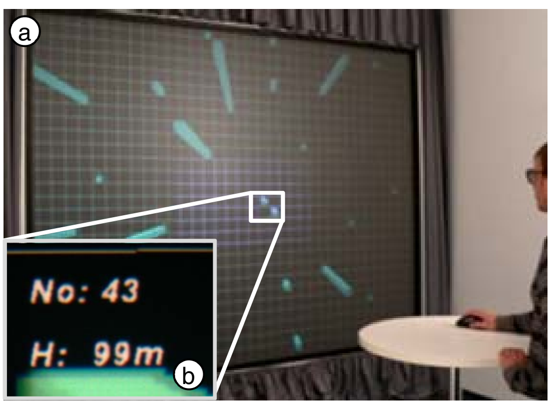


Fig. 6. (a) The *five* interface condition allowed participants to use 3D depth cues to locate candidate "buildings" of the sought height. (b) The exact height and the building's number were annotated onto the building's base (here shown in focus view)

exact height, which was written onto the building's base as shown in Fig. 6b. Whenever a building of the sought height was found, participants would call out the building number shown just above the building height.

Height and number annotations were scaled to fit the width of the building. We thereby simulated the space limitations of a real city where all cells would be filled, preventing text annotations to grow into the space of neighboring cells. Both annotations were readable while in the 2D focus area, but not readable while located in the 3D context.

Interfaces. There were three interface conditions: *pan-and-zoom*, *overview-and-detail*, and *five*. All interface conditions were implemented using the hardware setup described earlier. Participants controlled the system using a mouse located on a 1.1m high desk about 2.5 meters right of the center of the projection as shown in Fig. 6a.

In the *pan-and-zoom* condition, the skyscraper scene including buildings and annotations was rendered using the stereo projectors. In order to read annotations, participants had to zoom in by a factor of 2.8 or more by rolling the mouse wheel. As a result, participant alternated between zooming out to locate a candidate building and zooming in to examine the exact height and determine the building number. Alternatively, participants could pan the display by moving the mouse with the left button pressed.

Also the *overview-and-detail* used only the stereo projectors to render the scene. Fig. 7 shows how this interface condition was set up. The stereoscopic display wall always showed a zoomed out perspective on the 3D scene. The additional *detail* projection located right of the display wall was zoomed in deep enough to show annotations. A red marker frame in the 3D representation indicated which region of the scene was shown in the detail view. When the 3D scene was panned, the detail in the additional projector changed accordingly.

The *five* interface displayed the scene also using the stereo projectors. Annotations, however, were rendered using a focus projector directly onto the base of the respective buildings whenever that building was located inside the focus area; otherwise no

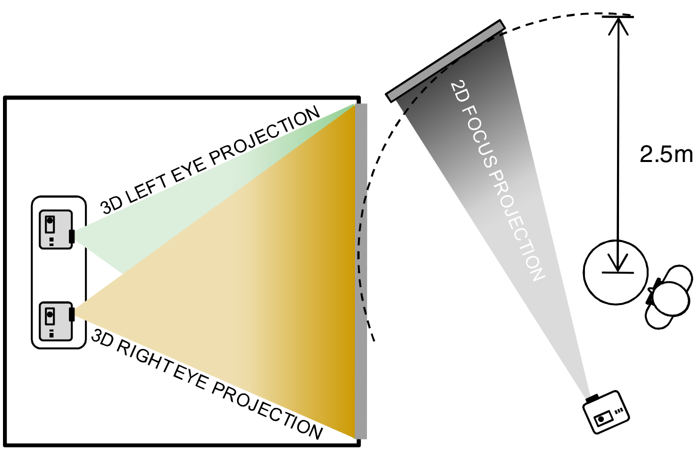


Fig. 7. Top view of the overview-and-detail interface

annotation was shown. As in the other conditions, participants panned using the mouse. The *five* interface did not support zooming.

Procedure. We used a within-subject experimental design, e.g. each participant carried out each task on all three interfaces. Interface order was counterbalanced following a Latin square of order 3. Maps were pre-computed and their order randomized. For each trial, we recorded task time and error rate (calling out the wrong building number).

Participants received a verbal explanation and training for each interface condition. Participants filled in a questionnaire after completing each interface condition. It contained questions on the participants' general impression of this interface (how comfortable / how easy to use / how fatigue-proof is it), its usability, and its suitability to the task, provided on a 7-item Likert scale. The study took less than 30 minutes per participant including training.

Results and Discussion. As expected, participants did not make any errors, across interface conditions. Fig. 8 shows overall task times.

We performed a one-way repeated measure ANOVA evaluating the within-subjects effect of display type (*five*, overview-and-detail and pan-and-zoom). The sphericity assumption was not met so the Huynh-Feldt correction was applied. The main effect of display type was significant, $F(1.476, 16.236)=16.997$, $p<0.005$.

Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons. The comparison revealed that the *five* interface yielded significant faster completion times than both the overview-and-detail ($p<0.001$) and the pan-and-zoom ($p<0.001$). However, the mean task completion time of overview-and-detail and pan-and-zoom was not significantly different ($p=0.182$).

As hypothesized, participants performed the visual search task faster when using the *five* interface than when using the two competing interface conditions. Subjective ratings reflect these findings. We noted that the participants definitely preferred the *five* interface over both other ones. Summarizing, all participants rated the *five* interface most applicable to the given task (Question 5). Only a slight difference occurred between the ratings of the overview-and-detail and pan-and-zoom interface.

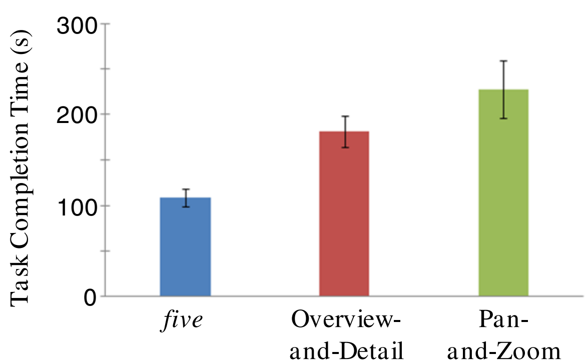


Fig. 8. Average task completion time in seconds (+/- standard error of the mean)

A repeated-measures ANOVA revealed a significant main effect of display ($F(2,22)=17.868$, $p<0.001$). Post-hoc comparisons using Bonferroni adjustment revealed that the *five* interface was more suitable for the task than the overview-and-detail one ($p<.001$) and the pan-and-zoom one ($p<.01$).

The performance benefit of the *five* interface over the pan-and-zoom condition was expected, as it was predictable that panning buildings into the focus area would be faster than zooming in and out.

The performance benefit compared to the overview-and-detail interface, however, does suggest that displaying information in situ leads to performance benefits. These findings are in accordance with [4], who found that requiring users to visually switch between overview-and-detail view in a 2D display system influences task performance.

The absence of time constraints allowed participants to complete the given task without error on all interface conditions. To investigate performance under time constraints, we conducted a second study. For this second study, we also used a more realistic scene geometry: in the first study, all annotations were located in the same plane; this prevented the introduction of depth error when flattening the annotation space to the 2D view. In the second study, we therefore used a scene with annotations along substantial depth.

7 User Study 2: Real-Time Interaction, Dynamic Scene

In the second user study, we investigated whether the *five* interface provides an advantage over overview-and-detail interface in the context of a dynamic scene. Participants performed a simplified real-time flight simulation that required them to simultaneously pay attention to both a 3D task (avoiding asteroids) and a 2D detail task (pick-up the right types of power-ups labeled in 2D located along the path).

We expected participants to obtain higher scores (more fuel, fewer trashcans and asteroids) when using the *five* condition. The reason was that the overview-and-detail condition would require participants to visually switch between views, introducing brief moments of reduced orientation. Again, we expected subjective satisfaction ratings to mirror this.

Participants. The same twelve participants from User Study 1 took part in User Study 2. We accepted that participants had additional experience using the different display setups, because potential training effects would have applied equally to both interface conditions.

Task. Participants were flying a simulated spacecraft, as shown in Fig. 9. The participants' task was to pick up fuel canisters (by simply flying through them) while avoiding trash canisters and asteroids. In order to do so, participants switched between four "lanes" arranged in a 2x2 grid (top left, top right, bottom left, and bottom right) using the arrow keys on a keyboard.

Fuel canisters and trashcans were floating in the middle of one of the lanes. Both were represented by similar looking blocks, labeled "fuel" or "trash" respectively. Fuel and trash canisters appeared nearly indistinguishable, then turned distinguished in the focus/detail view, and finally turned distinguishable also in the

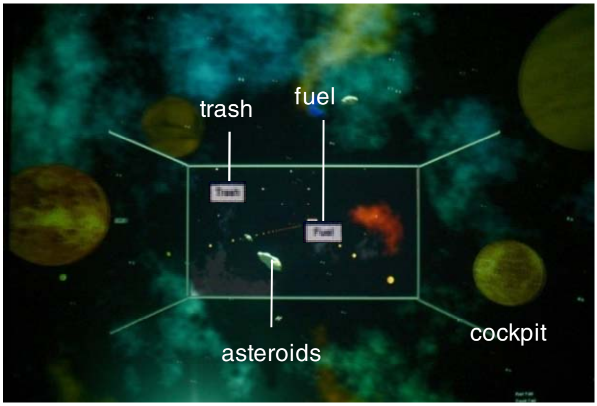


Fig. 9. Participant's task was to pick up fuel canisters while avoiding trashcans and asteroids in real-time (the four white labels and lines are part of the figure and were not shown on screen)

context/overview, approximately one second before flying through them. Asteroids appeared in pairs from different sides of the lane and tumbled across lanes.

The participants were informed that fuel canisters were worth one point, trash cans a minus point, and asteroid hits five minus points. For each interface condition, participants played a 90-second sequence containing 40 fuel canisters, 40 trashcans, and 50 asteroids (25 pairs). Sequences were generated randomly for every user and trial.

Interfaces. There were two interface conditions: The overview-and-detail interface corresponding to the overview-and-detail interface from the first user study. However, the only interaction possible was lane switching. Participants switched lanes using the arrow keys on the keyboard. The *five* condition corresponded to the *five* condition in the first user study. As in the overview-and-detail, participants navigated using arrow keys. We did not include the pan-and-zoom interface, because the speed of the dynamic nature of the task left no time for zooming.

Procedure. As in the first study, participants received instructions and training, completed the task on all interface conditions in counter balanced order, and filled in the same questionnaire. For each trial we recorded the number of picked–up fuel canisters as well as the collisions with asteroids and trash canisters. The study took between 10 and 15 minutes per participant.

Results and Discussion. Fig. 10 summarizes error rates. A repeated-measure ANOVA evaluating the within-subjects effects (*five,* O+D) found a highly significant mean effect of interface, for all three performance indicators. When using the *five* interface, participants collected more fuel packages (45% more, $F(1,11)=39.105$, $p<.001$) while colliding with fewer trashcans (47%, $F(1,11)=113.626$, $p<.001$) and asteroids (51%, $F(1,11)=55.579$, $p<.001$) than when using the *overview-and-detail* interface.

As expected, subjective preferences reflected these findings. Again, we could note that the participants definitely preferred the *five* interface over the other one. Summarizing, all participants rated the *five* interface most applicable to the given task. A

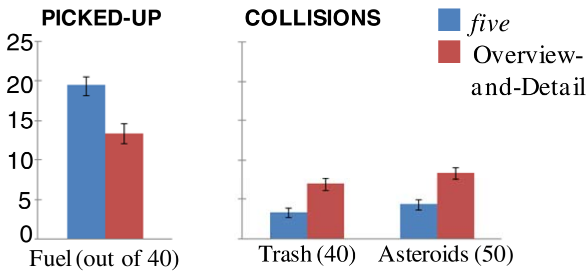


Fig. 10. Mean numbers of fuel packages picked-up (higher is better) and obstacles collided-with (lower is better) during flight task (+/- standard error of the mean)

repeated-measures ANOVA showed a significant effect for interface type $F(1,11)=17.963$, $p<.001$.

Unlike our first study, the real-time constraints caused the performance difference to impact error rate rather than task time (which was controlled). Our findings support our main hypothesis: Displaying 2D focus and 3D context in a single integrated view helps users keep track and stay in control in interactive applications. The overview interface, in contrast required users to switch visual attention back and forth, which interfered with the double attention task. These findings broaden the results given in [3] by adding dimensionality as an additional parameter.

8 Conclusion

In this paper, we have presented a hybrid display system. By complementing wide-angle stereoscopic projection with a 2D "focus projector" the resulting display setup allows to display detailed annotations *collocated with* the objects they are referring to, thus minimizing clutter and cognitive load. The findings of our two user studies suggest that a display like *five* can offer benefits when working with static as well as dynamic real-time applications.

Encouraged by our findings, we have started to adapt additional applications to the *five* display system. A commercial first person shooter (Unreal Tournament 2004) is now running on the *five* display. The additional 2D projector makes the game's heads-up display legible; alternatively, we could have used the increased clarity of the 2D projector to reduce the overlay size, thus reducing occlusion of 3D game contents.

As future work, we are planning to design visualization and interaction techniques that allow users to interact with objects and annotation on the *five* display system more easily.

Acknowledgments. We thank Patrick Baudisch (Microsoft Research) for the close cooperation, fruitful discussions, and constructive comments. We also thank all our study participants. This research has been funded by BMBF.

References

1. Apperley, M.D., Spence, R.: A bifocal display technique for data presentation. In: Proc. Eurographics 1982, pp. 27–43 (1982)
2. Ashdown, M., Robinson, P.: Escritoire: A Personal Projected Display. IEEE Multimedia 12(1), 34–42 (2005)
3. Baudisch, P., Good, N., Bellotti, V., Schraedley, P.: Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. In: Proc. CHI 2002, pp. 259–266 (2002)
4. Baudisch, P., Good, N., Stewart, P.: Focus plus context screens: combining display technology with visualization techniques. In: Proc. UIST 2001, pp. 31–40 (2001)
5. Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., Zierlinger, A.: Drag-and-Pop and Drag-and-Pick: Techniques for Accessing Remote Screen Content on Touch- and Pen-operated Systems. In: Proc. Interact 2003, pp. 57–64 (2003)
6. Beard, D., Walker, J.: Navigational Techniques to Improve the Display of Large Two-Dimensional Spaces. Behaviour and Information Technology 9(6), 451–466 (1990)
7. Bell, B., Feiner, S., Höllerer, T.: View management for virtual and augmented reality. In: Proc. UIST 2001, pp. 101–110 (2001)
8. Flider, M.J., Bailey, B.P.: An evaluation of techniques for controlling focus+context screens. In: Proc. GI 2004, pp. 135–144 (2004)
9. Furnas, G.W.: Generalized Fisheye Views. In: Proc. CHI 1986, pp. 16–23 (1986)
10. Furnas, G.W.: The Fisheye View: A New Look at Structured Files. In: Card, Mackinlay, Shneiderman (eds.) Information Visualization: Using Vision to Think, pp. 145–152. Morgan Kaufmann Publishers, San Francisco (1999)
11. Furnas, G.W., Zhang, X.: MuSE: a multiscale editor. In: Proc. UIST 1998, pp. 107–116 (1998)
12. Hornbaek, K., Frokjaer, E.: Reading of electronic documents: the usability of linear, fisheye, and overview+detail interfaces. In: Proc. CHI 2001, pp. 293–300 (2001)
13. Leung, Y.K., Apperley, M.D.: A review and taxonomy of distortion-oriented presentation techniques. ACM TOCHI 1(2), 126–160 (1994)
14. Mackinlay, J.D., Robertson, G.G., Card, S.K.: The perspective wall: detail and context smoothly integrated. In: Proc. CHI 1991, pp. 173–176 (1991)
15. Perlin, K., Fox, D.: Pad: An Alternative Approach to the Computer Interface. In: Proc. SIGGRAPH 1993, pp. 57–64 (1993)
16. Plaisant, C., Carr, D., Shneiderman, B.: Image-browser taxonomy and guidelines for designers. IEEE Software 12(2), 21–32 (1995)
17. Robertson, G.G., Mackinlay, J.D.: The Document Lens. In: Card, Mackinlay, Shneiderman (eds.) Information Visualization: Using Vision to Think, pp. 562–569. Morgan Kaufmann Publishers, San Francisco (1999)
18. Russel, D.M., Stefik, M.J., Pirolli, P., Card, S.K.: The cost structure of sensemaking. In: Proc. CHI 1993, pp. 269–276 (1993)
19. Sanneblad, J., Holmquist, L.E.: Ubiquitous Graphics: Combining Hand-held and Wall-size Displays to Interact with Large Images. In: Proc. AVI 2006, pp. 373–377 (2006)
20. Shneiderman, B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization. In: Proc. IEEE Symposium on Visual Languages (VL 1996), pp. 336–343 (1996)
21. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, 4th edn. Addison-Wesley, Reading (2005)
22. Woodruff, A., Landay, J., Stonebreaker, M.: Constant Information Density in Zoomable Interfaces. In: Proc. AVI 1998, pp. 110–119 (1998)

Improving Window Switching Interfaces

Susanne Tak[1], Andy Cockburn[1], Keith Humm[1], David Ahlström[2],
Carl Gutwin[3], and Joey Scarr[1]

[1] Computer Science and Software Engineering, University of Canterbury,
Private Bag 4800, Christchurch 8140, New Zealand
susanne.tak@pg.canterbury.ac.nz, andy@cosc.canterbury.ac.nz,
{khu23,jls129}@student.canterbury.ac.nz
[2] Department of Informatics Systems, Klagenfurt University,
Universitätsstrasse 65-67, A-9020 Klagenfurt, Austria
david@isys.uni-klu.ac.at
[3] Department of Computer Science, University of Saskatchewan,
110 Science Place, Saskatoon, Saskatchewan, S7N 5C9, Canada
gutwin@cs.usask.ca

Abstract. Switching between windows on a computer is a frequent activity, but current switching mechanisms make it difficult to find items. We carried out a longitudinal study that recorded actual window switching behaviour. We found that window revisitation is very common, and that people spend most time working with a small set of windows and applications. We identify two design principles from these observations. First, spatial constancy in the layout of items in a switching interface can aid memorability and support revisitation. Second, gradually adjusting the size of application and window zones in a switcher can improve visibility and targeting for frequently-used items. We carried out two studies to confirm the value of these design ideas. The first showed that spatially stable layouts are significantly faster than the commonly-used recency layout. The second showed that gradual adjustments to accommodate new applications and windows do not reduce performance.

Keywords: window switching, revisitation patterns, spatial constancy.

1 Introduction

Switching between windows like email applications, word processors, and Web browsers is a very common task. A previous study [1] found that the mean time between window switches is only 20.9 seconds and that users have more than eight windows open more than 78% of the time. It has also been reported that the average number of simultaneously opened windows increases with available display space: from four for single monitor users to up to 18 for users with multiple monitors [2].

Current interface methods for window switching have changed relatively little since early graphical user interfaces – clicking on a window brings it into focus, as does selecting the window from a spatial iconic representation (e.g., the Windows Taskbar) or from a recency list (e.g., the Windows Alt+Tab display). Recent research and commercial systems demonstrate alternatives to these mechanisms, but problems

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 187–200, 2009.

still exist – with more than a few windows in the display, users must often carry out a laborious search to find the desired window, even if that window is used frequently.

Despite the different designs that have been proposed, no window switching tools are based on empirical evidence about how people revisit windows in actual desktop work. To address this limitation, and to identify new design principles for window switching tools, we carried out three studies. We first conducted a longitudinal study where window switching behaviour was recorded; this study showed that revisitation is frequent, both to applications and to specific windows, and that most switches are to a small number of applications. From this study, we identified the design principles of spatial constancy and morphing target sizes. Spatial constancy supports window revisitation by keeping thumbnails for activating windows and applications in the same place in the switching display, allowing users to build up spatial memory of frequently-used items. Morphing target sizes allocates more space to frequently-used applications and windows, improving Fitts' Law targeting time for the most frequent items, and allowing for the addition of new items.

Our second and third studies tested the value of these two principles. The second showed that spatial constancy is effective: stable layouts are significantly faster than recency layouts (similar to Windows Alt+Tab). The third study showed that gradual adjustments in the display's layout do not reduce (or improve) performance. We discuss how these successful results can be deployed in new window switching tools.

2 Related Work

Two areas of related work inform our investigation: research on task and window switching interfaces; and studies of user behaviour with desktops and windows.

2.1 Window Switching Interfaces

The importance and frequency of window switching has led to extensive research and development into improved interfaces for the task. Two interface properties help distinguish window switching approaches: first, the type of information used to form and display relationships between windows, such as temporal, spatial, or semantic data about the windows; and second, the degree to which systems try to automatically establish these relationships, with some being entirely manual while others use sophisticated predictive methods to automatically establish window relationships.

Henderson and Card's seminal work with the *Rooms* [3] virtual desktop manager was almost entirely manual, with the user assuming all responsibility for the spatial placement of windows within a room based metaphor. *Scalable Fabric* [4] also uses manual relationship controls, including extensive support for zooming, but unlike *Rooms* it provides little explicit structure for grouping windows. The default desktop behaviour of Microsoft Windows XP/Vista is also largely manual (users place windows where they wish), although it automatically groups windows belonging to each application in the Taskbar. Microsoft's *GroupBar* [2] maintains manual control, but replaces the Taskbar's application-based grouping with user-defined task groups.

The primary limitation with manual control of window relationships is that users must carry out explicit actions to gain potential benefits. To remove the dependence on manual actions many systems automatically form window relationships. The most basic and widely deployed automatic relationship system is the familiar Alt+Tab key binding in Microsoft Windows operating systems. Alt+Tab allows users to rapidly flip through a temporally based 'z-ordering' of windows on the display. Kumar *et al.* [5] observe that Alt+Tab is very efficient when the number of windows is low, but researchers have also labelled the method 'tedious' [6]. Mac OS X's Exposé also uses an automatic layout, displaying thumbnails of all windows at once; but the layout is not spatially constant, so users cannot accurately predict where specific windows will be located, demanding visual search to find them.

SWISH [7], *UMEA* [8], *TaskTracer* [9], *RelAltTab* [10] and *Taskposé* [11] all use sophisticated methods to automatically adapt to user activities. *SWISH* uses temporal relationships between window focus events as well as window titles to establish semantic relationships. Their evaluations suggested 70% accuracy rates in assigning windows to task groups. Similarly, *TaskTracer* uses machine learning to modify the Microsoft Start menu, Taskbar and Windows Explorer. *RelAltTab* uses similar methods to modify the Alt+Tab window order. *Taskposé* provides an overview in which windows drift towards each other dependent on their temporal relationships, based on the results of a 'WindowRank' algorithm. *Taskposé* windows also gradually enlarge to reflect their relative importance as calculated by the algorithm.

The primary limitations of automatically adaptive systems are that they can incorrectly predict the user's intention and that users can fail to understand or anticipate the system's adaptation [12]. When this happens users must resort to time-consuming visual search of candidate targets. Of all the previous designs, however, only two – *WindowScape* [13] and *Elastic Windows* [14] – used stable layouts to help improve memorability of previously-used windows.

2.2 Studies of Window Use

Gaylin [15] provides an analysis of window activities from the early days of graphical user interfaces, based on 22 minute observations of nine participants. Observations show that window switching activities were far more frequent than window creation, deletion, or geometry management. Hutchings *et al.* [1] update these findings using automatic logs of 39 participants over a 3-week period. They report on how window management activities differ across single- and multi-monitor display use (in common with Grudin's earlier field study [6]) but their data also highlights important general characteristics of window management. This includes the finding that window switching is extremely frequent, with a mean window activation time of 20.9 seconds, and a median of only 3.77 seconds. This frenetic frequency of window switching is confirmed by Mackinlay and Royer [16], who also conducted a log analysis of window switching. Hutching *et al.'s* data also shows that users normally have many windows open, with eight or more windows open 78.1% of the time.

Many types of human behaviour are highly repetitive, as observed by Zipf's Law [17] and the Pareto Principle [18] (also called the "80-20" rule), where a large portion of effects comes from a small portion of causes. Zipfian distributions have been observed in many areas of computer use [19], such as frequency of command use and menu use. Although window and application switching activities are among the most frequent in computer use, with one study showing a mean time of 21 seconds between actions [1], we are unaware of any empirical studies of users' patterns of revisitation to windows and applications. To address this limitation, we carried out the longitudinal study described in the following section.

3 Log Study of Application and Window Switching

To find out how users switch between applications and windows, we carried out a longitudinal study that recorded windowing behaviour as users went about their everyday tasks. We developed logging software for Windows XP that unobtrusively monitors window switches, window creation/destruction events, changes in window geometry, and the method used to switch windows (e.g., Alt+Tab, Taskbar, and direct mouse click). Nine frequent computer users (18 to 55 years old) took part in a study during which we recorded between 8 and 117 days of data per participant. Four participants used dual monitors; five, single monitors (see Table 1).

Overall, we obtained 241 person-days of data. Only manual window switches were included in the analysis: automatic switches (such as window/dialog pop-ups) were removed from the data. This left a total of 45,377 switch events.

3.1 Results

Results are divided into three categories. First, we present data describing revisitation patterns for the windows used each day. This characterises the users' main activities in window switching. Second, we describe long term revisitation patterns to applications. While many windows are transient, existing only for immediate work requirements (such as a window containing an email message during its composition) applications are relatively stable. For many tasks, such as "search on the Web for topic X" or "check my email inbox" the user's target is likely to be the application (e.g., Firefox or Outlook) rather than a specific window, so an application based analysis is potentially informative. Third, we analyse the interface mechanisms used to carry out window switching activities to gain insights to how current interfaces are used and to determine whether users adopt similar or divergent patterns of behaviour.

Daily Window Revisitation. For each participant on each day, we analysed how frequently each window was revisited. This was conducted by forming a ranked order of windows according to their percentage of total daily switches.

The number of window switches per day for participants ranged from 5 to 807, with a cross participant mean of 219 per day (*s.d.* 91). The number of distinct windows switched to per day ranged from 3 to 177, with a cross participant mean of 39 (*s.d.* 19). In their study of Web revisitation, Tauscher and Greenberg [20] define the *recurrence rate R* as the probability that any URL visit is a repeat of a previous visit, giving *R=(total URLs visited – distinct URLs visited)/total URLs visited* ×100. Adapting this formula for window revisitation gives a mean recurrence rate of 82%; much higher than the 58-61% rate reported for the Web. This data shows that window revisitation is a very common activity.

Analysing the same data with respect to the Pareto principle shows that 80% of window switches were triggered by between 24 and 40% of windows for the different participants (see Table 1), with a mean of 35.1%.

Table 1. Application and window switching data for the nine participants

					Pareto. 80% of switches by what %		**Interface %**		
P.	*single/ dual*	*# days*	*# switch*	*# apps*	*Windows*	*Apps*	*Click*	*Task-bar*	*Alt+ Tab*
1	single	10	1396	40	40	15	23.4	76.7	0.0
2	dual	117	17088	45	34	7	78.4	21.4	0.2
3	single	23	2588	23	37	17	37.0	47.7	15.2
4	single	16	3019	49	37	16	35.3	61.3	3.4
5	dual	9	2454	37	40	22	64.5	35.5	0.0
6	dual	30	10373	54	31	13	64.1	35.5	0.4
7	dual	8	2922	34	35	18	82.7	14.5	2.8
8	single	13	2989	39	24	18	22.7	75.3	1.9
9	single	15	2548	23	37	17	7.6	7.7	84.7

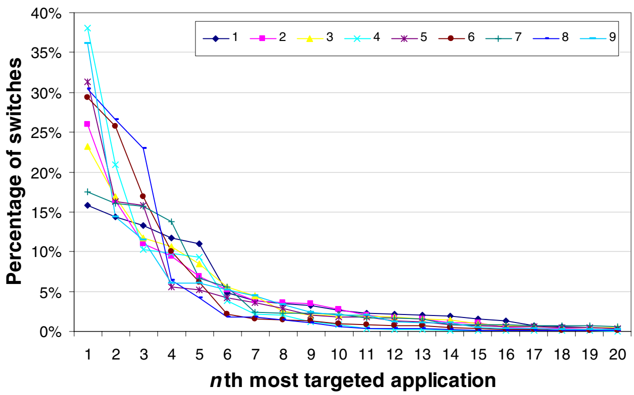


Fig. 1. Percentage of switches to the 20 most frequent applications for each participant

Revisitation to Applications across Sessions. As mentioned above, many user tasks involve targeting an application rather than a specific window. For example, a user may have several Web browser windows concurrently open, any of which can be used for a pressing search task; or the user may need to target a calculator, file explorer, etc. In each case the application is the target, not a specific window. We therefore analysed each participant's revisitation to applications for the duration of the study.

The number of distinct applications used by the participants ranged from 23 to 54, with a mean of 38 (*s.d.* 10.6). For each participant the total number of switches to each application was calculated, ranked by frequency, and converted to a percentage of their total application switches. Fig. 1 shows this data for each participant. It is clear that a few applications are used a lot, and a lot are used relatively little.

Table 1 shows Pareto principle data, revealing that most users' application revisitation roughly adheres to the 80-20 rule. For example, 22% of participant five's applications accounted for 80% of switches. All other participants' application revisitation was *more* pronounced than the Pareto principle predicts: e.g., only 7% of participant two's application switches accounted for 80% of switches.

3.2 Interface Mechanisms Used to Switch Windows

We analysed the main three interface mechanisms currently used to visit and revisit windows: direct window clicks, selection from the Taskbar, and Alt+Tab. This analysis was conducted to determine whether users made similar or divergent use of the tools available. Table 1 shows each participant's use of these tools. Two important observations from this data are as follows. First, there are substantial differences between participants in their use of Alt+Tab. Seven of the participants used it very lightly (less than 3.5% of window activations) or not at all; one used it fairly often (15.2%); and one used it almost exclusively (84.7%). The two participants who used Alt+Tab heavily had a single monitor, suggesting that it might be most valuable for users with constrained screen real-estate (see Fig. 2). Second, direct clicks and Taskbar use were also influenced by screen real-estate, with dual monitor participants using direct clicks more and the Taskbar less than single monitor participants.

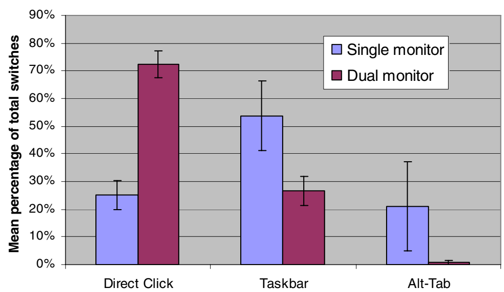


Fig. 2. Cross participant means of the percentage of window switches activated by clicking on the window, Taskbar selections, and Alt+Tab. Error bars show ±1 standard error.

4 Design Principles for Window Switchers

The patterns observed in the log study strongly suggest that supporting revisitation should be a main design goal in window switching tools. More specifically, this support should allow users to quickly find and distinguish between previously-visited windows in a switcher's display. There are several possible ways to provide this support (for example, previously-visited windows could be highlighted in the display, or the most frequently-used windows could be shown first), but previous research in psychology argues for an approach that makes use of *spatial constancy*, discussed in Section 4.1. However, user's patterns of behaviour change over time: for instance, a user may replace an application with a different vendor's system. Complete spatial constancy does not allow for the addition/removal of items. We address this problem with the use of *size morphing* (Section 4.2). Morphing target sizes allow for the addition of new items while maintaining as much spatial stability as possible. Also, allocating more space to frequently-used applications and windows will reduce their Fitts' Law targeting times. Given that some applications and windows are used much more frequently than others this might increase overall performance.

4.1 Supporting Window and Application Revisitation with Spatial Constancy

Spatial location memory is a person's memory of where objects are in space. It is well developed and can be extremely fast: studies have shown that items can be found in time proportional to the logarithm of the size of the set, which can be much faster than the linear search time needed for unorganised sets [19]. This fast performance is enabled by spatial constancy – that is, items remaining in the same location over time. This idea has been known since the first interface design guidelines [21], and many studies have demonstrated its effectiveness.

There is also evidence that spatial location memory in user interfaces is surprisingly robust. In studies of the Data Mountain's spatial layout of thumbnail images [22] users were able to quickly and accurately recall the location of specific targets four months after originally creating a layout of 100 images. Furthermore, their retrieval performance was not significantly harmed when the images were replaced with blank outlines. Spatial constancy, however, has received little attention in research on task- and window switching tools (with a few exceptions [13, 14]).

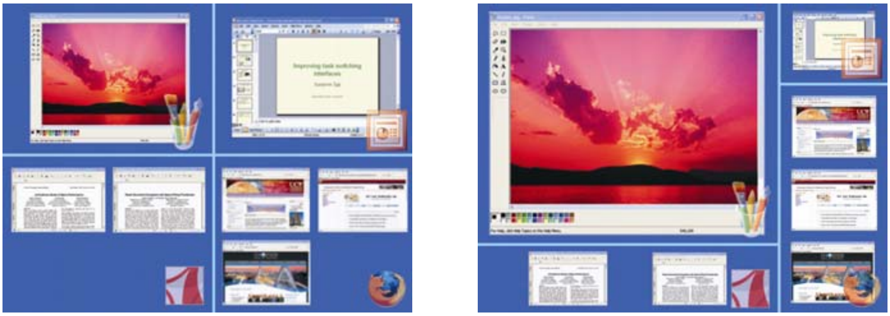
Applying the idea of spatial constancy in a window switcher implies that windows and applications should not move in the switcher's display. Within a work session, this design approach has clear advantages: since revisitation is strong, people will quickly learn the locations of the most frequently-used windows. Spatial stability offers similar advantages across work sessions. Since people regularly return to a relatively small number of applications, the problem of where to place individual windows in the switcher display can be solved by creating 'application zones' that are themselves spatially constant, and that reflect people's longer-term repeating work patterns. Spatial constancy can thus be applied in a hierarchical fashion: applications are given stable zones in the switcher display, since these change slowly over the long term; and within each zone, the application windows used in the current work session are placed in stable spatial locations. The application-based organisation provides an initial guide to the location of a new window, but as the window is used more and more frequently, users will start to remember its location as a separate entity.

4.2 Size Morphing to Accommodate Change and Optimise Performance

The design principle of spatial constancy potentially conflicts with changes in patterns of behaviour. For example, when users replace one application with another how can spatial stability be maintained? Also, given that some applications and windows are used much more frequently than others, how can the switching interface depict the relative importance of applications and windows, and optimise the acquisition of frequent targets?

Our proposed solution is to use gradual size 'morphing' to adjust the sizes of application zones and window thumbnails. Morphing avoids abrupt changes in layout that would damage spatial memory, while allowing the introduction of new items and enabling frequent targets to be enlarged to enhance their visibility and to reduce pointing time.

Fig. 3 shows a mock-up of our design, which uses the entire screen when activated. Clicking on a zone or thumbnail immediately switches to the associated application or window. Each zone contains thumbnails representing all windows associated with the



(a) Initial view. (b) After morphing.

Fig. 3. Prototype application zones, before and after size morphing. The application and window in the top left have been used most frequently.

application, scaled and tiled to fit. Initially all application zones are of equal size (see Fig. 3a), but they gradually morph in size to reflect frequency of use (see Fig. 3b).

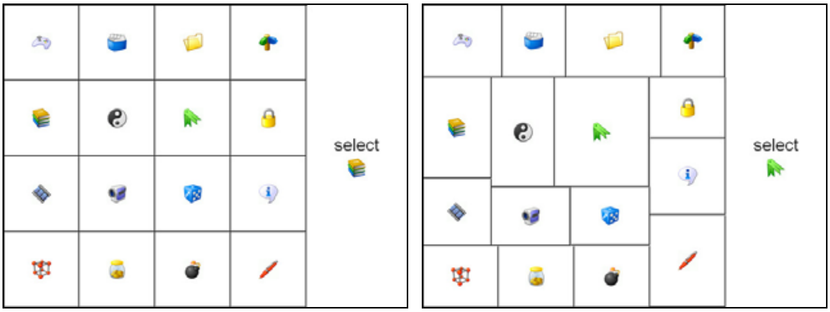
5 Evaluating Spatial Constancy and Morphing

Our design principles assume that users will perform better at switching to windows when thumbnails remain in spatially stable locations than when they are rearranged according to recency or other properties. It also assumes that size morphing will allow users to maintain their spatial memory while the interface adapts. We conducted two experiments to separately investigate these issues.

5.1 Experiment 1: Spatial, Recency, or Frequency Ordering?

A window switching interface can order items in a variety of ways, including spatially or by recency (like Alt+Tab). Other orders are also possible: frequency reordering might work well for Zipf-like distributions; random orders may perform well if visual popout effects are strong. This experiment, therefore, investigates the performance impact of four different orderings (spatially stable, recency order, frequency order, and random order) for tasks involving acquisition of targets in a Zipf-like distribution.

The experimental interface consisted of a grid of distinct icons (Fig. 4a) with a cued target on the right. Participants were instructed to click on the target icon region as quickly and accurately as possible, with each successful acquisition immediately cueing the next. The *spatially stable* layout used an arbitrarily but stable order (i.e., icons never moved). In the *random layout* all items were randomly repositioned after each selection. The *recency reordering* condition moved the most recently selected item to the top left of the grid, pushing earlier items along in row-major order (similar to Alt+Tab). Finally, *frequency reordering* repositioned items according to their cumulative selection counts (most frequent at top left, in row-major order). Note that frequency reordering rapidly stabilizes with a Zipf-like distribution of targets, while recency ordering is less stable.



(a) Experiment 1. (b) Experiment 2 (post morphing).

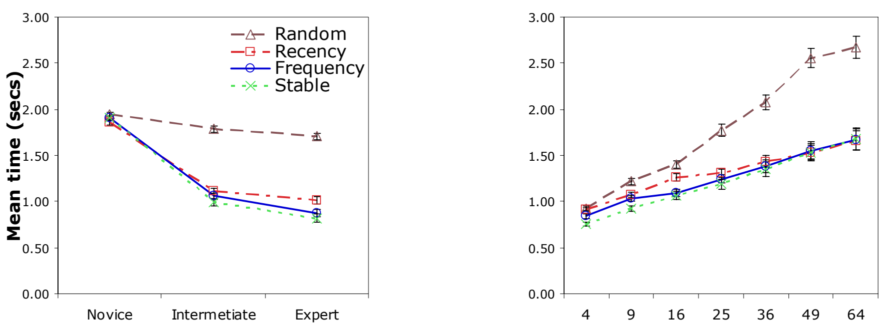
Fig. 4. Experimental interfaces showing 16 icon regions and the target cue on the right

To understand how performance with these interfaces is influenced by number of targets, participants completed trials with 4, 9, 16, 25, 36, 49 and 64-item grids. The Zipfian distribution of targets was generated by randomly selecting eight targets from among the candidates (all four for the 4 item grid): one was cued 10 times, one 5 times, then 3, 2, 2, 1, 1, and 1 for the others.

Twenty-six students volunteered for the experiment (16 male, 10 female, 16-44 years old). Their participation lasted approximately 45 minutes.

Results and Discussion. The selection time dependent measure was analysed using a 4×3×7 RM-ANOVA for factors *layout* (stable, frequency, recency, random), *experience* (novice, intermediate or expert) and *items* (4-64). Experience was determined by assigning first-time icon selections as novice, 2nd-7th selections as intermediate, and 8th-10th as expert.

All factors showed significant main effects: *layout* ($F_{3,60}$=72, p<.001), *experience* ($F_{2,40}$=409, p<.001) and *items* ($F_{6,120}$=135, p<.001) (see Fig. 5). Stable layouts were the fastest (mean 1.2s) followed by frequency reordering (1.3s), recency reordering (1.3s), and random (1.8s). Post hoc comparisons show pairwise differences between all layouts except frequency and recency reordering, and frequency reordering and



(a) Experience. (b) Number of items.

Fig. 5. Experiment 1 mean selection times (±1 standard error) with the four layouts

stable (Bonferroni, $p<.05$). Fig. 5a shows a significant *layout*×*experience* interaction ($F_{6,120}$=47, $p<.001$) caused by relatively constant performance across *experience* with random layouts in contrast to marked improvement with other layouts. Fig. 5b shows a significant *layout*×*items* interaction ($F_{18,360}$=14, $p<.001$), caused by the random layout worsening much more rapidly across increased number of items than the other three layouts. The results show that spatial constancy is beneficial. The stable, frequency, and recency layouts all supported expertise development; random did not. The recency layout (similar to Alt+Tab order) was outperformed by the stable layout.

5.2 Experiment 2: The Effect of Size Morphing

The second experiment examined the performance impact of the morphing behaviour used for two purposes: to allow new windows/applications to be introduced, and to change target sizes in response to the Zipf-like frequency distribution. Experiment 1 suggests that total stability is the 'gold standard', but total stability would prohibit the addition of zones for new windows/applications, as well as prohibiting morphing size adaptation to reduce the Fitts' Law targeting time.

We therefore experimentally compared target acquisition performance using totally stable placement (the gold standard control condition) as well as morphing behaviour implemented with *squarified* [23] and *spiral* [24] treemaps. Treemaps recursively divide 2D spaces into rectangles of various sizes, with size representing an underlying quantitative data attribute. Various algorithms for creating treemaps exist, and two main properties are the average aspect ratio and spatial stability (see [24] for a recent review). *Squarified* and *spiral* treemaps were used in this experiment because they respectively support low and high spatial stability. Fig. 4b shows the spiral experimental condition after morphing. The experimental method, procedure and design were similar to Experiment 1. There were seventeen participants (fifteen male, two female, 21-35 years old).

Results and Discussion. Task time was analysed using a 3×3×7 RM-ANOVA for factors *layout*, *experience*, and *items*. There were significant main effects for *items* ($F_{6,96}$=204, $p<.001$) and *experience* ($F_{2,32}$=444, $p<.001$), but not for *layout* ($F_{2,32}$=1, p=.3). Mean times for the three layouts were similar at 1.3s, 1.4s and 1.4s with stable, squarified and spiral respectively.

The absence of a significant difference between layouts is interesting and warrants further analysis. The stable interface is not a practical solution to window switching because it prohibits the addition of new applications and windows. However, total stability best supports users in exploiting their spatial memory, so it is interesting that it did *not* significantly outperform either of the two treemap layouts. It is reasonable to suspect that the treemap interfaces allowed users to benefit from reduced Fitts' Law targeting performance with frequently selected items (which gradually morphed to larger sizes), and that this counteracted the penalties associated with reduced absolute stability, but this explanation is not supported by a significant *layout*×*experience* interaction ($F_{4,64}$=2, p=.1). More powerful experimental analysis is required.

6 Discussion

The studies provide important findings for the design of window switching tools:

- Within a work session, people frequently revisit previously-used windows;
- Revisitation to a small set of applications is also strong across sessions;
- Most users rely heavily on mouse-based window switching methods, but some use keyboard methods almost exclusively;
- Spatially-constant item layouts are significantly faster than recency-based layouts for a Zipfian distribution of target stimuli;
- Gradual size 'morphing' of item areas in the display did not affect performance – importantly, morphing did not substantially harm the participant's spatial memory for target locations.

In the following paragraphs, we consider reasons why these design principles succeeded and discuss design issues for real-world window switching tools.

6.1 Why the Principles Succeeded

The results of Experiment 1 add to previous findings showing performance advantages for spatial constancy in comparison to alternatives. Furthermore, these are the first results (that we know of) comparing spatially stable 2D layouts with frequency and recency based layouts.

The stable layout was significantly faster than the recency layout, with performance benefits increasing with expertise. This is easily explained – after each selection in the recency layout users must either calculate the new position of their targets or visually search for them, both of which demand time.

Performance with stable and frequency layouts was similar to each other, raising the question of whether window switching interfaces should use frequency ordering. However, the experiment used a Zipfian distribution of stimuli, causing the frequency layout to quickly settle to spatial stability, so we believe that the frequency layout's comparative success is also explained by its spatial stability (after a short period of reorganisation). In practical use, a frequency layout for windows is unlikely to succeed due to the transient nature of most windows. Consequently, some other basis for organisation is required, such as application zones. Although application zones could be placed in a frequency layout, doing so would create placement instability during early use, which might cause users to discard the system due to its unpredictable behaviour. We therefore believe that using spatial stability as the main layout principle is a preferable solution. Another reason for believing application zones to be a useful placement strategy is that it quickly narrows the user's search space when there are several candidate windows from the same application – rather than having to search the whole display, users can quickly dismiss the majority of candidates by homing in on only those in the appropriate application zone.

The finding from the second experiment that morphing was no worse than absolute spatial stability is important because it suggests that gradual layout changes allow users to successfully use spatial memory while the display adapts to changes in their behaviour. Although we have not yet tested item addition and deletion, we believe that the spiral treemap algorithm used in Experiment 2 will allow users to maintain their spatial understanding of the layout. We will test this in future work. We will also model and test the Fitts' Law performance advantages gained from morphing.

6.2 Design Considerations for Real-World Switching Tools

The log study and experiments suggest that our design principles can be successful, but several design issues arise in translating these findings into a real-world switching tool, as discussed below.

Number of windows and applications. Fig. 3 portrays our design intention, showing four application zones and seven windows. However, the log study revealed that users actually work with dozens of applications and windows. Can our design scale? We are confident it can for several reasons. First, prior work (e.g., [22]) shows that users can successfully learn and remember the locations of hundreds of targets. Second, our experiments included up to 64 targets, and the benefits of spatial stability increased with the number of targets. Third, our design uses the full display space to show and access all windows and applications at once, allowing much larger targets than is possible with visually compact tools such as the Windows Taskbar.

Flexibility in input device and interface behaviour. The log study showed that most users relied on mouse input for window switching, but that one user used the Alt+Tab key combination almost exclusively. Convenience will clearly influence the choice of input device (e.g., using the mouse when a coffee cup is held in the other hand, or using Alt+Tab when typing to avoid repositioning the hands). However, Alt+Tab is also powerful when the user needs to 'flip' between a few recent windows. It is therefore desirable that next generation window switching tools continue to support flexibility in input devices and recency-based traversal. Our design can easily accommodate both. For example, the tool could be activated by a key combination (such as Alt+Tab) or by a dedicated mouse button or wheel; and the traditional Alt+Tab recency list can be traversed by highlighting items on subsequent key combination or button/wheel press.

Integrated support for application launch and window switching. Another potential advantage of our design (yet to be evaluated) is that it integrates support for different types of window switching tasks. In current interfaces, application launch facilities are largely partitioned from window switching tools, yet the user activities are closely related. For example, to "search on the Web for topic X" the user needs to acquire a browser window. If the user starts by searching the Taskbar they will be unsuccessful if no browser windows are active, necessitating a second action such as navigating through the Start menu hierarchy to launch the browser. Alternatively, if the user begins by launching the application, they will often gain a superfluous window (when others were already available) adding to their window management load. Our design, in contrast, provides a single interface mechanism for all window and application activities: if the application zone is empty, the user clicks to launch; but if windows are already available, any candidate can be immediately brought into focus.

Tailored layouts. Finally, the design could be adapted to allow users various manual controls, such as ensuring that certain application zones are always displayed in specific locations (perhaps to ensure consistency between a desktop and laptop display), or excluding particular applications from appearing in the layout.

7 Conclusions and Future Work

Log analysis of real window and application switching showed that revisitation (returning to previously used windows and applications) is an extremely frequent activity in computer use. It also showed that most users activate their window and application switches using the mouse as the input device, but that some users rely on keyboard methods almost exclusively. These empirical characterisations of window and application switching are new, and although they may confirm (or refute) designers' intuitions, empirical confirmation is necessary for informed design. The main design message of the log analysis is that window and application switching interfaces should improve efficiency by explicitly supporting revisitation.

We proposed two design principles to better support revisitation: spatial constancy of the interface controls to activate applications and the windows belonging to them; and size morphing to allow the spatial display to adapt to changes of behaviour and to optimise selection of frequent targets. Two empirical studies showed, first, that spatially stable layouts allow faster acquisition than recency and random layouts for skewed distributions such as those occurring in window switching tasks, and second, that size morphing is not significantly slower than the (idealistic but impractical) gold standard of absolute spatial stability.

In future work we will conduct a longitudinal study of user performance with a complete tool implementing the principles derived and tested in this paper.

References

1. Hutchings, D., Smith, G., Meyers, B., Czerwinski, M., Robertson, G.: Display Space Usage and Window Management Operation Comparisons between Single Monitor and Multiple Monitor Users. In: Proc. AVI 2004, pp. 32–39. ACM Press, New York (2004)
2. Smith, G., Baudisch, P., Robertson, G., Czerwinski, M., Meyers, B., Robbins, D., Horvitz, E., Andrews, D.: GroupBar: The TaskBar Evolved. In: Proc. OzCHI 2003, pp. 34–43 (2003)
3. Henderson, A., Card, S.: Rooms: The use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. ACM Transactions on Graphics 5(3), 211–243 (1986)
4. Robertson, G., Horvitz, E., Czerwinski, M., Baudisch, P., Hutchings, D., Meyers, B., Robbins, D., Smith, G.: Scalable Fabric: Flexible Task Management. In: Proc. AVI 2004, pp. 85–89. ACM Press, New York (2004)
5. Kumar, M., Paepcke, A., Winograd, T.: EyeExposé: Switching Applications with Your Eyes, Stanford University (2007), `http://hci.stanford.edu/cstr/reports/2007-02.pdf`
6. Grudin, J.: Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use. In: Proc. CHI 2001, pp. 458–465. ACM Press, New York (2001)
7. Oliver, N., Smith, G., Thakkar, C., Surendran, A.: SWISH: Semantic Analysis of Window Titles and Switching History. In: Proc. IUI 2006, pp. 194–201. ACM Press, New York (2006)
8. Kaptelinin, V.: UMEA: Translating Interaction Histories into Project Contexts. In: Proc. CHI 2003, pp. 353–360. ACM Press, New York (2003)

9. Dragunov, A., Dietterich, T., Johnsrude, K., McLaughlin, M., Li, L., Herlocker, J.: TaskTracer: A Desktop Environment to Support Multi-tasking Knowledge Workers. In: Proc. IUI 2005, pp. 75–82. ACM Press, New York (2005)
10. Oliver, N., Czerwinski, M., Smith, G., Roomp, K.: RelAltTab: Assisting users in switching windows. In: Proc. IUI 2006, pp. 385–388. ACM Press, New York (2006)
11. Bernstein, M., Shrager, J., Winograd, T.: Taskposé: Exploring Fluid Boundaries in an Associative Window Visualization. In: Proc. UIST 2008, pp. 231–234. ACM, New York (2008)
12. Shneiderman, B.: Direct Manipulation for Comprehensible, Predictable, and Controllable User Interfaces. In: IUI 1997, pp. 33–39. ACM Press, New York (1997)
13. Tashman, C.: WindowScape: A Task Oriented Window Manager. In: Proc. UIST 2006, pp. 77–80. ACM Press, New York (2006)
14. Kandogan, E., Shneiderman, B.: Elastic Windows: Improved Spatial Layout and Rapid Multiple Window Operations. In: Proc. AVI 1996, pp. 29–38. ACM Press, New York (1996)
15. Gaylin, K.: How are Windows Used? Some Notes on Creating an Empirically-Based Windowing Benchmark Task. In: Proc. CHI 1986, pp. 96–100. ACM Press, New York (1986)
16. Mackinlay, J., Royer, C.: Log-based Longitudinal Study Finds Window Thrashing, Palo Alto Research Center, UIS-2004-06 (2004)
17. Zipf, G.: Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology. Addison-Wesley, Reading (1949)
18. Juran, J.: Quality Control Handbook. McGraw-Hill, New York (1951)
19. Cockburn, A., Gutwin, C., Greenberg, S.: A Predictive Model of Menu Performance. In: Proc. CHI 2007, pp. 627–636. ACM Press, New York (2007)
20. Tauscher, L., Greenberg, S.: How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems. IJ. Hum. Comp. Stud., 47(1), 97–138 (1997)
21. Hansen, W.: User Engineering Principles for Interactive Systems. In: Barstow, D., Shrobe, H., Sandewall, E. (eds.) Interactive Programming Environments, pp. 288–299. McGraw-Hill, New York (1984)
22. Czerwinski, M., van Dantzich, M., Robertson, G., Hoffman, H.: The Contribution of Thumbnail Image, Mouse-Over Text and Spatial Location Memory to Web Page Retrieval in 3D. In: Proc. INTERACT 1999, pp. 163–170. Springer, Heidelberg (1999)
23. Bruls, M., Huizing, K., van Wijk, J.: Squarified Treemaps. In: Proc. Eurographics and IEEE TCVG Symposium on Visualization, pp. 33–42. IEEE Press, Los Alamitos (2000)
24. Tu, Y., Shen, H.: Visualizing Changes of Hierarchical Data Using Treemaps. IEEE Trans. on Visualization and Computer Graphics 13(6), 1286–1293 (2007)

The Panopticon and the Performance Arena: HCI Reaches within

Ann Light and Peter Wright

Sheffield Hallam University, Sheffield, S1 1WB
{a.light,p.c.wright}@shu.ac.uk

Abstract. The impact of new technologies is hard to predict. We suggest the value of theories of performativity in understanding dynamics around the convergence of biomedical and information technology. Drawing on the ideas of Butler and Foucault, we discuss a new, internal, context for HCI and raise potentially disturbing issues with monitoring health. We argue that by adopting explicitly social framings we can see beyond the idea of medical interventions to tools for wellbeing and recognize more of the implications of looking within.

Keywords: performativity, bodies, embodiment, biomedical, convergence.

1 Introduction

Research is being conducted on wellbeing and healthcare to provide new interfaces with the body (eg [1]), such as sensors and monitors. Although much work has been focused on the ethics and the practicalities of privacy in medical contexts, such as data protection for patient records [2], there has been less consideration of socio-political issues such as the effects of living wired up to the local medical centre. Rodden notably quipped during a keynote vision of the future [3] that he didn't want a fridge conspiring with his kids to stop him having a beer because his blood sugar was high. Awareness of health, with the attendant expectation that one will look after it, might seem an incontrovertible good, but carries with it the seeds of new normative behaviors with socio-political potentials. The complex relation between the human body (particularly its internal function) and such socio-political behaviors are difficult to articulate within our traditional conceptions of humans as technology users. In this paper, we engage theories of *performativity* [4] as a way of articulating these relations. And we open a space to discuss whether we can explore engagement with bodily practices fruitfully by seeing them as a performance for self and others.

To set the scene for this discussion, we identify the following developments:

- Networked/ubiquitous computing: the means of joining together information from different sources, transmitting it and using it differentially.
- Processing power for data mining: extracting specified information (such as heart rate, brain waves and gastric processing) from vast reserves.
- Nanotechnology: intervention on a scale that allows new sites of monitoring, particularly of the internal human state through implantation.
- Advances in Genetic Science including the human genome project: giving rise to new understandings of causal relations (like disease and genetic coding) and genetic difference across race and gender.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 201–204, 2009.

1.1 Performance and Identity

Performance has been used to understand interaction in many ways. There is the use of actual performances or methods drawn from drama (eg [5],[6]), but this does not concern us here. Another strand has used performance as a metaphor for people's engagement with technology ([7], [8]). HCI has drawn from Goffman [9] (eg [10]) for insights into how people construct their relations with others and more specifically how they present themselves. For instance, they may perceive themselves to be back- or front-stage in the company of others, and we can observe that most people feel the insides of their bodies should be backstage, ie a topic only for those they know and trust. Healey and Light make a crucial distinction between 'performing' as the designer of a tool intended and the potential appropriation of the tool for performing as social display [7]. Both they and [8] give the example of Dance Dance Revolution (DDR). DDR is not used as conceived because the shared aesthetics of dancing in public was overlooked in the design. Consequently, features are used in unanticipated ways for acquiring social standing. In reflecting upon the social engagement around the tool, [7] look beyond a common use of performance for exploring human-machine relations and implicitly consider the related area of performativity, or the creation of identity through repeated enactment [4]. "One is not simply a body, but, in some very key sense, one does one's body and, indeed, one does one's body differently from one's contemporaries" [4:272]. To date, most performance of identity is outside the skin[1]. Here we wish to go further, beyond the traditional arenas for 'doing one's body', to consider what happens when monitors and sensors turn identity inside out and give us access to 'do' one's body publicly from within. We invoke performance arenas - and the panopticon – to look at the effect of internal monitoring on identity.

2 Performing as a Healthy Being

One feature of Butler's analysis of how identity (for her, specifically gender) becomes created is the way that it is reinforced by social norms. This has a correlation in medical contexts. In treating conditions such as diabetes, the uses of automatic devices which monitor and inject insulin unquestionably provide support and safety, and may save lives. What happens, though, when such devices routinely report on the body's condition, either to the person or to professionals? Whose knowledge is it that a run of fish and chips has affected insulin levels? How does public health policy interact with people's choices about lifestyle? We must ask if the monitored person has the 'right' to live badly in the eyes of the wider populace and whether 'bad' behavior should affect access to treatment or the cost of life insurance. We could argue that, designed in or not, health gadgets persuade (take your tablets; eat less fat). Just as the wattson [11] makes one aware of use of fuel, so body monitoring can lead to societal pressures akin to those toward green living.

Foucault [12], in his analysis of power relations, suggests mere awareness that a controlling agent might be monitoring everyone at all times promotes individuals' compliance with the rules of the system. He conceives of this in a public world with

[1] Though let us note work, primarily in art, where internal modifications are manifested outwardly (eg [13]) and literature on the Cyborg (eg [14]).

centrally located viewing posts that, in turn, those monitored can see. By putting human-technology interfaces into the body, we are led to consider how previously hidden aspects of a person are made available and to whom. By likening the introduction of observation technologies to this panopticon, we highlight both the new arena for inspection these tools introduce and the performative attributes of internal readings, thereby raising the issue of conformity. In taking this perspective, questions arise which reach beyond intervention, safety and privacy, to consider understandings of observation and agency, such as:

- Who is the audience for an internal view of what my body does?
- How do I want to portray myself and what scope exists to choose a portrayal?
- Am I empowered to resist others' readings of my condition?
- What is the impact of my new visibility upon my sense of self and wellbeing?

2.1 Performing in Networks Where Machines Watch over Us

So who is looking into our bodies? In the networked world, the performance of health may be interpreted by society, but its first audience is machines with capacity to make connections between streams of data, note intersections and store patterns. In other words, this audience is forming judgments about the nature of the individual. But, whereas when I dance, how I dance and what I wear are all familiar choices that I make in using DDR, I have less awareness about or creativity in how I 'do' my body. We can reflect that, as audience, a machine is impersonal and apparently apolitical in the face of my body's performance: a machine does not process identity based on appearance and other qualitative aspects, but finds rationale in metrics. However, ostensibly neutral quantitative approaches operate with embedded political readings of information and how it should be manipulated. For instance, statistical genetic information will be more interesting to a machine than skin color. But the use of statistical genetic data relating to race will, in itself, be politically charged, as will any embedded interpretative formulae used in analytic systems. What kind of body would a person want to 'do' in this context? It is not easy to predict, even if a warning about sickle cell anaemia would be useful, but asking brings us closer to an answer. What sort of audience do we want to construct? Reflective? Empathetic? One sharing embedded co-experiences? Or is that to mislead people with a 'kind' front-end?

3 Discussion

We have presented a developing trend in mediation and a way of framing it so that the implications can be seen in social terms. By calling on an (inter)active reading of the construction of identity and its relation to changes in biomedical practices, we offer a demonstration of how medical interfaces can affect public aesthetics. Just as DDR was not conceived to be about performing socially, neither are the tools here and yet we see that they offer this opportunity. However, whereas the dancers using DDR can choose a public image to 'do' even if they cannot choose their spectators, the panopticon is imposed from outside into a space where we are genetically programmed and have less recourse to (re)inventing ourselves. It is an arena in which we are not used to appearing and have little understanding of presentation.

We offer a distinction, and it is subtle, between wellbeing tools and health tools. In designing health interventions, is our highest purpose to support individual wellbeing? (Instead, it may be to keep welfare costs down.) People's greatest good may not be served by the pressure to conform in yet another arena: anxiety about obesity may not be its best cure. So we suggest the device of exploring implanted and embodied health technology through performativity to encourage new discourses around presentation and control. Such measures will be critical if we are not thoughtlessly to embed new orthodoxies of behavior in the code of our creations. We offer this analysis as providing an opportunity to design tools - and connections between them - that allow people to perform themselves inside and out with discretion, decorum and control.

References

1. Leonhardt, S.: Personal Healthcare Devices. In: Mukherjee, S., Aarts, R.R., Widdershoven, F., Ouwerkerk, M. (eds.) amIware Hardware Techbology Drivers of Ambient Intelligence, part. 6, Springer, Heidelberg (2006)
2. Brown, I., Adams, A.A.: The ethical challenges of ubiquitous healthcare. Int. Review of Information Ethics 8, 53–60 (2007)
3. Rodden, T.: Ubiquitous computing in the real world. In: HCI 2006 Keynote address (2006)
4. Butler, J.: Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory. In: Case, S. (ed.) Performing Feminisms: Feminist Critical Theory and Theatre. Johns Hopkins UP, Baltimore (1990)
5. Newell, A.F., Morgan, M.E., Gregor, P., Carmichael, A.: Theatre as an intermediary between users and CHI designers. In: Proc. CHI 2006 (2006)
6. Light, A., Weaver, L., Healey, P.G., Simpson, G.: Adventures in the Not Quite Yet: using performance techniques to raise design awareness about digital networks. In: Proc. DRS (2008)
7. Healey, P.G., Light, A.: When Scoring doesn't Matter: The aesthetics of performance in arcade games. Co Design suppl. 1 (2007)
8. Dalsgaard, P., Hansen, L.K.: Performing Perception—Staging Aesthetics of Interaction. ToCHI 15(3), 1–33 (2008)
9. Goffman, E.: The Presentation of Self in Everyday Life. Penguin Books (1959, 1990)
10. Hulme, M., Truch, A.: The role of interspace in sustaining identity. In: Glotz, P., Bertschi, S., Locke, C. (eds.) Thumb Culture: The Meaning of Mobile Phones for Society, pp. 137–148. Transcript Verlag, Bielefeld (2005)
11. Wattson, http://www.diykyoto.com/uk
12. Foucault, M.: Discipline and Punish: the Birth of the Prison. Random House, New York (1975)
13. Stelarc, http://www.stelarc.va.com.au/
14. Haraway, D.A.: Cyborg Manifesto. In: Simians, Cyborgs and Women: The Reinvention of Nature, pp. 149–181. Routledge, New York (1991)

Exploring the Use of Discrete Gestures for Authentication

Ming Ki Chong and Gary Marsden

Department of Computer Science, University of Cape Town,
Rondebosch, Cape Town, South Africa
{mchong,gaz}@cs.uct.ac.za

Abstract. Research in user authentication has been a growing field in HCI. Previous studies have shown that peoples' graphical memory can be used to increase password memorability. On the other hand, with the increasing number of devices with built-in motion sensors, kinesthetic memory (or muscle memory) can also be exploited for authentication. This paper presents a novel knowledge-based authentication scheme, called gesture password, which uses discrete gestures as password elements. The research presents a study of multiple password retention using PINs and gesture passwords. The study reports that although participants could use kinesthetic memory to remember gesture passwords, retention of PINs is far superior to retention of gesture passwords.

Keywords: User authentication, gesture passwords, discrete gestures.

1 Introduction

User authentication is a fundamental requirement for most remote access services. Most often, a user is required to provide a knowledge-based password as evidence of the user's identity. Traditionally, alphanumerical passwords or PINs (Personal Identification Numbers) are used for user authentication. However, the use of a text-based password requires a trade-off between security and memorability. This trade-off arises from the limitation of human memory and, as a result, 'secure' passwords (long and random sequences of characters) are easily forgotten. To avoid the risk of forgetting passwords, users often adopt insecure behaviours such as writing down their passwords or disclosing their passwords to perceived trusted parties [1].

Usable security is a growing field of research in HCI. Concepts and ideas from earlier research have suggested that alternative authentication schemes, other than the knowledge-based schemes, are possible. Schemes such as token-based and/or biometric-based solutions were introduced as alternatives to passwords. Since these schemes do not require the users to memorize passwords, they have an advantage over knowledge-based systems by reducing the load on users' memory. However, these solutions are introduced at the expense of accessibility and cost of hardware [2], and, as a result, knowledge-based verification remains the preferred form of user authentication.

In the search to improve password usability, the use of graphical passwords has emerged as an alternative solution. Graphical passwords exploit the *picture superiority effect* [3]. Cognitive studies have shown that people are much better at recognizing

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 205–213, 2009.

previously seen images than at recalling text precisely [4]. Graphical images provide a rich and detailed representation in memory, which also make the images distinctive at time of retrieval [3]. And hence graphical passwords can be particularly useful in addressing the weakness of memorability in text-based passwords. This is confirmed by findings from earlier research which have shown that graphical images can increase password memorability through users' visual memory [5]. Moreover, a previous study by Moncur and Leplâtre [6] examined subjects' retention of multiple PINs compare to multiple graphical passwords. Their results demonstrate that multiple graphical passwords are substantially more memorable.

Although previous studies on graphical passwords have shown that using visual aids can help users to encode passwords into their long-term memory, there are other retention approaches that can be exploited to increase password memorability, and kinesthetic memory (or muscle memory) is a one of those approaches. Instead of using text-characters or images, a password can be made up of multiple gestures: through practice and repetition, the password movements can gradually consolidate into the user's memory. However, so far no research was done to investigate how kinesthetic memory can assist users in remembering passwords.

This paper addresses a new authentication approach, called gesture password, which exploits kinesthetic memory for password retention. It aims to answer the following question:

Can people remember gesture passwords successfully?

In addition, we are particularly interested in exploring the use of gesture passwords for mobile authentications such as mobile banking; therefore, mobile phones are selected as the input devices in this research.

In the rest of this paper, we present a new gesture password design that uses arm movements as a set of password elements. We report an empirical study of gesture passwords and a discussion of the results before conclusions are drawn and future directions for research are identified.

2 Design

In recent years, accelerometers have been increasingly integrated into mobile phones (such as Apple's iPhone, Nokia N95, etc.). A built-in accelerometer allows the mobile device to sense users' movements and, as a result, it provides a new modality for user input. At the moment (2008), there are a small number of mobile applications which make use of this: for example, Williamson et al. introduced Shoogle, an interface for sensing data within a mobile device as the device is shaken [7]. As more uses of accelerometers are being discovered, we confidently predict that more mobile phones will be equipped with built-in accelerometers in the near future.

In the design of this research, an accelerometer is used to detect directional movements as gesture inputs for user authentication for mobile devices.

2.1 Related Work

Although there is no research that investigates the use of kinesthetic memory to increase password memorability, there is some work which exploits the use of gesture movements for authentication. In an attempt to create PIN-less authentication environments, researchers have designed methods that use gestures for pairing devices. In [8], Patel et al. suggested an authentication scheme by shaking a device. Their authentication scheme is based on a user authenticating his/her device when using a public terminal by mimicking a sequence of gestures that is generated by the device and displayed on the terminal. Similar work by Mayrhofer and Gellersen suggests a method which uses accelerometer data to pair two mobile devices [9]. Their authentication method requires a user to hold the pairing devices tightly together and shake the devices for a short period. The built-in accelerometers within the devices are used for measuring the movement as the devices are shaken; the devices will only be paired if the accelerometers readings are similar.

Research by De Luca et al. discovered that many people memorize their PINs by remembering the resulting shape of the spatial relations on the number pad [10]. From their findings, Weiss and De Luca introduced a concept, called PassShapes, of replacing PIN numbers with directional strokes in a two-dimensional plane. The strokes of a password are made up of many horizontal, vertical, and diagonal strokes. The strokes can form specific shapes, and the shapes are suggested as password mnemonics [11].

2.2 Gesture Password Elements

The gestures used for this study are discrete gestures. A discrete gesture can be distinguished as a movement from a starting position to a stopping position. Here, a discrete gesture is defined as *a distinctive singular movement that can be perceived individually and not connected to, or part of another motion*. In other words, a motion that cannot be further decomposed as units of actions can be classified as a discrete gesture. Since it has the property of a discrete structure, multiple singular gesture units can be combined to form a string of discrete gestures.

In this study, we apply a concept similar to PassShapes. Gesture strokes in a three-dimensional space are used as passwords. Ten discrete gestures (illustrated in Fig. 1) are defined as password elements. The elements were designed based on the spatial orientation of a mobile phone, and they were also designed with the intention that each gesture must have a symmetrical gesture in the mirror direction. The *forward* gesture (Fig. 1.a), for example, has a symmetrical gesture element *backward* (Fig. 1.b) in the mirror direction. The reason for the design with symmetry is to ease the process of learning the gestures. As novice users learn a gesture element they can apply the reverse movement to learn the mirror gesture, thus simplifying the learning process for the users.

Due to the articulation structure of a human arm, motion is limited in certain ways. This implies that some strings of gestures could be impossible for people to reproduce. For example, when a user holds a device perpendicular to the ground, the maximum tilting angle the user can rotate the device is about 180 degrees. Beyond that point it is uncomfortable or impossible to tilt further. Thus a string of tilt left or tilt right is not replicable by a user (Fig. 2 illustrates an example of a string of tilt left gestures. Tilting left beyond fig. 2.c is impossible). This problem manifested during our initial design stage, so a decision had to be taken that after each element entry, the device must be moved back to its starting position; hence the next element entry must start from the initial position. As a result, a valid password element entry is a string of paired discrete gestures, where a given pair is made up of a gesture element and its reverse.

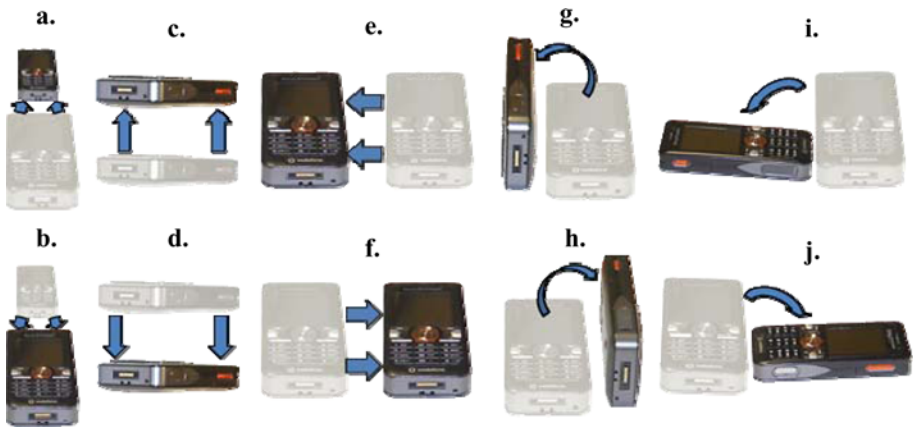


Fig. 1. Gesture password elements. (a) *Forward*, (b) *Backward*, (c) *Up*, (d) *Down*, (e) *Left*, (f) *Right*, (g) *Tilt Left*, (h) *Tilt Right*, (i) *Swing Left*, (j) *Swing Right*

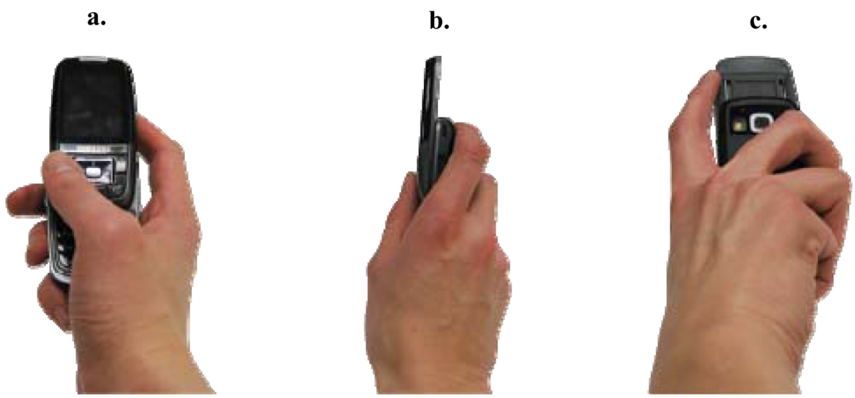


Fig. 2. A string of tilt left gestures before adjustment. (a) *Initial position*, (b) *Tilt left from position (a)*, (c) *Tilt left from position (b).*

2.3 Security

A *string* is an ordered list of elements in which the same elements can appear multiple times at different positions. In this study, a *gesture password* is defined as a string of multiple gesture elements. For a user to authenticate, that user is required to produce a string of gesture password elements in the correct order. Since our system supports ten different gestures, it has the same password space as standard PINs, thus both systems have the same statistical guessability.

One of the drawbacks of gesture passwords compares to PINs is that the system is susceptible to shoulder surfing attacks. Due to the nature of using movements as inputs, an attacker can observe and record a user's gesture password entry easily. We therefore recommend gesture passwords are only to be used in private secure areas.

2.4 User Interaction

A gesture entry is entered as a movement that starts and stops in the same position. A complete gesture is defined as the change of state from a motionless state to a moving state and then a motionless state to identify the end of the gesture. Consequently, for the system to register a motionless state, the user needs to pause between each gesture elements during password entry.

3 Study

Human memory has a limited capacity to remember the arbitrary text and number strings that make up a password. This results in password retention deficiency and to overcome this deficiency, people select passwords to which they can attach meaning. Although memory cues help users to memorize their passwords, the passwords can also become more guessable and raise security vulnerabilities. To avoid this weakness, many security systems disallow the use of personalized passwords, generating and enforcing random passwords for their users.

In this study, we therefore restrict ourselves to the issue of retention of multiple system generated passwords. As previously mentioned, a study by Moncur and Leplâtre [6] compared subjects' retention of multiple PINs compared to multiple graphical passwords. Following on from their study, we investigate users' retention of multiple passwords of PINs and gesture passwords. We therefore conducted a one week longitudinal study to measure passwords retention. Furthermore, we are particularly interested in the memory strategies of users of the evaluating authentication systems.

3.1 Passwords Retention Test

In this study, we tested for users' retention of multiple passwords. 12 learners from a skill training centre in Khayelitsha, Cape Town, were recruited as participants. The subjects participated in the study during their class hours, so each subject was paid R60[1] at the end of the study for remuneration. We recruited adult subjects; their ages range between 17 to 39 years old and with a mean age of 24.9.

One of the aims of this study is to learn the strategies that the subjects adopt to memorize passwords of each testing system. To avoid any learning effects (such that the subjects applying a strategy learned from one system to another), the between-groups experiment approach was adopted. Therefore, six subjects were assigned to a group using PINs (Group 1) and the other six were assigned to a group using gesture passwords (Group 2).

To simplify the experiment process, the experiment was conducted without using digital prototypes. Instead, the *wizard of Oz* prototyping technique [12] was adopted. All participants first undertook a training session to familiarize themselves with the allocated password system; each subject was trained by the experiment facilitator. For the gesture password group, the movements of the ten gestures elements were illustrated to the subject by the facilitator. To simulate inputting gesture passwords using a mobile device, a block of a rectangular object (we used a mobile phone) was given to

[1] The symbol "R" represents South African currency "Rand".

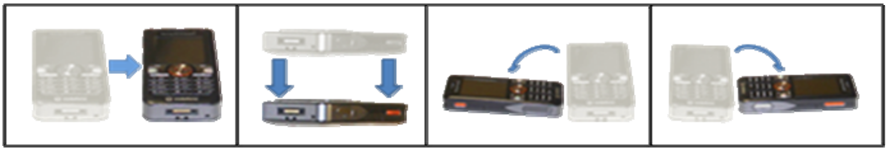


Fig. 3. An example of a random gesture password. The password is read from left to right (i.e. the left most gesture is the first element).

the subject to represent a sensor device. After demonstrations, the subject was requested to reproduce the ten gestures; this is to ensure the subject understands the system and knows how to enter gesture passwords.

Once the subject was familiar with the system, he/she was given three pre-generated random passwords (see fig. 3 for an example of a gesture password); each password was made up of four elements.

In [6], their participants undertook a rehearsal session by entering each of their assigned passwords correctly twice. However, we believe that entering the passwords twice is not sufficient for the subjects to register the gesture passwords into their kinesthetic memory. Instead, our subjects were given a 24 hour rehearsal period to memorize the passwords. To enable the subjects to have access to the passwords during the rehearsal period, the passwords were given to the subjects on paper.

After the rehearsal period expired, a facilitator returned to the subjects and collected the papers that contain the passwords. Six days later the facilitator returned again and requested the subjects to recall their given passwords. This was followed by a questionnaire session to determine the methods the subjects applied to remember the passwords.

Although 12 subjects were initially recruited for this study, one subject from group 1 had dropped out. Therefore, a total of 15 PINs and 18 gesture passwords were tested (the results are listed in Table 1). Given the participants with the defined periods, the results show that retention in group 1 was superior to retention in group 2. The result of group 1 could be influenced by the subjects' familiarity with PINs. Hence, the subjects have already adopted an effective strategy to remember PINs beforehand. The low scores archived by group 2 are suspected to have been caused by the subjects not having enough time to practise their passwords. One of the subjects had explicitly mentioned to us that she needed more time to memorize her given passwords.

Table 1. Results of the retention test

PINs		Gesture Passwords	
Subject ID	Score (Out of 3)	Subject ID	Score (Out of 3)
PIN1	3	GES1	0
PIN2	3	GES2	1
PIN3	3	GES3	1
PIN4	3	GES4	0
PIN5	3	GES5	0
PIN6	*dropped out*	GES6	1

3.2 Retention Strategies

Subsequent to the experiment, a questionnaire on users' retention strategies was conducted. Three subjects from group 1 indicated that they remembered PINs through visualizing the images of the numbers. One subject said he adopted a strategy by grouping the PINs' digits into groups of two numbers. He memorizes a PIN by remembering the first two digits and then the next two digits. Two subjects from group 1 adopted a strategy of rehearsing the numbers audibly in their mind. None of the subjects adopted the strategy of constructing a story using the given PINs, and neither did they adopt the strategy of memorising the PINs through the mnemonic of the numbers' position on a keypad. The former can be explained by the difficulty of constructing a story using only numbers, and the latter is because the PINs were given on paper, thus the subjects never entered the PINs on a number pad. The questionnaire with group 2 reported that all six participants have practised their gesture passwords by moving their hands in the passwords' directions. Five subjects reported to have adopted the strategy of rehearsing the gesture passwords audibly by speaking the directions, and three subjects memorized the passwords by visualising the directions, like the arrows of the gestures. The results show that participants attempted to use their kinesthetic memory to memorise gesture passwords through practice, in addition, some participants also opted to use their audio and graphical memory for gesture passwords. In other words, some people use their alternative memory to increase the memorability of gesture passwords. This could be because visual and audio memory complements muscle memory for gesture passwords; however, further study is required to confirm this.

4 Conclusions and Future Work

In conclusion, we suggested an alternative password authentication system for mobile devices that exploits users' kinesthetic memory to recall passwords. We defined a new password scheme which uses discrete movements as password elements, and we investigated the memorability and users' retention strategies of gesture passwords. The results from the empirical study suggest that users are more effective in remembering PINs than gesture passwords.

Currently, we can argue that computers and mobile phones are the most common devices that require authentication, but we should not limit ourselves in researching suitable authentication solutions for those devices only. As ubiquitous computing is becoming more popular, many new types of devices also require authentication. Although negative results of gesture passwords were shown in this research, it should not limit the adoption of gesture passwords. Our gesture password system is ideal for devices that are equipped with a built-in accelerometer and limited user input capabilities. Using Apple's fourth generation iPod Nano [13] for an example, instead of using its click-wheel to enter PINs, the system can adopt gesture passwords for authentication. One can argue that graphical authentication is a good option for iPods. However, much of the details of the graphical display may be lost due to the small resolution screen of the device; therefore, some pictures of a graphical authenticator may not be displayed properly.

In this project, the gesture elements were designed for mobile phones. This decision has restricted the gestures to be large and slow arm movements. Smaller sensors may be more appropriate to detect movements. In such a case, faster and more discreet gestures could be adopted as password elements and more flexible muscles could be used; for example, detecting finger movements as gestures.

One aspect of security not covered in Moncur and Leplâtre's work [6], nor any other HCI study we could find, was an investigation into the users' perception of trustworthiness of the systems being evaluated. Currently, password authentication is the most commonly used verification scheme and users have adapted to use passwords for authentication. Although alternatives, such as graphical passwords, have been proven to be more usable, it is arguable that users may prefer to use text-based passwords from the standpoint of familiarity. Therefore, investigation of perceived trust and preferences between password systems is essential.

Acknowledgements

We would like to thank Learn to Earn, Khayelitsha Branch, for their support during the study of this project. We would also like to acknowledge the National Research Foundation (NRF) for funding this research. Thanks also to Rob Mori of Sun Microsystems for dontaing the SunSpot equipment used to capture the gestures. This work was also supported by the Telkom/NSN Centre of Excellence in Broadband Networks and their Applications.

References

1. Adams, A., Sasse, M.: Users are not the enemy. Communications of the ACM 42(12), 41–46 (1999)
2. Renaud, K.: Evaluating Authentication Mechanisms. In: Cranor, L., Garfinkel, S. (eds.) Security and Usability, pp. 103–128. O'Reilly Media, Inc., Sebastopol (2005)
3. Nelson, D.L., Reed, U.S., Walling, J.R.: Picture Superiority Effect. Journal of Experimental Psychology: Human Learning & Memory 2, 523–528 (1976)
4. Paivio, A., Csapo, K.: Picture superiority in free recall: Imagery or dual coding? Cognitive Psychology 5(2), 176–206 (1973)
5. Dhamija, R., Perrig, A.: Déjà vu: A user study using images for authentication. In: Proceedings of the 9th USENIX Security Symposium (2000)
6. Moncur, W., Leplâtre, G.: Pictures at the ATM: Exploring the usability of multiple graphical passwords. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 887–894. ACM Press, New York (2007)
7. Williamson, J., Murray-Smith, R., Hughes, S.: Shoogle: Excitatory multimodal interaction on mobile devices. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 121–124. ACM Press, New York (2007)
8. Patel, S., Pierce, J., Abowd, G.: A gesture-based authentication scheme for untrusted public Terminals. In: Proceedings of the 17th annual ACM symposium on User interface software and technology, pp. 157–160. ACM Press, New York (2004)

9. Mayrhofer, R., Gellersen, H.: Shake well before use: Authentication based on accelerometer data. In: Proceedings of the 5th International Conference on Pervasive Computing, pp. 144–161. Springer, London (2007)
10. De Luca, A., Weiss, R., Hussmann, H.: PassShape: stroke based shape passwords. In: Proceedings of the 19th Australasian conference on Computer-Human Interaction, pp. 239–240. ACM Press, New York (2007)
11. Weiss, R., De Luca, A.: PassShapes: utilizing stroke based authentication to increase password memorability. In: Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges, pp. 383–392. ACM Press, New York (2008)
12. Sharp, H., Rogers, Y., Preece, J.: Interaction Design: Beyond Human-Computer Interaction, 2nd edn. John Wiley & Sons, Chichester (2007)
13. Apple iPod Nano, http://www.apple.com/ipodnano/

AirMouse: Finger Gesture for 2D and 3D Interaction

Michael Ortega[1] and Laurence Nigay[2]

[1] PRIMA - INRIA Rhône-Alpes
michael.ortega@inria.fr
[2] IIHM - University of Grenoble, CNRS, LIG
laurence.nigay@imag.fr

Abstract. This paper presents AirMouse, a new interaction technique based on finger gestures above the laptop's keyboard. At a reasonably low cost, the technique can replace the traditional methods for pointing in two or three dimensions. Moreover, the device-switching time is reduced and no additional surface than the one for the laptop is needed. In a 2D pointing evaluation, a vision-based implementation of the technique is compared with commonly used devices. The same implementation is also compared with the two most commonly used 3D pointing devices. The two user experiments show the benefits of the polyvalent technique: it is easy to learn, intuitive and efficient by providing good performance. In particular, our conducted experiment shows that performance with AirMouse is promising in comparison with a touchpad and with dedicated 3D pointing devices. It shows that AirMouse offers better performance as compared to FlowMouse, a previous solution using fingers above the keyboard.

Keywords: AirMouse, interaction, 2D/3D pointing, computer vision, Fitts' law.

1 Introduction

Interaction devices such as a mouse require an additional surface to operate on. Laptops are widely used today, and additional space is an issue in a context where space is at a premium as is the case for example of the table in a train. More generally, the cumbersome problem is one of the «grand challenge» questions described by Bowman et al. [2]; they argue that being cumbersome "has a huge impact on the usability of the systems". Touchpads and key-joysticks are solutions for this problem, however the pointing performances are low and they are not efficient for 3D interaction.

For 3D interaction, the current existing devices are bulky and expensive. Previous works, like [22, 17, 28], proposed solutions using a finger above the keyboard. However, the solutions imply arm tiredness and cognitive load due to transformation between the manipulation space and the display space. Moreover, quoting the authors of Flowmouse [28]: "pointing performance with FlowMouse was significantly worse than with a trackpad".

Facing these issues, we present AirMouse for 2D and 3D interaction that is based on finger gestures performed above the keyboard. AirMouse is a mix of efficient pointing techniques (namely Ray-casting and Virtual-Hand techniques [2]) adapted for 2D and 3D pointing. The key features of AirMouse include:

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 214–227, 2009.

No additional surface: The additional surface, beside the laptop, is suppressed.

2D/3D interaction: The technique supports mixed 2D and 3D interaction in the same application. This feature is important since 2D and 3D views are more and more common and their combinations valuable as explained in [26]. With AirMouse no additional time is needed for the user to move her/his hand from one input 2D device to another 3D one.

Reduced homing time: The switching time between the keyboard and the pointing device is drastically reduced. This feature is particularly important for limited-mobility users.

Easy to learn and easy to use: Based on a natural and direct way of pointing and manipulating, the concept is easily adopted by new users.

Reasonably low cost: The technique can be implemented at reasonably low cost.

Low tiredness: Compared to similar existing methods [22,28,17], AirMouse does not force the user to move his forearm. The hand palm can be left beside the keyboard, and only forefinger movements are needed. Other fingers can stay on the keyboard.

Low cognitive load: No rotation between the manipulation space and the display space decreases the cognitive load implied by previous methods [14].

Good performance for 2D and 3D pointing: The performance of the AirMouse is better than the FlowMouse technique and is promising in comparison with a touchpad and with dedicated 3D pointing devices.

This paper presents a low-cost vision-based implementation of the AirMouse technique in order to validate the technique itself. A final product could be industrially implemented on all the laptops with an unobtrusive and no bulky system of cameras integrated in the corners of the laptop screens. Indeed our technique on a laptop is promising for 3D games or professional 3D applications (architecture, interior design applications used by sellers at clients home) and may foster even more 3D applications on laptops.

The paper is organized as follows: we first discuss related work before explaining the AirMouse technique and its vision-based implementation. We then present a formal evaluation and its results of the two implemented pointing techniques (2D and 3D) that we compare with more traditional 2D and 3D pointing devices.

2 Related Work

On the one hand, the mouse is the most commonly used device for desktop applications, for experts and occasional users. The device is easy-to-learn, low cost, and, compared to other desktop pointing devices like trackballs, touchpads or key-joysticks, it offers the best performance for time completion of pointing tasks [16, 7]. On the other hand, it is natural to designate an object using the forefinger which is routinely performed in everyday life. Towards naturalness and intuitiveness of the interaction, several studies therefore focused on using a finger for showing, selecting and manipulating objects displayed on screen instead of the mouse [24].

With multi-touch interactive surfaces as in [10] the mouse is replaced by the forefinger. Multi-touch interactive surfaces define a very dynamic research area. When focusing on large surfaces such as a table, such setting cannot be used in everyday work environment. Moreover, although interaction with multi-touch interactive surfaces is natural and intuitive, there are also identified limitations. One of the main issues is the tiring effect of lifting and moving the arm between the different points of the surface. Moreover such movements are not possible for limited-mobility users.

In our work and as opposed to multi-touch interactive surfaces, we focus on fingerbased interaction for a laptop setting including a traditional keyboard. Previous studies have been conducted on using fingers while keeping the keyboard for interaction. First, in [3], the mouse is replaced by a joystick placed in the middle of the keyboard. The keyboard and joystick combination reduces the homing time in comparison with the keyboard and mouse combination. However, the performance and the usability of the joystick are far from the mouse capacities [7]. A different approach is presented in [19] and [22] with the FingerMouse: this freehand pointing technique is based on computer vision for tracking a fingertip. The screen cursor moves according to the user gestures in a horizontal plane just above the keyboard. The selection, equivalent to a mouse click, is performed by pressing the SHIFT key. More recently, another computer vision-based pointing gesture technique, namely the FlowMouse, has been proposed in [29]. FlowMouse uses one camera, and detects the complete hand 2D horizontal movements (using optical flow cues) above the keyboard. Tha hand translates and then force the user to move its arm. The technique has been experimentally evaluated: a Fitts' law study demonstrated that while pointing performance was worse than a touchpad, the interaction was intuitive, easy to learn and easy to use. Finally the Visual Touchpad [17] is a mixed technique that combines the "Finger-Mouse" technique, 2D hand gesture techniques and a virtual keyboard. Two cameras are used to detect if the fingertip is on or above the virtual keyboard. The system could be considered as a low-cost tabletop display or touch-screen, but with a dissociation of the horizontal tracked surface (i.e., a quadrangle surface replacing the keyboard) from the vertical computer screen.

The aforementioned techniques do not require additional space to operate on, which is an important issue for laptops used in various contexts. They also reduce the homing time which is responsible for 42% of the time required to move the hand from the keyboard to the mouse, point, and go back to the keyboard [4]. However these techniques have two main limitations. First, the forearm has to move above the keyboard which is tiring. Letting the palm beside the keyboard, and only moving the fingers should be less tiring. Secondly, the transformation, here a rotation, between the plane of the finger gesture and the one of the cursor movements increases the cognitive load, which decreases performances and increases tiredness of the user [14].

These techniques aim at replacing the mouse. These techniques as well as the mouse are not adapted for 3D pointing or manipulation, due to their lack of dimensions. Additional modifier keys are then required for 3D interaction. Specific devices exist for manipulating objects in three dimensions, like PHANTOMs [18] and the spacemouse [5]. Most of them are expensive, bulky and also involve switching time when changing from using the 3D pointing device to the keyboard. Moreover, the 3D pointing device being next to the keyboard, the action workspace defined by the position of the device is deported from the screen that defines the virtual workspace. A large translation between the action and virtual workspaces decreases the interaction performances [20].

The AirMouse technique, for which an implementation is presented in the next section, extends the FingerMouse and FlowMouse possibilities by considering 3D finger gestures. AirMouse therefore supports both 2D and 3D pointing.

3 Technique Overview and Implementation

The AirMouse technique consists of using fingers over the keyboard for interacting in two and/or three dimensions with desktop applications. The first goal is to decrease the tiredness implied by previously presented techniques. One constraint is then to allow users only to move the fingers, letting the palm beside of the keyboard, as shown in figure 1. The second goal is to reduce the cognitive load implied by a transformation between the manipulation space and the working space. By only moving the fingers, it is not possible to perfectly fit the two spaces, and a scale as well as a translation are required. However, the rotation, which is the transformation with a strongest impact on cognitive load, is suppressed.

These two aforementioned goals implies that we initially assumed that it was not useful for the users to visually identify the 3D interaction volume, i.e. the volume in which gestures are possible since fingers are tracked by the system. An informal evaluation showed that our assumption was right and that it is not important for users to know the limits of the tracking area. Indeed, users do not look at their fingers but only at the screen (same as when using a mouse). They visually understand the interaction workspace, seeing the limits on the screen. A scale transformation between the tracking area and the displayed area is therefore possible. This allows us to reduce the gesture amplitude and therefore the tiring effect while preserving enough precision for pointing tasks. The tracking area has been defined considering the medium size of a hand: for right handed users, the defined area horizontally corresponds to the right half part of the keyboard, and vertically corresponds to the lower half part of the screen (see Figure 2). However, a calibration step allows to readjust this area for users with very small or very big hands.

The main technological issue is first to track a part or the whole fingers in 3 dimensions. The vision-based implementation proposed in this paper is based on vision reconstruction algorithms [6], and is able to track many points in 3D. Two trackIR [12] camera devices are placed on top of the laptop screen (see Figure 2). Each trackIR device is composed of one infra-red camera that is circled by infrared LEDS.

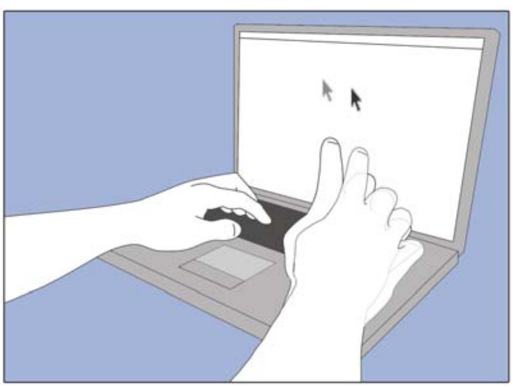


Fig. 1. AirMouse is an interaction technique, which consists of moving fingers in 3D above the keyboard for directly pointing and manipulating objects on screen

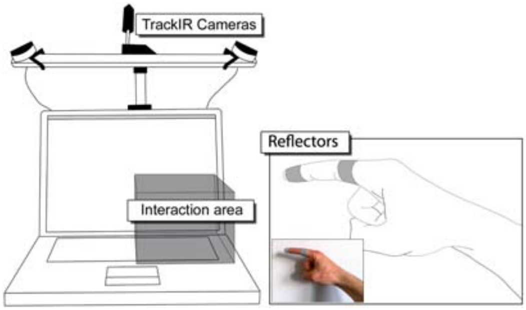


Fig. 2. One set-up for the AirMouse technique. Two trackIR devices are placed on top of the laptop in order to provide a large 3D interaction space.

Thus, a reflector placed in front of the device allows us to reflect the infrared light from the LEDS back to the camera. Using the trackIR SDK, we implemented an algorithm for reconstructing the three dimensional position of the reflector area. Many reflectors can be used.

As shown in Figure 2, the prototype, that has been done for validating the technique, includes two TrackIR cameras far from the top of the screen. For sure this prototype is not usable in all the usage contexts of a laptop. However, we focus here on validating the AirMouse technique. Using cameras with large enough focal distance will enable us to fix the cameras at the two top corners of the screen. Another future solution would be to use a single camera providing depth cues [13].

In this implementation of the AirMouse technique, we only focus on 2D and 3D pointing. In order to provide an intuitive and natural way of pointing, we decided to use isotonic interaction instead of isometric interaction.

2D Pointing. The forefinger is commonly used for designating far or proximal objects. For this action, we can consider the finger as defining an infinite ray which intersects with the designated object. This technique [1] is commonly called Ray-Casting [2]. It is a natural and conceptually simple [27] pointing technique. The Ray-Casting technique sounds adapted for AirMouse: while using the keyboard, the user moves her/his forefinger and the cursor will be displayed on the designated position on screen. For this technique, two reflector ring are used (see Figure 2): one on the forefinger tip, and another one on the forefinger third phalanx. Thus, the two recorded 3D points allow us to define a line whose intersection with the screen plan gives the 2D position of the cursor.

3D pointing. Only the forefinger tip reflector is used here. The tracking by the two cameras, combined with a classical reconstruction algorithm [6], gives one three-dimensional point. The 3D cursor then moves according to the three-dimensional position of the fingertip. In order to preserve the directness of the interaction, no rotation transformation is applied between the tracking area and the displayed area. There is a direct mapping: left/right and up/down movements of the fingertip are directly mapped to cursor movements in the same direction. Direct mapping is also provided for the depth: while the fingertip is going away from the screen, the cursor moves ”closer” to the user. Nevertheless as explained above, we apply a scale transformation between the tracking area and the displayed area for reducing gesture amplitudes and therefore tiredness while maintaining enough precision for pointing tasks (see Figure 2 Interaction area).

Selection. The selection, equivalent to a mouse click, is performed by clicking the touchpad button of the laptop.

The goal of this implemented technique of AirMouse is to experimentally study the 2D and 3D pointing tasks. Nevertheless we point out that AirMouse intends to replace the mouse by providing a large range of possible interaction techniques that need to be further studied. For example for this implementation, the clutching aspects are not examined. We nevertheless show in the last section of this paper an implementation of AirMouse for an existing 3D modeler that supports smoothly integrated 2D and 3D pointing tasks with no need for activation/deactivation.

4 2D Evaluation

In this controlled experiment, we evaluated the performance of the above implemented technique of AirMouse as a pointing device, using 2D and 3D Fitts' law studies. The two tasks have been performed by 15 subjects with no prior experience with 3D interaction devices. They were right handed, all rated themselves as advanced computer users and had normal or corrected normal vision. The two tasks are based on the recommendation given by Soukoreff et al. in [25], using the Fitts' law [8]. Finally, at the end of each experiment, we asked participants to freely comment on the techniques and then to rank-order each of the experimented pointing devices respectively in terms of performance, satisfaction and tiring effect. As pointed out in [23] subjective satisfaction may be the key determinant of success.

4.1 Pointing Task

The goal of this evaluation is to position the 2D pointing performance of AirMouse in relation with the performance of other traditional device. We therefore compared

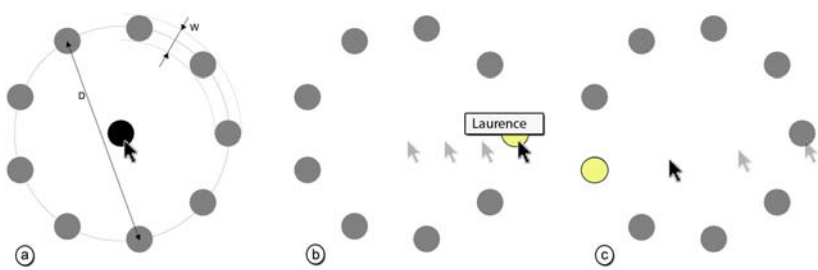


Fig. 3. Combination of multidirectional tapping with click-and-write. a: Initial cursor position, with amplitude and width visual representation. b: First target is reached (and clicked), the subject wrote her first name. c: 'RETURN' has been pressed, the cursor is going to the next target.

Air-Mouse with the three well-known and commonly used pointing devices: the traditional mouse, the touchpad and the key-joystick. Traditional mouse is an isotonic device and every advanced computer user can be considered as an expert, i.e. the time performance of this device are partly due to the advanced knowledge of the users. The touchpad is also an isotonic device with a limited interaction space. The efficiency of use of such a device can be optimized by improving the scaling factor: however, the limit is fixed by the corresponding obtained precision quality. In the experiment, the scaling factor has been chosen empirically, computing the mean of three users parameters. The same value has been kept for all subjects. The key-joystick is an isometric device, i.e., it controls the cursor by speed. It can be found on a large variety of laptop, but it is not often used. None of the subjects has regularly used this device before the experiment, or just a few times for testing it.

This two-dimension Fitts'law task is administrated using the multidirectional tapping task paradigm [25], described in the ISO9241-9 standard, in which the subjects have to successively clicked on circular targets placed along a circle (see Figure 3). This paradigm presents the advantage of controlling the effect of direction. The distance between two successively clicked targets corresponds to the amplitude (D) of the movement, and the size of each target is the width (W). Combining different widths, four different difficulties (ID), from 3 to 6, are proposed to the subjects (according to the formula $ID=log2(D/W + 1)$).

For starting the trial, the subject must click on the centered target. The first target is then highlighted. Because of the amplitude difference, the movement, which consists of reaching the first target from the starting point, is not kept in final results. In order to prove that the proposed device can reduce the homing time [3] (the time needed to reach the device from the keyboard, and vice versa), we decided to use the click-and-write technique, as proposed in [19, 7, 3], and combine it with the multidirectional tapping task [25]. Thus, a click on a target opens a small command line, in which the subjects were asked to write her/his first name and to press the 'RETURN' key for closing the command line. Since we do not implement a mode switching between pointing and typing, the cursor disappears at the bottom of the screen when the user starts typing. This evaluation is composed of 12 sessions: 3 sessions per device. Each session is composed of 12 trials: 3 trials per ID. One trial is composed of 9 clicks corresponding to the multidirectional tapping tasks. We then obtained 9x12x3 = 324 pointing events per device. For each device, the first session is considered as a training session, but subjects do not know it. The results of this session could be kept for analyzing learning effects of each device but they are not considered for the comparison of the device time performance. Considering the (ID, Session) couples, all the arrangements are used, in order to avoid learning/tiring or influence effects between the tested devices.

4.2 Results

Quantitative Results. For each trial, three times are recorded:

1. **Homing1**: elapsed time between the "RETURN" key press and pointing start
2. **Pointing Time**: elapsed time between pointing start and the click on the target
3. **Homing2**: elapsed time between the click and the first letter key press.

The experience data are analyzed within the framework of General Linear Model Procedure from SAS Software. There is a significant effect of device for all the recorded time types (Homing1: $F = 1976.3; p < 0.0001$; Pointing Time: $F = 283.41; p < 0.0001$; Homing2: $F = 24.04; p < 0.0001$). The classification of the device performances can be deduced from Figure 4. For Homing1, AirMouse is close to zero while mouse and touchpad are quite similar, but faster than key-joystick. Homing2 is similar for each device, from 0.655s for AirMouse to 0.548s for the mouse. This similarity is due to the time needed to find the first key to press on the keyboard. This value is not dependent to the ponting device. In Pointing Time, mouse is faster, followed by the touchpad. Then, AirMouse and key-joystick are close but key-joystick is slower. This result confirms the work presented by Douglas and al. [7].

Considering the total time (Homing1 + Pointing + Homing2), the device parameter has a significant effect ($F = 300; p < 0.0001$) and pairwise comparisons show a significant difference between all the devices ($p < 0.0001$ for each combination). The corresponding curve presented in Figure 4 presents the final classification of the devices. Despite of the very good performances of AirMouse for Homing1, Figure 4 allows us to point up the very good performances of the mouse for total time, maybe because of the expertise of the subjects. However, key-joystick is the slower device. AirMouse is slower but comparable to touchpad performances.

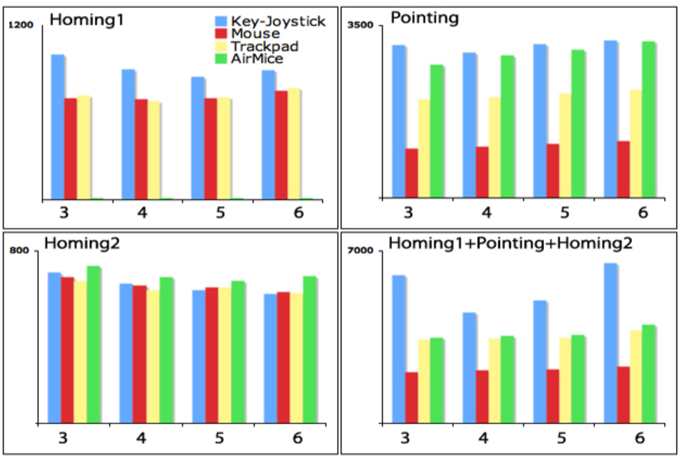


Fig. 4. Mean Time (in milliseconds) needed for each device and each difficulty for Homing1, Pointing, Homing2 and the sum of the 3 time values

Qualitative Results. Concerning the perceived performances, all the subjects consider the key-joystick as the slowest one. This is confirmed by the quantitative results previously presented. This could be explained by the fact that they never used the device before the experiment. The AirMouse is higher ranked, although it was the first time the participants use it. The mouse is considered as the faster device by 77% of the subjects. They are all experts, and the quantitative results confirm it. The AirMouse and the touchpad are quite similar, with a slight advantage for the AirMouse. However, subjects have never classified the touchpad as the faster device, in constrast with the AirMouse that has been classified as the faster device by 23% of the subjects.

Concerning the preference classification, the key-joystick is unanimously the less appreciated device (lower rates for all the subjects), mainly because of the fastidious learning time required. For the touchpad, results show the same pattern as previously, i.e. it is never the favorite device, but the second or the third selected device. AirMouse seems to be the most favorite device, i.e. it is classified as the most comfortable, intuitive and easy to learn device by 70% of the subjects. However, 77% of the subjects has classified the mouse in first or second position. Again, this could be explained by the mouse expertise of all the subjects.

Considering the tiring effect, only 10% of the subjects express a small feeling of tiredness in the hand for the AirMouse. But they consider it as a side effect of the experiment, i.e. the technique is new, and the hand is contracted for moving as fast as possible.

5 3D Evaluation

5.1 Pointing Task

As for the above 2D Fitts' law study, the goal of this evaluation is to position the 3D pointing performances of AirMouse in relation with other traditional device performances. We then compared our technique for 3D translation pointing with two well-known devices: the PHANTOM [18], an isotonic arm-based pointing device, which can provide haptic feedback (not used in the experiment) ; the SpaceNavigator [5], an isometric joystick.

Because of their isotonic property, and excluding the grasping action of the stylus, the movements with the PHANTOM and with the AirMouse for selecting and manipulating an object in translation are close. However, compared to theAirMouse, the PHANTOM is expensive and bulky. The comparison between the AirMouse and the spacemouse is interesting because of the popularity of the spacemouse. It is a low cost device, commonly used by designers for manipulating objects in 3D modelers. However, mainly because of its isometric property, its pointing time performances are lower than the ones of the PHANTOM [30]. Moreover, despite the tuning possibilities, the manipulation of such a device is not easy to learn and implies training time.

The PHANTOM and the AirMouse have a limited workspace, that we can consider as a cube. After preliminary tests, an empirical scale of each device workspace has been defined in relation to human skills, and then used for all the subjects. The scale of the AirMouse workspace has been defined in order to avoid subjects to move their hands for moving the cursor from the left to the right of the screen, allowing them to do it with only finger movements. Similarly the scale of the PHANTOM has been fixed in order to minimize arm movements.

3D pointing devices are usually used in manipulation tasks. Then, in order to fit with reality, the 3D pointing performances are evaluated with the same principle presented in [31] and recently used in [11]: subjects have to manipulate a tetrahedron and bring it inside another bigger one. Fitts' law studies can be used according to the Prince technique proposed in [15], the cursor being an area cursor.

Because 3D rotation is not considered and only 3D translation is used, the tetrahedrons are thus replaced by spheres. After selecting a green sphere with the end-effector of the device (represented by a small radius sphere, the same for each device), the subject has to bring it into a transparent spherical area (see Figure 5 for a snapshot of the 3D environment of the evaluation). Before the selection, the target area is not displayed, in order to avoid any anticipation of the recorded movement. The radius of the green manipulated sphere R is fixed. This manipulated object corresponds to an area cursor. While the distance D between the two spheres is fixed during the experiment, the radius of the target area R' is modified in order to define 3 different levels of difficulties (IDs): 3,4,5, according to the Fitts' law formula used in Prince:

$$ID = \log_2(D/(R'-R) + 1) . \quad (1)$$

The evaluation has been performed on a traditional laptop without simulation of visual stereoscopy. In order to improve the depth perception, real-time shadows have been added. The horizontal position of each sphere is represented by a black disk projected onto the bottom plan (see Figure 5) and perceived by the subjects in their peripheral vision. This evaluation is composed of 9 sessions: 3 sessions per device. Each session is composed of 21 trials: 7 trials per ID. We then obtained 9x21=189 events per device. For each device, the first session is considered as a training session, but subjects do not know it. The results of this session could be kept for analyzing learning effects of each device but they are not considered for the comparison of the device time performances. Considering the (ID, Session) couples, all the arrangements are used, in order to avoid learning/tiring or influence effects between the tested devices.

5.2 Results

Quantitative Results : Pointing Time. The experience data are analyzed within the framework of General Linear Model Procedure from SAS Software. For each trial, Pointing Time (PT) has been recorded between the date of the click required for selecting the green sphere to be manipulated and the date of sphere disappearance into the spherical target area. PT values (in seconds) are presented in Figure 5. There is a

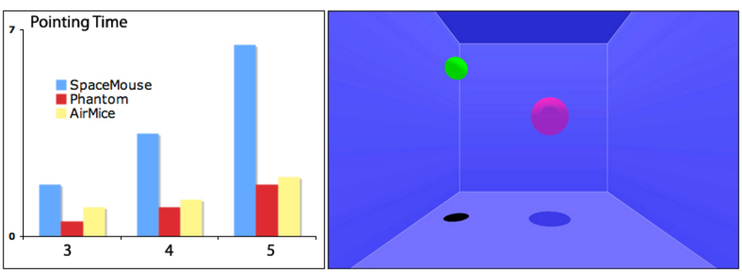


Fig. 5. Left: Mean Pointing Time (in seconds) for each device and for each ID. Subjects are slower with the SpaceMouse and faster with the PHANTOM. AirMouse is placed in between, but closer to the PHANTOM. Right: Snapshot of the 3D pointing evaluation environment. The green sphere on the left must be moved as fast as possible into the transparent red area. The transparent area size is modified during the experiment for defining different difficulties (IDs).

significant effect of the device (*F = 122.11; p < 0.0001*), with mean times decreasing from 4.9s (*SD = 4.4*) with the spacemouse, through 2.6s (*SD = 1.9*) with the AirMouse, to 1.8s (*SD = 1.2*) with the PHANTOM ; a 63% reduction in PT across the three conditions. The spacemouse is the slower device. Observing subjects during the experiment, we could notice that pointing movements are natural and intuitive for the PHANTOM and the AirMouse. However, using the spacemouse, subjects usually try to decompose the pointing movement: first, following the shadow cues, they try to adjust the position in the horizontal plan, then they adjust the height. This decomposition is not observed with other tested devices, so we suppose that it is linked to the spacemouse capabilities. This could explain the spacemouse low performances. Moreover, after the learning effect of the first session, the movement is more direct, but still slower than the movements with the two other tested devices.

Post-hoc pairwise comparisons showed a significant difference between the AirMouse and the spacemouse (*p < 0:0001*) and a less significant difference between the AirMouse and the PHANTOM (*p = 0.0168*). As expected by the Fitts'Law study, the difficulty (ID) has a significant effect on task performance (*F = 95.84; p < 0.0001*). Figure 5 shows both results, the effect of the device and the effect of the ID on PT. Mean PT increases with the ID for all the devices. It is higher with the spacemouse and lower with the PHANTOM. AirMouse is placed in between, but closer to the PHANTOM.

There is a significant effect of the session (*F = 23.05; p < 0.0001*), with mean times decreasing from 3.8s (*SD = 3.7*) for session 1, through 3.0s (*SD = 3.4*) for session 2, to 2,5 (*SD = 1.8*) for session 3; a 34% reduction of PT across the three conditions. However, session 1 and session 3 are significantly different (*p = 0.0003*), but session 2 and session 3 are not (*p = 0.06*). This effect is explained by the learning effect between session 1 and session 2, that seems to disappear between session 2 and session 3.

Qualitative Results. Concerning the perceived performances, without ambiguity, all subjects estimated that they are slower with the spacemouse. This is confirmed by the quantitative results previously presented. The main quoted reason is the learning stage linked to the sensitivity of the sensors. Concerning the PHANTOM and the AirMouse, 55% of the subjects are not able to know which of the two devices offer the best performance, and 33% took a decision and said that the PHANTOM is faster.

Concerning the preference classification, the spacemouse is unanimously the less favorite device. In contrast, the PHANTOM and the AirMouse are considered as intuitive, without learning stage. The viscosity of the PHANTOM (i.e. the resistance provided by the arm mechanism) is considered as helpful for precision by 44% of the subjects, but 33% consider it as a disturbing side effect. Compared to AirMouse, the PHANTOM is considered as less transparent and more invasive.

Concerning the tiring effects, 10% of the subjects express a feeling of tiredness in the hand for the AirMouse. 10% of the subjects also express a feeling of tiredness using the PHANTOM, because of the movements of the hand and the forearm. However, as for the previous evaluation, they explained it by the stress of the experiment, trying to perform the tasks as fast as possible.

6 Discussion and Future Work

As expected, these evaluations allow us to position AirMouse in relation to other existing and commonly used pointing devices. Results show that the performances of

AirMouse in 2D pointing are not better than the ones of the mouse, but are comparable to the ones of the touchpad, and are better than the ones of key-joystick. Since the pointing performance with FlowMouse has been reported to be significantly worse than with a touchpad [28], we conclude that AirMouse offers better pointing performance than FlowMouse. In 3D pointing, AirMouse is not so far from the PHANTOM and a lot faster than the spacemouse. Finally, qualitative results show that the performance is not the most important criterion. Subjects prefer to use a device that is intuitive and easy to learn while providing correct performance. Based on these criteria, AirMouse is appreciated by most of the users. They consider the technique as promising, and useful for laptop configuration.

Based on these encouraging results obtained for 2D /3D pointing tasks, it is now possible to further investigate AirMouse for other object manipulation tasks. AirMouse opens a vast world of possibilities in terms of interaction techniques. For example we plan to explore two-handed AirMouse interaction and gesture recognition as in [11, 9, 28] or based on real world metaphor: for example the user can perform a gesture similar to the one of turning a page in a book in order to scroll to the next page of a document. While for pointing tasks, mode switching between pointing and keyboard was not a key issue since the technique supports a direct designation of the objects on screen and therefore the cursor can move and disappear while typing, natural and efficient mode switching [28] is a primarily issue for the other tasks that we study and envision.

Since AirMouse seems very promising for 3D pointing, we first study the use of the AirMouse for full 3D manipulation and we started to investigate 3D rotation. A prototype has been designed and will be described in a next paper.

Finally an interesting feature of AirMouse is to support both 2D and 3D interaction in the same application. In order to informally evaluate the combined usage of 2D and 3D interaction, we tested the vision-based implementation of AirMouse in the context of a 3D modeler called Autodesk 3D Studio Max (3DSMax). Figure 6 shows a screenshot of the application. The two pointing techniques (2D and 3D) have been mixed: the 2D pointing technique, used in the experiment and based on two reflectors, has been plugged to the mouse cursor. The switch between 2D and 3D is based on the application mode defined by the cursor position. When the cursor of the mouse is over the 3D view, 3D pointing becomes active and 3D movements of the fingerTip are

Fig. 6. Left: Use of the two hands for 3D rotation: metaphor of planting a pin inside the object to be rotated. Right: Mixed 2D and 3D Interaction with AirMouse in Autodesk 3D Studio Max. the 2D pointing technique has been plugged to the mouse cursor while 3D pointing becomes active when the user moves over the perspective 3D view.

used. To sum up, the user can interact with 3DSMax, by moving the desktop cursor with its forefinger and clicking on icons. The user can also perform 3D manipulation as soon as the cursor is within the 3D scene: the arrow cursor is then replaced by a small sphere. The transition between 2D and 3D interaction is therefore observable as well as implicit and smooth based on the application mode activated by the position of the cursor and more importantly based on the same AirMouse technique.

7 Conclusion

In this paper, we have introduced and studied a new technique, namely AirMouse, for 2D and 3D interaction using finger gestures above the keyboard. The controlled experiment of a vision-based implementation of AirMouse shows the promising pointing performance of the technique compared with existing and commonly used devices for a pointing task. Subjects pointed out the intuitive, easy to learn and comfortable aspects of AirMouse that does not require additional surface for interaction using a laptop. In addition to our current studies of other tasks than pointing using AirMouse, a longitudinal evaluation of the 2D pointing is under investigation. We plan to test AirMouse with three regular computer users (scientists in the lab) in their everyday work, replacing the mouse by AirMouse.We hope to observe an improvement, which will make the technique comparable with the mouse in terms of time performance.

Acknowledgments. The authors would like to thank R. Blanch for extensive comments, A. Demeure for his help in the code for evaluation and all the participants in the experiments. Special thanks to N. Mandran for her precious help and advise on statistics for the evaluation and G. Serghiou for reviewing the paper. The work presented in the article is partly funded by the European Commission under contract OpenInterface (FP6-35182), www.oi-project.org.

References

1. Smith, R.B.: Put-that-here: Voice and gesture at the graphics interface. Computer Graphics 14, 262–270 (1980)
2. Bowman, D.A., Kruijff, E., LaViola, J.J., Poupyrev, I.: 3D User Interfaces: Theory and Practice. Addison-Wesley Professional, Reading (2005)
3. Card, S.K., English, W.K., Burr, B.J.: Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys, for text selection on a crt. In: Human-Computer Interaction, pp. 386–392. Morgan Kaufmann Publishers Inc, San Francisco (1978)
4. Card, S.K., Morans, T.P., Newell, A.: The Psychology of Human Computer Interaction. Lawrence Erlbaum Associates, Inc, Mahwah (1983)
5. 3D Connexion. Space navigator, `http://www.3dconnexion.com`
6. Dementhon, D., Davis, L.S.: Model-based object pose in 25 lines of code. International Journal on Computer Vision (1995)
7. Douglas, S.A., Mithal, A.K.: The effect of reducing homing time on the speed of a finger-controlled isometric pointing device. In: Computer-Human Interaction, pp. 411–416. ACM Press, New York (1994)
8. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47(6), 381–391 (1954)

9. Grossman, T., Widgor, D., Balakrishnan, R.: Multi-finger gestural interaction with 3d volumetric displays. In: User Interface Software and Technology. ACM Press, New York (2004)
10. Han, J.Y.: Low-cost multi-touch sensing through frustrated total internal reflection. In: Symposium on User Interface Software and Technology. ACM Press, New York (2005)
11. Hancock, M., Carpendale, S., Cockburn, A.: Shallow-depth 3d interaction: Design and evaluation of one-, two- and three-touch techniques. In: Computer-Human Interaction (2007)
12. Natural Point `http://www.naturalpoint.com`. Trackir `http://www.naturalpoint.com/trackir`
13. Iddan, G.J., Yahav, G.: 3d imaging in the studio. SPIE (2001), `http://www.3dvsystems.com.il`
14. Jacob, R.J.K., Sibert, L.E.: The perceptual structure of multidimensional input device selection. In: SIGCHI conference on Human factors in computing systems (1992)
15. Kabbash, P., Buxton, W.: The "prince" technique: Fitts' law and selection using area cursors. In: Computer-Human Interaction, pp. 273–279 (1995)
16. MacKenzie, I.S., Sellen, A., Buxton, W.: A comparison of input devices in elemental pointing and dragging tasks. In: Computer-Human Interaction, pp. 161–166 (1991)
17. Malik, S., Laszlo, J.: Visual touchpad: A two-handed gestural input device. In: ICMI (2004)
18. Massie, T.H., Salisbury, J.K.: The phantom haptic interface: A device for probing virtual objects. In: Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. ASME winter Meeting (1994)
19. Mysliwiec, T.A.: Fingermouse: A freehand computer pointing interface. Technical report for the degree of Masters of Science (October 1994)
20. Paljic, A., Burkhardt, J., Coquillart, S.: A study of distance of manipulation on the responsive workbench (2002)
21. Pierce, J.S., Stearns, B.C., Pausch, R.: Voodoo dolls: Seamless interaction at multiple scales in virtual environments. In: Symposium on Interactive 3D graphics (1999)
22. Qhueck, F.K.H.: Unencumbered gestural interaction. In: IEEE Multimedia (Winter 996)
23. Schneiderman, B.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley Longman Publishing Co., Inc, Amsterdam (1986)
24. Shneiderman, B.: The future of interactive systems and the emergence of direct manipulation. Human Factors in Interactive Computer Systems (1983)
25. Soukoreff, R.W., MacKenzie, I.S.: Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. International Journal of Human-Computer Studies 61, 751–789 (2004)
26. Tory, M., Moller, T., Atkins, M.S., Kirkpatrick, A.E.: Combining 2d and 3d views for orientation and relative position tasks. In: CHI 2004: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 73–80. ACM Press, New York (2004)
27. Vogel, D., Balakrishnan, R.: Distant freehand pointing and clicking on very large, high resolution displays. In: User Interface Software and Technology (2005)
28. Wilson, A.D.: Robust computer vision-based detection of pinching for one and twohanded gesture input. In: User Interface Software and Technology. ACM Press, New York (2006)
29. Wilson, A.D., Cutrell, E.: Flowmouse: A computer vision-based pointing and gesture input device. In: INTERACT: human-computer interaction (2005)
30. Zhai, S.: Human Performance in 6dof Input Control. PhD Thesis (1995)
31. Zhai, S., Milgram, P.: Quantifying coordination in multiple dof movement and its application to evaluating 6 dof input devices. In: Computer-Human Interaction, pp. 320–327. ACM Press, New York (1998)

Follow My Finger Navigation

Rami Ajaj, Frédéric Vernier, and Christian Jacquemin

LIMSI-CNRS and Université Paris-Sud 11
B.P. 133 / 91403 Orsay Cedex
France
{rami.ajaj,frederic.vernier,christian.jacquemin}@limsi.fr

Abstract. This paper presents a novel interaction technique called Follow my Finger (FmF) for navigation in 3D virtual environments using a 2D interactive view on a table-top device. FmF consists in moving a camera icon that represents the 2D subjective position and orientation of a viewpoint in the 3D world. Planar, tactile, and direct manipulation of the camera icon facilitates navigation in the 3D environment. From the user's perspective the camera icon follows her/his finger trajectory to interactively modify the horizontal location, inclination, and orientation of the 3D point of view.

Keywords: Table-top, navigation, interaction technique, 3D environments.

1 Introduction

Navigation techniques in 3D virtual environments have been developed for decades but each present advantages and drawbacks. A mouse and/or a keyboard are often used for navigation tasks in desktop 3D environments such as video games or computer-aided design applications. The use of these input devices seems less intuitive but faster and more accurate than direct touch of table-top devices [1]. This article studies the combination of 3D and 2D visualization and a corresponding interaction technique that facilitates intuitive navigation in virtual worlds through direct touch interactions (see Fig. 1).

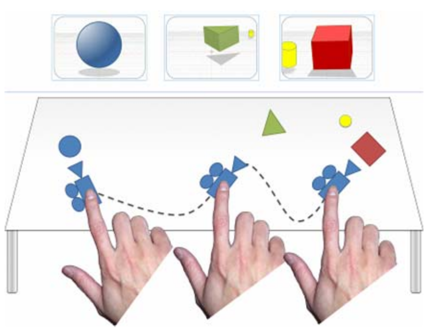


Fig. 1. FmF navigation. **Bottom**: The user's finger control of the camera icon. **Top**: The corresponding 3D view at 3 camera key positions and orientations.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 228–231, 2009.

2 Setup Description

The new interaction technique introduced in this article has been developed in a more complete framework for the combination of 2D and 3D interaction named Management Cabin (MC). MC is an architecture for multimodal interaction in 3D virtual worlds. It is made of a 2D interactive horizontal view of a virtual scene projected on a table-top device and a 3D view of the same scene presented vertically close to the table-top device in front of the user (see Fig. 2 left). A camera icon is displayed in the 2D table-top view and represents the point of view (POV) position and orientation in the 3D view. The figure 2 illustrates a user interacting through the table-top device and looking at the 3D view presented in front of her/him.

The MC software architecture is based on three components (see Fig. 2 right) that manage complementarity, consistency, and concurrency between the two interfaces (2D and 3D view and interaction). The central component controls the communication between the 2D and the 3D components.

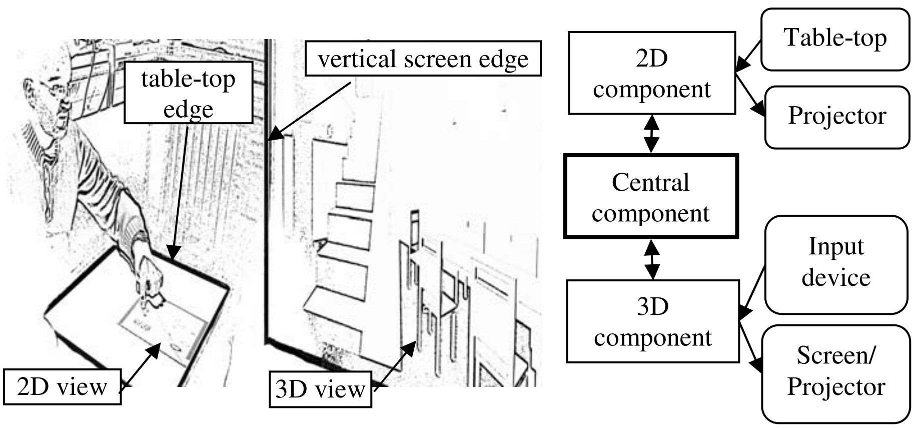


Fig. 2. Left: A user interacting with the 2D table-top view. **Middle**: The vertical screen in front of the user with the 3D view. **Right**: The software components and their connections to the input and output devices in MC.

3 Follow My Finger Navigation Technique

FmF Interaction Technique Description. At the 2D only level, the FmF navigation technique is similar to the Rotate'N Translate (RNT) interaction technique [2]. In RNT, window or 2D object orientation is automatically linked to the trajectory of interaction. In 3D, orientation can be automatically linked to the relative position of the camera and a virtual object like in HooverCam [3]. In FmF and RNT, orientation is deduced from the trajectory of the drag and drop. RNT is used directly for window manipulation while FmF uses a 2D camera icon linked to a 3D point of view for 3D navigation (like the HooverCam). FmF allows a user to select a camera icon with her/his finger and, while dragging the icon on the table, the camera orientation follows the trajectory. This technique controls both 2D rotation and 2D translation transformations of the 3D viewpoint. However translations and rotations can-not be performed separately.

Planar and direct manipulations of the camera icon allow easy and intuitive navigation in a 3D virtual environment. Moreover, simultaneous rotation and translation offer fast navigation task accomplishment when compared with interaction techniques that alternate translation and rotation tasks. A convincing example to illustrate these advantages is to consider a user navigating in a 3D scene that contains three objects of interest (see Fig. 1) in which she/he wants to visualize different key POVs in the 3D view.

The effect of the FmF interaction technique on the 3D navigation is a fluid and coherent navigation. Moreover, it allows a 2D direct manipulation of the 3D POV. RNT applied to 3D navigation does not cover the full range of possible degrees of freedom (DoF): vertical translation, pitch and roll are not covered. Among the three missing DoFs, the pitch is the most important for walk through navigation. Based on our observations, pitch and orientation seem related to trajectory features: while running fast we tend to look downward and looking upward is related to back movements. We have transposed these observations to FmF to control inclination of the viewpoint through fast or slow motions along the trajectory (forward=looking down and backward=looking up). To allow backward motions, we disable orientation linking to the trajectory for very sharp rotation angles (±60° around the opposite direction to trajectory). To better match walk through navigation, we apply a threshold to the pitch angle to fit human neck capabilities (±60° around horizontal). We manually adapted the pitch angle to the navigation speed according to table size and resolution. The pitch angle is modified linearly with respect to speed when navigating in the trajectory direction. We choose not to apply this linearity between pitch angle and speed to backward navigation to avoid users' dizziness. Therefore, pitch angle goes from 0° to the maximum threshold and stays fix during backward navigation. At last, two possibilities exist at finger release: the last pitch angle is preserved or it is put back to the relaxed angle (0°) with or without animation.

Multiple Cameras. Several camera icons can be displayed in the 2D view. Finger-based manipulation allows fast and easy POV change. In the case of one vertical display, only one camera, which represents the POV of the current 3D view, can be active. The POV change is made by choosing a button on the selected camera that will activate it. The button use allows an "offline" POV positioning and orientation (without inclination) by manipulating an inactive camera icon. The 3D POV is automatically moved to the new position that corresponds to the new current camera position and orientation in the 3D scene.

Multiple camera representations facilitate fast and easy POV modifications in scenes with multiple canonical views. For example a given POV might be important for a user who also wishes to explore other POVs in the virtual scene (e.g. in figure 1 the user wants to easily return to the POV in front of the sphere); In this case the user adds a new camera icon and navigates using this camera and return to the selected view through just one click. The use of multiple camera icons helps to overcome the lack of lateral translation in FmF.

Translated and Hybrid Interaction Techniques. In addition to FmF, we have implemented two other interaction techniques: the translated and the hybrid interaction techniques. The translated navigation technique allows the translation of the 3D POV horizontally through direct manipulation of the camera icon on the 2D view with user's finger. For rotation tasks, a widget is presented to the user next to the camera icon. The direct selection and manipulation of this widget leads to the 3D POV rotation. In comparison with the FmF technique, the translated technique allows separate and not simultaneous rotation and translation of the 3D POV. Moreover, the inclination is not managed in the translated technique.

The hybrid interaction technique combines both, translated and FmF, interaction techniques and offers the user the possibility to choose easily between both. Inspired by [2, 4], we divide the selection area of the camera icon into two parts. The front selection area of the camera icon corresponds to the FmF interaction technique whereas the back selection area corresponds to the translated navigation mode. This interaction technique allows simultaneous as well as separate 3D POV rotation and translation. The drawback is the toggle between two interaction techniques that could disorient users.

4 Discussion and Conclusion

We have implemented FmF in MC and used it to navigate in a small, two floors house model. The resulting interaction technique matches our expectations and the four DoFs are seamlessly integrated. However the parameters must be adjusted with end users and a user experiment must be conducted to study the assimilation of the four DoFs. FmF should also be compared to other navigation techniques that allow four DoFs. At last, we foresee how our extension to RNT [2] can also be applied to TNT [4] and provide a new family of very intuitive 3D navigation techniques by using the physical finger orientation.

References

1. Shen, C., Ryall, K., Forlines, C., Esenther, A., Vernier, F.D., Everitt, K., Wu, M., Wigdor, D., Morris, M.R., Hancock, M., Tse, E.: Informing the Design of Direct-Touch Tabletops. IEEE Comput. Graph. Appl. 26(5), 36–46 (2006)
2. Kruger, R., Carpendale, S., Scott, S.D., Tang, A.: Fluid integration of rotation and translation. In: Proc. of the SIGCHI Conference on Human Factors in Computing Systems CHI 2005, April 02 - 07, 2005, pp. 601–610. ACM, New York (2005)
3. Khan, A., Komalo, B., Stam, J., Fitzmaurice, G., Kurtenbach, G.: HoverCam: interactive 3D navigation for proximal object inspection. In: Proc. of the 2005 Symposium on interactive 3D Graphics and Games, I3D 2005, Washington, District of Columbia, April 03 - 06, 2005, pp. 73–80. ACM, New York (2005)
4. Liu, J., Pinelle, D., Sallam, S., Subramanian, S., Gutwin, C.: TNT: improved rotation and translation on digital tables. In: Proc. of Graphics interface 2006, Quebec, Canada, June 07 - 09, 2006, vol. 137, pp. 25–32. Canadian Information Processing Society, Toronto (2006)

DGTS: Integrated Typing and Pointing

Iman Habib[1], Niklas Berggren[1], Erik Rehn[1], Gustav Josefsson[1], Andreas Kunz[2], and Morten Fjeld[1]

[1] CSE; ETA Lab; t2i Lab, Chalmers TH, Rännv. 6, SE-412 96 Gothenburg, Sweden
fjeld@chalmers.se
[2] ICVR, ETH Zurich, Tannenstr. 3, CH-8092 Zurich, Switzerland
kunz@iwf.mavt.ethz.ch

Abstract. Capacitive sensing is used in many different fields of application. It has been implemented in such devices as mobile phones and remote controls. However, up until now the physical sensing area has remained limited despite the widespread use of larger input devices such as keyboards. We present DGTS, which seamlessly integrates keyboard typing and cursor pointing. This input device offers multi-finger operation for scrolling and other specialized input commands. The objective of this work is to replace computer mice and touchpads by integrating capacitive sensing into a layer within the keyboard thereby reducing the space required for pointing devices. This paper gives the technical background, shows our contribution, and concludes with initial tests.

Keywords: Capacitive sensing, touchpad, multi pointer, typing, pointing.

1 Introduction

Ever since Graphical User Interfaces (GUIs) were developed for computers, the mouse has been the most common pointing device; however, the physical space occupied by a mouse is considerable. Most laptop computers come with a pointing stick or a touchpad. A pointing stick is an isometric joystick that is situated close to or in the middle of the keyboard and is controlled by the user's index finger. Capacitive sensing was introduced in PreSense [1] and SmartPad [2] as a technology to apply with a large sensor for controlling a mouse pointer on a mobile device. Another example is IMPAD, a paper-thin, pressure sensitive multitouch device [3]. Capacitive touch-pads have also been used in combination with a keyboard by ThumbSense [4] to change the keyboard's operation mode. Other input technologies such as the layered touch panel [5] or the PointingKeyboard [6] employ additional infrared-grid sensors to realize more functionality. Most of these systems offer limited interaction space.

Integrating capacitive sensing into a keyboard presents at least two issues. Firstly, there is the height detection problem regarding *touchdown* and *liftoff;* that is, the sensors might have difficulties detecting whether or not a finger has touched the keyboard [7]. Secondly, there is the issue of the key's sensor mechanics. In order to detect when a key is pushed, the capacitive sensors must not be in the way of the keyboard's regular circuitry, thus requiring a thin and flexible sensor surface.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 232–235, 2009.

2 Contribution

We present DGTS[1] employing capacitive sensors to detect the position of human digits on/over the keyboard. The name is at once an abbreviation of "digits" (i.e. human fingers) and an allusion to digital data, thus encapsulating the concept of finger-based input for digital data. Capacitive finger sensing can be done in multiple ways, each with their respective pros and cons [8][9]. The sensor mode influences the required electrode grid and thus the available resolution. Table 1 presents three alternative sensor modes with the pros (+) and cons (–) that determine their uses as capacitive buttons and large multi-touch surfaces (shunt mode), multi-user systems (transmit mode), and touchpad and capacitive buttons (loading mode).

Table 1. Sensor modes with pros [+] (left) and cons [–] (right)

Shunt mode + Delimited area of sensitivity + Multiple senders and receivers + Electrode can have sender and receiver status	– Difficult to shield – 2 electrodes/sensors
Transmit mode + User-sensitive + 1 electrode/sensor	– Less quality in measurements with increasing user grounding
Loading mode + Easy to shield + 1 electrode/sensor	– Sensitive to all conductive objects – Short detection range

By combining information from several sensors, an image of hands and fingers on the keyboard can be created. Depending on the sensor patterns [10], the sensors create two-dimensional images with different characteristics. Table 2 shows a number of sensor patterns with pros (+), major-pros (+ +), cons (–), and other relevant facts.

Table 2. Comparison of sensor patterns, [–, +, + +]

Sensor Pattern	Surface/Electrode	Pattern Type	Sensor Modes	Multi-touch
CellMatrix	–	Matrix	Loading, Shunt	+ +
SnakePit	+	X-Y	Loading	–
Synaptics	+	X-Y	Loading	–
Hexagonal	+	X-Y	Shunt	–
TT [7][11]	+	X-Y	Loading	–

Unlike SmartSkin [12] working in shunt mode, DGTS is based on the CellMatrix pattern working in loading mode. We chose the loading mode principle together with the CellMatrix layout, because it best suits our requirements for simplicity of the setup and precision of the position readout. We use 1 cm^2 sensors, then average multiple sensors' readouts, and achieve a resolution of ~1 mm combined with a high update rate for the sensor readout. The sensor size had to fulfill these requirements:

[1] Video presentation of DGTS: http://www.t2i.se/pub/media/dgts.wmv

Fig. 1. Left: Color value map and finger position on the keyboard. Right: DGTS electronics.

- Since a finger's position is calculated from multiple sensor values, many sensors should be used for blob detection, i.e. they should be as small as possible.
- To realize a reasonable capacity and measure capacity changes when a finger approaches the sensor, the electrodes should be set to the size of a fingertip.

If the electrode's size is too large, the proximity of a finger only causes small capacity changes compared to the sensor's basic capacity. Thus, the change might not be detectable by the electronics (resolution of the A/D-converter). In our prototype, a 96 cm^2 sensor surface with several functionalities was realized (Figure 1). It controls the cursor, detects when keys are pressed, recognizes simple gestures, and discerns when mouse clicking is performed on the keyboard's surface. The prototype's update frequency is 13.68 Hz. The data read is sent to the computer via a USB port.

Electronics: The electrodes are connected to the capacitance-to-digital converter (CDC) from Analog Devices AD7147 via signal selectors [13]. The AD7147 consists of a sigma-delta-based CDC with 12 analog input channels and communicates with a microcontroller via an I^2C bus. Since 96 sensors exist so far, the 12 channels of the AD7147 are not enough. We use an 8 channel signal selector (analog multiplexer) to connect eight sensors to each of the CDC's 12 analog input ports. A microcontroller enables communication between the sensor board and the PC. The PC-software continually dispatches requests for data and then processes feedback.

Blob Detection: The blob detection algorithm starts with the sensor that has the highest readout. Next, all neighboring cells are assumed to belong to the same blob as long as they have a readout that is smaller than the previous one but still above a certain threshold (coming from the calibration). The algorithm stops if a neighboring cell's readout is below this threshold or if the readout increases again (Figure 1).

Software Design: The software performs several tasks. It handles incoming packets via USB and generates field images from them. A three-step calibration consists of i) sampling background-noise level when the system is in idle mode, ii) setting a hi-pass filter by holding a finger fixed for a few seconds, and iii) normalization of a maximum sampling level by hovering a finger above the sensor area. Based on this, a sensor value map is provided showing detected fingers and is used to control the cursor (see Fig.1). In order to address existing applications like a browser, it is possible to use the sensor data to control a mouse pointer. Signal duration is used to distinguish between mouse movement (long) and mouse clicking (short). Although controlling the mouse pointer requires only a single touch, we implemented a dual finger gesture for other functions e.g. scrolling on a webpage.

3 Discussion and Outlook

To gather initial feedback from potential users, we had six experienced computer users experiment with the system's functionality[1], each user for about fifteen minutes. We observed a noticeable effect of keyboard texture as test subjects controlled the mouse pointer. Standard keyboards are structured with indentations or gaps between the keys through which the user's fingers can navigate. However, the uneven shape of keycaps is not an optimal surface for precisely controlling a mouse pointer. A future solution to make the keyboard smoother may be the use of a foil keyboard.

The current prototype has a width-/height-ratio of 1.5, which is between the typical screen ratios of 4:3 and 16:9. Although this ratio does not fit exactly, none of the users noticed this difference. We note that these findings might change if we apply the sensing capability to a complete keyboard, thus generating a larger mismatch. Future work will focus on the question of whether it makes more sense to use a complete keyboard as an absolute pointing device or just its separate number block.

Since DGTS sensor surface can be used to detect both pointer movements and button strokes, the technology could be used in Ultra-Mobile Personal Computers (UMPC), which have inherently compact designs. Combining the keyboard and the touchpad in such a way that they use the same surface is worthwhile in such devices.

References

1. Rekimoto, J., Ishizawa, T., Schwesig, C., Oba, H.: PreSense: interaction techniques for finger sensing input devices. In: UIST 2003, pp. 203–212. ACM, New York (2003)
2. Rekimoto, J., Oba, H., Ishizawa, T.: SmartPad: a finger-sensing keypad for mobile interaction. In: Proc. CHI 2003, pp. 850–851. ACM Press, New York (2003)
3. Rosenberg, I.D., Grau, A., Hendee, C., Awad, N., Perlin, K.: IMPAD: an inexpensive multi-touchpressure acquisition device. In: Proc. CHI EA 2009, pp. 3217–3222. ACM, New York (2009)
4. Rekimoto, J.: ThumbSense: Automatic input mode sensing for touchpad-based interactions. In: CHI 2003 Late braking (2003)
5. Tsukada, Y., Hoshino, T.: Layered touch panel: the input device with two touch panel layers. In: Proc. CHI 2002 Interactive Poster, pp. 71–80 (2002)
6. Tsukuda, Y., Hoshino, T.: PointingKeyboard: An input device for typing and pointing. In: Proc. Workshop on Interactive Systems and Software (WISS 2002) (2002) (in Japanese)
7. Fallot-Burghardt, W., Fjeld, M., Speirs, C., Ziegenspeck, S., Krueger, H., Läubli, T.: Touch&Type: a novel pointing device for notebook computers. In: Proc. NordiCHI 2006, vol. 189, pp. 465–468. ACM Press, New York (2006)
8. Baxter, L.K.: Capacitive Sensors, Design and Applications. IEEE Press, NY (1997), `http://www.capsense.com/capsense-wp.pdf`
9. Mackey, B.L., Golovchenko, M.: Capacitive Sensing Apparatus Designs, patent (2007)
10. Hinckley, K., Sinclair, M.: Touch-sensing input devices. In: Proc. CHI 1999, pp. 223–230 (1999)
11. Fallot-Burghardt, W., Speirs, C., Ziegenspeck, C., Krueger, H., Läubli, T.: Touch&Type^{TM}: a Novel Input Method for Portable Computers. In: Proc. INTERACT 2003, pp. 954–957 (2003)
12. Rekimoto, J.: SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. In: Proceeding of the SIGCHI conference on Human factors in computing systems: Changing our World, changing ourselves, pp. 113–120 (2002)
13. Analog Devices Data Sheet AD 7147, `http://www.analog.com/static/imported-files/Data_sheets/AD7147.pdf` (accessed, 22.05.2009)

Understanding Multi-touch Manipulation for Surface Computing

Chris North[1], Tim Dwyer[2], Bongshin Lee[2], Danyel Fisher[2], Petra Isenberg[3], George Robertson[2], and Kori Inkpen[2]

[1] Virginia Tech, Blacksburg, VA, USA
north@vt.edu
[2] Microsoft Research, Redmond, WA, USA
{t-tdwyer,danyelf,bongshin,ggr,kori}@microsoft.com
[3] University of Calgary, Alberta, Canada
petra.isenberg@ucalgary.ca

Abstract. Two-handed, multi-touch surface computing provides a scope for interactions that are closer analogues to physical interactions than classical windowed interfaces. The design of natural and intuitive gestures is a difficult problem as we do not know how users will approach a new multi-touch interface and which gestures they will attempt to use. In this paper we study whether familiarity with other environments influences how users approach interaction with a multi-touch surface computer as well as how efficiently those users complete a simple task. Inspired by the need for object manipulation in information visualization applications, we asked users to carry out an object sorting task on a physical table, on a tabletop display, and on a desktop computer with a mouse. To compare users' gestures we produced a vocabulary of manipulation techniques that users apply in the physical world and we compare this vocabulary to the set of gestures that users attempted on the surface without training. We find that users who start with the physical model finish the task faster when they move over to using the surface than users who start with the mouse.

Keywords: Surface, Multi-touch, Gestures, Tabletop.

1 Introduction

The rapidly-developing world of multi-touch tabletop and surface computing is opening up new possibilities for interaction paradigms. Designers are inventing new ways of interacting with technology and users are influenced by their previous experience with technology.

Tabletop gestures are an important focal point in understanding these new designs. Windowing environments have taught users to experience computers with one hand, focusing on a single point. What happens when those constraints are relaxed, as in multi-touch systems? Does it make sense to allow—or expect—users to interact with multiple objects at once? Should we design for users having two hands available for their interactions? Both the mouse-oriented desktop and the physical world have constraints that limit the ways in which users can interact with multiple objects and users come to the tabletop very accustomed to both of these.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 236–249, 2009.

There is no shortage of applications where users might need to manipulate many objects at once. From creating diagrams to managing files within a desktop metaphor, users need to select multiple items in order to move them about. A number of projects in the visual analytics [11] and design spaces [6] have attempted to take advantage of spatial memory by simulating sticky notes—a mixed blessing when rearranging the notes is expensive and difficult. As it becomes simpler to move objects and the mapping between gesture and motion becomes more direct, spatial memory can become a powerful tool.

We would like to understand what tools for managing and manipulating objects the tabletop medium affords and how users respond to it. Particularly, we would like to understand the techniques that users adopt to manipulate multiple small objects. What techniques do they use in the real world and how do those carry over to the tabletop context? Do they focus on a single object—as they do in the real world—or look at groups? Do they use one hand or two? How dexterous are users in manipulating multiple objects at once with individual fingers?

The problems of manipulating multiple objects deftly are particularly acute within the area of visual analytics [13], where analysts need to sort, filter, cluster, organize and synthesize many information objects in a visualization. Example systems include In-Spire [16], Jigsaw [12], Occulus nSpace [10], or Analyst's Notebook [4], i.e. systems where analysts use virtual space to organize iconic representations of documents into larger spatial representations for sensemaking or presenting results to others. In these tasks, it is important to be able to efficiently manipulate the objects and it is often helpful to manipulate groups of objects. Our general hypothesis is that multi-touch interaction can offer rich affordances for manipulating a large number of objects, especially groups of objects.

A partial answer to these questions comes from recent work by Wobbrock et al. [17]. Users in that study were asked to develop a vocabulary of gestures; the investigators found that most (but not all) of the gestures that users invented were one-handed. However, their analysis emphasized manipulating single objects: they did not look at how users would handle gestures that affect groups of items.

In this paper we explore how users interact with large numbers of small objects. We discuss an experiment in which we asked users to transition from both a mouse and a physical condition to an interactive surface, as well as the reverse. We present a taxonomy of user gestures showing which ones were broadly used and which were more narrowly attempted. We also present timing results showing that two-handed tabletop operations can be faster than mouse actions, although not as fast as physical actions. Our research adds a dimension to Wobbrock et al.'s conclusions showing that two-handed interaction forms a vital part of surface gesture design.

2 Background

Typical interactions on groups of items in mouse-based systems first require multi-object selection and then a subsequent menu selection to specify an action on the selected objects. Common techniques for multi-object selection include drawing a selection rectangle, drawing a lasso, or holding modifier keys while clicking on several objects. In gestural interfaces this two-step process can be integrated into one motion. Yet, the design of appropriate gestures is a difficult task: the designer must develop gestures that can be both reliably detected by a computer and easily learned by people [5].

Similar to the mouse, pen-based interfaces only offer one point of input on screen but research on pen gestures is relatively advanced compared to multi-touch gestures. Pen-based gestures for multiple object interaction have, for example, been described by Hinckley et al. [3]. Through a combination of lasso selection and marking-menu-based command activation, multiple targets can be selected and a subsequent action can be issued. A similar example with lasso selection and subsequent gesture (e.g., a pigtail for deletion) were proposed for Tivoli, an electronic whiteboard environment [9].

For multi-touch technology, a few gesture sets have been developed which include specific examples of the types of multi-object gestures we are interested in. For example, Wu et al. [18] describe a Pile-n-Browse gesture. By placing two hands on the surface, the objects between both hands are selected and can be piled by scooping both hands in or browsed through by moving the hands apart. This gesture received a mixed response in an evaluation. Tse et al. [14] explore further multi-touch and multi-modal group selection techniques. To select and interact with multiple digital sticky notes, users can choose between hand-bracketing (similar to [18]), single-finger mouse-like lasso-selection, or a speech-and gesture command such as "search for similar items." Groups can then be further acted upon through speech and gestures. For example, groups of notes can be moved around by using a five-fingered grabbing gesture and rearranged through a verbal command. Using a different approach, Wilson et al. [15] explore a physical based interaction model for multi-touch devices. Here, multiple objects can be selected by placing multiple fingers on objects or by "pushing" with full hand shapes or physical objects against virtual ones to form piles.

Many of the above multi-selection gestures are extremely similar to the typical mouse-based techniques (with the notable exception of [15]). Wobbrock et al. [17] present a series of desired effects, and invite users to "act out" corresponding gestures in order to define a vocabulary. Participants described two main selection gestures—tap and lasso—for both single and group selection. This research also showed a strong influence of mouse-based paradigms in the gestures participants chose to perform. Similarly, our goal was to first find out which gestures would be natural choices for information categorization and whether a deviation from the traditional techniques of lasso or selection rectangles would be a worthwhile approach.

Previous studies have examined the motor and cognitive effects of touch screens and mouse pointers, and the advantages of two-handed interaction over one-handed techniques, primarily for specific target selection tasks (e.g., [1,7]). Our goal is to take a more holistic view of multi-touch interaction in a more open-ended setting of manipulating and grouping many objects.

3 Baseline Multi-touch Surface Interaction

Our goal is to study tasks in which users manipulate large numbers of small objects on screen. For our study, we abstracted such analytic interactions with a task involving sorting colored circles in a simple bounded 2D space.

Our study tasks, described below, involved selecting and moving colored circles on a canvas. We were particularly interested in multi-touch support for single and group selection of such objects. To provide a study platform for comparison with standard mouse-based desktop and physical objects conditions, we had to make some interaction design decisions for our baseline multi-touch system. Our design incorporates several assumptions about supporting object manipulation for surface computing:

- One or two fingers touching the surface should select individual objects.
- A full hand, or three or more fingers touching the surface, should select groups of objects.
- Contacts far apart probably indicate separate selections (or accidental contact) instead of a very large group. Unintentionally selecting a large group is more detrimental than selecting small groups.
- Multiple contacts that are near each other but initiated at different times are probably intended to be separate selections. Synchronous action might indicate coordinated intention.

The system (Fig. 1) is implemented on the Microsoft Surface [8], a rear-projection multi-touch tabletop display. The Surface Software Development Kit provides basic support for hit testing of users' contact points on the display. It also provides coordinates and an ellipsoidal approximation of the shape of the contact, as well as contact touch, move, and release events.

Our testing implementation supports selecting and dragging small colored circles both individually and in groups. The interaction design was intentionally kept simple to support our formative study goals. Contacts from fingers and palms select all the circles within their area. As feedback of a successful selection, the circles are highlighted by changing the color of their perimeters, and can be dragged to a new position. From there, they can be released and de-selected. A (small) fingertip contact selects only the topmost circle under the contact, enabling users to separate overlapping circles. Large contacts such as palms select all circles under the contact. Using multiple fingers and hands, users can manipulate multiple circles by such direct selection and move them independently. Such direct selection techniques are fairly standard on multi-touch interfaces.

We also provide an analogue to the usual mouse-based rectangular marquee selection of groups of objects. However, a simple rectangular marquee selection does not make effective use of the multi-touch capability. Instead, the users can multi-select by defining a convex hull with three or more fingers. If three or more contacts occur

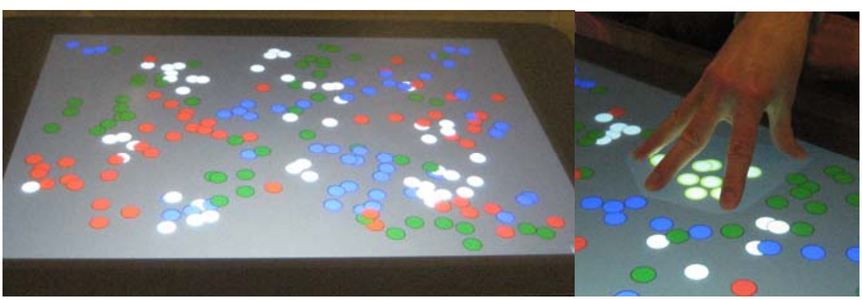


Fig. 1. Left: The system at the start of Task 1. **Right**: One-handed hull selection technique.

within 200ms and within a distance of 6 inches from each other (approximately a hand-span), then a convex-hull is drawn around these contacts and a “group” selection is made of any circles inside this hull (Fig. 1, right). The background area inside the hull is also colored light grey to give the user visual feedback. These hulls, and the circles within them, can then be manipulated with affine transformations based on the users’ drag motions. For example, users can spread out or condense a group by moving their fingers or hands together or apart. While the group selection is active, users can grab it with additional fingers to perform the transformations as they desire. The group selection is released when all contacts on the group are released.

4 Study Design

The goal of this study is to discover how users manipulate many small objects, in three different interaction paradigms: physical, multi-touch, and mouse interaction. To support our formative design goals, we took a qualitative exploratory approach with quantitative evaluation for comparisons.

4.1 Participants

We recruited 32 participants (25 males and 7 females) and 2 pilot testers via email from our institution. We screened participants for color blindness. They were mainly researchers and software developers who were frequent computer users. The average age of participants was 34, ranging from 21 to 61. None of the participants had significant experience using the Surface. Participants had either never used the Surface before or had tried it a few times at demonstrations. Participants each received a $US10 lunch coupon for their participation. To increase motivation, additional $10 lunch coupons were given to the participants with the fastest completion time for each interface condition in the timed task.

4.2 Conditions and Groups

We compared three interface conditions: Surface, Physical and Mouse. For both the Surface and Physical conditions, we used a Microsoft Surface system measuring 24" × 18". For the *Surface* condition (Fig. 1, left), we ran the multi-touch implementation described in Section 3 with 1024 × 768 resolution. For the *Physical* condition (Fig. 2, Left), we put 2.2cm diameter circular plastic game chips on top of the Microsoft Surface tabletop with same grey background (for consistency with the Surface condition). The circles in the Surface condition were the same apparent size as the game chips in the Physical condition.

For the *Mouse* condition (Fig. 2, right), we ran a C# desktop application on a 24" screen. This application supported basic mouse-based multi-selection techniques: marquee selection by drawing a rectangle as well as control- and shift-clicking nodes. Circles were sized so that their radii as a proportion of display dimensions were the same on both the desktop and surface.

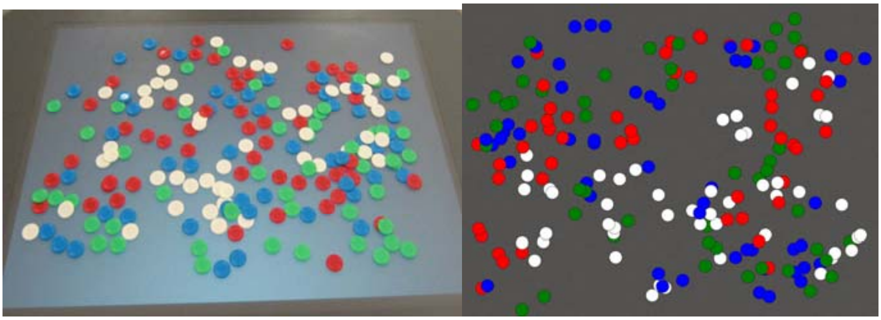


Fig. 2. (Left) Physical condition and **(Right)** Mouse condition

Since our goal is to compare the Surface condition against the other two conditions, each participant used only two conditions: Surface and one of the others. Users were randomly divided into one of four groups: Physical then Surface (PS), Surface then Physical (SP), Mouse then Surface (MS), Surface then Mouse (SM). This resulted in participants' data for 32 Surface, 16 Physical and 16 Mouse.

4.3 Tasks

Participants performed four tasks, each task requiring spatially organizing a large number of small objects. The first and second tasks were intended to model how analysts might spatially cluster documents based on topics, and manage space as they work on a set of documents, and were designed to capture longer-term interaction strategies. The tasks required a significant amount of interaction by the participants and gave them a chance to explore the interface.

All participants worked on the four tasks in the same order, and were not initially trained on the surface or our application. Participants were presented with a table of 200 small circles, with 50 of each color: red, green, blue, and white. Fig. 1 illustrates the 200 circles on a Surface at the start of the first task, positioned randomly in small clusters.

With the exception of Task 3, which was timed, we encouraged participants to think aloud while performing the tasks so that we could learn their intentions and strategies.

Task 1: Clustering task. This task was designed to elicit users' intuitive sense of how to use gestures on the surface. The task was to organize the blue and white circles into two separate clusters that could be clearly divided from all others. Participants were told that the task would be complete when they could draw a line around the cluster without enclosing any circles of a different color. Fig. 3. shows one possible end condition of Task 1.

Task 2: Spreading Task. Participants spread out the blue cluster such that no blue circles overlap, moving other circles to make room as needed. Participants start this task with the end result of their Task 1.

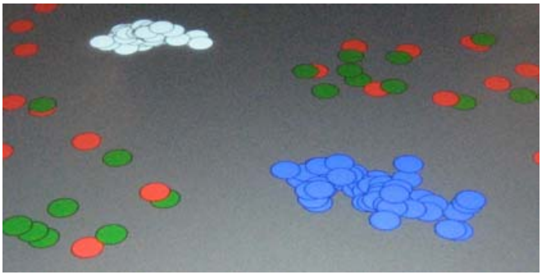


Fig. 3. Example end condition of Task 1

Task 3: Timed Clustering Task. This task was designed to evaluate user performance time for comparison between interface conditions and to examine the strategies which users adopt over time. Task 3 repeated Task 1, but participants were asked to complete the task as quickly as possible. They were not asked to 'think aloud' and a prize was offered for the fastest time.

Task 4: Graph Layout Task. Inspired by the recent study of van Ham and Rogowitz [2], we asked participants to lay out a social network graph consisting of 50 nodes and about 75 links. In the Physical condition, participants did not attempt this task. Due to the broader scope and complexity of this task, the analysis of the results of Task 4 will be reported elsewhere.

4.4 Procedure

Each participant was given an initial questionnaire to collect their demographics and prior experience with the Microsoft Surface system. Participants completed Tasks 1 and 2 without training, in order to observe the gestures they naturally attempted. Participants in the Surface and Mouse condition were given a brief tutorial about the available interaction features after Task 2. At the end of each condition participants answered a questionnaire about their experience. They then repeated the same procedure with the second interface condition. At the end of the session participants answered a final questionnaire comparing the systems. Each participant session lasted at most an hour.

We recorded video of the participants to capture their hand movements, their verbal comments and the display screen. The software also recorded all events and user operations for both the Surface and Mouse conditions.

5 Results

We divide our results into an analysis of the set of gestures that users attempted for Tasks 1 and 2, timing results from Task 3, and user comments from the after-survey.

5.1 Gestures

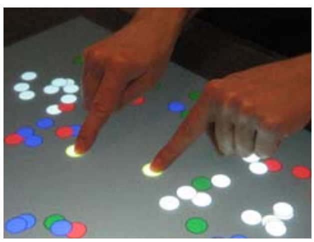
The video data for Task 1 and 2 (clustering and spreading) were analyzed for the full set of operations users attempted in both the Physical and Surface conditions. We first used the video data to develop a complete list of all gestures, both successful and unsuccessful. For example, if a participant attempted to draw a loop on the surface, we coded that as an unsuccessful attempt to simulate a mouse “lasso” gesture. The gestures were aggregated into categories of closely-related operations. Once the gestures were identified, the videos were analyzed a second time to determine which gestures each user attempted.

Table 1 provides a listing of all classes of gestures that participants performed during the study; six of them are illustrated in Fig. 4. These gestures are divided into several categories: single-hand operations that affect single or groups of objects, two-handed gestures that affect multiple groups of objects, and two-handed gestures that affect single groups. Last, we list gestures that apply only to one medium: just surface, and just physical.

In order to understand how gestures varied by condition, we classed gestures by which participants attempted them. Table 2 lists all of the gestures that were feasible in both the Physical and Surface conditions. This table also lists the percentage of participants who utilized each gesture at least once during the session. This data is aggregated by the Physical and Surface conditions, followed by a further classification by which condition was performed first (Physical, Mouse, Surface). Table 3 lists additional gestures that were only feasible for the Surface condition, while Table 4 lists gestures that were only used in the Physical condition.

Across all participants the most popular gestures were those that entailed using fingertips to move circles across the table—all participants moved at least some items around that way. While all participants realized they could move physical objects with

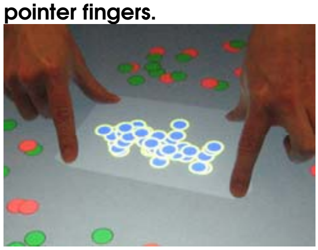
(a) One-hand shove.



(b) Drag two objects with pointer fingers.

(c) Two hands grab groups.

(d) Add/remove from selection.



(e) Two-hand transport.

(f) Both hands coalesce large group to small.

Fig. 4. Six selected one- and two-handed gestures attempted by participants during the study (see Table 1 for the full list)

Table 1. Descriptions of gestures

	Individual Items	**Groups**
One hand	**Drag single object**. Drag a single item across the tabletop with a fingertip. **Drag objects with individual fingers.** Using separate fingers from one hand, drag individual items across the table **Toss single object**. Use momentum to keep an object moving across the tabletop.	**Splayed hand pushes pieces**. (Fig. 1) An open hand pushes pieces. Could define a hull. **Hand and palm**. A single hand is pressed flat against the table to move items underneath it. **One hand shove** (Fig. 4a). Moves many points as a group. **Pinch a pile.** Several fingers "pinch" a group of pieces together. In the Surface condition, this would define a (very small) hull.
	Coordinated, >1 Group	**Coordinated, 1 Group**
Two hands	**Drag two objects with pointer fingers** (Fig. 4b). Does not entail any grouping operations. **Two hands grab points in sync**. Each hand has multiple fingers pulling items under fingers. **Rhythmic use of both hands.** "Hand-over-hand" and synchronized motion, repeated several or many times. **Two hands grab groups** (Fig 4c)**.** Hands operate separately to drag groups or individual points.	**Both hands coalesce large group to small** (Fig 4f). **Two-hand transport** (Fig 4e)**.** Use two hands to grab a group and drag across the region. **Add/remove from selection** (Fig 4d)**.** Use one hand to pull an object out of a group held by the other.
	Surface only	**Physical Gestures**
By condition	**One hand expand/contract**. Use a single hand with a convex hull to grow or shrink a hull. **Two-hand hull tidy**. Use fingers from two hands with a convex hull to shrink the hull to make more space. **Two-hand hull expand/contract**. Use fingers from two hands with a convex hull to manipulate the hull. **Expand hull to cover desired nodes**. Define a hull first, then expand it to cover more nodes. Does not work on our Surface implementation. **Treat finger like a mouse**. Includes drawing a lasso or marquee with one or two hands, different fingers of the hand for "right" click, or holding down one hand to "shift-click" with the other. **Push hard to multi-select**. Press a finger *harder* into the table to hope to grow a selection or select more items in the near vicinity.	**Lift Up**. Pick up chips in the hand, carry them across the surface, and deposit them on the other side. **Go outside the lines**. Move, stack, or slide chips on the margin of the table, outside the screen area. **Slide around objects**. When sliding circles, choose paths across the space that avoid other circles. **"Texture"-based gestures.** Slide chips under palms and fingers and shuffle them, using the feel of the chip in the hand. **Toss items from one hand to other**. Take advantage of momentum to slide chips from one hand to the other. **Drag a handful, dropping some on the way**. Intentionally let some chips fall out of the hand, holding others, to either spread out a pile or sort them into different groups.

two hands, six of them never thought to try that with the Surface (three that started in Surface condition; three from group MS). Closer examination of the gesture data revealed that participants who started with the physical condition were much more likely (88%) to try multiple fingers with both hands than users who started with the mouse (56%) or the surface (50%).

When participants worked with two hands on the surface they almost always used them on separate groups: only 30% of participants performed operations that used

Table 2. Gestures that apply to both the physical and surface conditions. Values indicate the percentage of subjects who used the gesture at least once. Values over 50% are boldfaced.

	Physical	Surface	Surface (by 1st Condition)		
			After Mouse	After Physical	Surface 1st
	(n=16)	(n=32)	(n=8)	(n=8)	(n=16)
1 Hand, Individual Items					
Drag single object	**75%**	**94%**	**100%**	**75%**	**100%**
Drag objects with indiv fingers	**81%**	**69%**	50%	50%	**88%**
Toss single object	38%	19%	0%	13%	31%
1 Hand, Groups					
Splayed hand pushes pieces (Fig 1)	50%	28%	25%	25%	31%
One hand shove (Fig 3a)	**75%**	47%	38%	38%	**56%**
Hand and palm	31%	41%	25%	25%	**56%**
Pinch a pile	6%	38%	13%	25%	**56%**
2 Hands, Coordinated, > 1 Group					
Drag 2 objects with pointer fingers (3b)	**63%**	**63%**	50%	**88%**	**56%**
Two hands grab points in sync	**88%**	50%	38%	**88%**	38%
Rhythmic use of both hands	**56%**	41%	50%	**63%**	25%
Both hands grab groups (3c)	**81%**	**34%**	38%	50%	25%
2 Hands, Coordinated, 1 Group					
Both hands coalesce large group to small (3f)	**75%**	9%	13%	13%	6%
Two-hand transport (3e)	**69%**	41%	38%	63%	31%
Add/remove from selection (3d)	25%	19%	0%	13%	31%

Table 3. Gestures that apply only to Surface condition

	Surface (by 1st Condition)		
	Mouse 1st (n=8)	Physical 1st (n=8)	Surface 1st (n=16)
Hull Resizing			
One Hand Hull Expand/Contract	13%	13%	25%
Two hand hull Tidy	0%	25%	6%
Two-hand hull Expand/Contract	25%	63%	56%
Expand hull to cover desired nodes (doesn't work)	13%	25%	6%
Other (failures)			
Treat finger like a mouse	50%	25%	38%
Push hard to multi-select	25%	13%	31%

both hands at once to affect a single group. However, both hands were often used to move groups separately.

We observed several habits from the other conditions that crept into the Surface interactions. For example, 56% of users tried to use their fingers as a mouse, experimenting using a different finger on the same hand for a "multi-select" or trying to draw marquees or lassos. Half of the users who started with the mouse continued to try mouse actions on the surface, while 25% of users who started with the physical condition tried mouse actions. More results are summarized in Table 3.

Table 4. Gestures that refer to the Physical condition

	Physical (n=16)	Surface (n=32)
Physical Gestures		
Lift Up	**75%**	3%
Go outside the lines	**69%**	0%
Slide *around* objects	**88%**	34%
"Texture"-based gestures (e.g. flattening a pile)	44%	3%
Toss items from one hand to other	38%	0%
Drag a handful, dropping some on the way	25%	6%

We wanted to understand what additional physical operations might be applied to a digital representation. In Table 4, we list operations that users performed in the physical condition that do not have a direct digital analogue. For example, 75% of all participants in the physical condition lifted the chips off the table; and 69% also pushed chips outside of the bounds of the table. Some of these gestures were attempted in the surface condition, but participants quickly realized that they were not supported on the surface. The one exception to this was a gesture to *slide objects around other objects* when moving them, which *was* possible in the surface condition although it was unnecessary since selected circles could be dragged *through* unselected circles.

5.2 Timing Results for Task 3

In addition to articulating the set of possible operations, we also wanted to understand which ways of moving multiple objects were most efficient. Do participants do better with the two-handed grouping operations of the surface, or the familiar mouse? We analyzed the task time data with a 2 (Condition) × 2 (Group) mixed ANOVA. Table 5 shows mean completion times with standard deviations for Task 3.

Table 5. Mean completion times (with std deviations) for Task 3 in seconds

Condition \ Order	MS (n=8)	PS (n=8)	SM (n=8)	SP (n=8)
Physical	-	71.0 (14.5)	-	107.6 (13.8)
Mouse	123.9 (30.9)	-	144.5 (32.5)	-
Surface	116.7 (21.8)	94.9 (30.3)	118.7 (31.2)	146.4 (37.5)

Surface is faster than Mouse. For the 16 participants who completed the Surface and Mouse conditions, we ran a 2 x 2 mixed ANOVA with *condition* {Surface, Mouse} as the within subjects variable and *order of conditions* as the between subjects variable. A significant main effect of condition was found ($F_{1,14}$=6.10, p=.027) with the surface condition being significantly faster (116 sec) than the mouse condition (134 sec). No significant effect of order was found ($F_{1,14}$=.928, p=.352) and there was no interaction effect between condition and order ($F_{1,14}$=1.38, p=.260).

Physical is faster than Surface, and trains users to be faster. For the 16 participants who completed the Surface and Physical conditions we again ran a 2 x 2 mixed ANOVA with *condition* {Surface, Physical} as the within subjects variable and *order* *of conditions* as the between subjects variable. A significant main effect of condition was found ($F_{1,14}$=11.96, p=.004) with the physical condition being significantly faster (89 sec) than the surface condition (120 sec). In addition, a significant effect of condition order was found ($F_{1,14}$=11.482, p<.001) where participants who started with the physical condition were significantly faster than the participants who started with the surface condition. No significant interaction effect was found between condition and order ($F_{1,14}$=0.655, p=.432).

Impact of First Condition. We hypothesized that users' performance on the surface would be impacted by whether they started with the Mouse condition first, or the Physical condition first. Two participants data were classified as outliers (> Average + 1.5 * SD). An independent samples t-test revealed that participants who performed the physical condition first were significantly faster on the surface, than participants who performed the mouse condition first (t_{12}=2.38, p=.035).

Number of Group Operations. In attempting to understand the time difference reported in the previous section, we found that the physical-surface (PS) group used more group operations than the mouse-surface (MS) group: an average of 33 group operations across participants in the PS group against 26 for the MS group. However, this difference was not significant enough to reject a null-hypothesis (t_{14}=0.904, p=.381). Of course multi-touch interaction on the surface affords a number of other types of interaction that may increase efficiency in such a clustering task, e.g., simultaneous selection with multiple fingers or independent "hand-over-hand" gestures.

5.3 User Comments

We asked participants to rate the difficulty of the clustering task on a 7 point Likert scale (1=Very difficult, 7=Very easy). We ran a 2 x 2 mixed ANOVA with *condition* as the within subjects variable and *order of conditions* as the between subjects variable. We found a significant main effect of condition ($F_{1,14}$=5.8, p=.03) with the surface condition being significantly easier (5.5) than the mouse condition (4.9). No significant effect was found between Physical and Surface.

Participants seemed to appreciate the manipulation possibilities of the Surface: when we asked which condition they preferred to perform the clustering task, 14 participants (88%) prefer Surface to Mouse. However, only 7 (44%) prefer Surface to Physical. Interestingly, the performance advantage of the Surface over the Mouse was greater than some participants thought. When we asked which condition they felt faster, only 9 participants (56%) felt Surface was faster than Mouse even though 12 (75%) actually did perform faster with Surface. However, 4 participants (25%) felt Surface was faster than Physical even though 3 (19%) were actually faster with Surface.

In verbal comments from participants who used both Physical and Surface, the most commonly cited advantage of Physical was the tactile feedback, i.e. selection feedback by feel rather than visual highlights. Whereas, the most cited advantage of the Surface was the ability to drag selected circles through any intervening circles instead of needing to make a path around them. For the participants who used both Mouse and Surface, the most cited advantage of the Mouse was multi-selecting many dispersed circles by control-clicking, while the most cited advantage of Surface was the ability to use two hands for parallel action.

6 Discussion and Conclusions

Tabletop multi-touch interfaces such as the Microsoft Surface present new opportunities and challenges for designers. Surface interaction may be more like manipulating objects in the real world than indirectly through a mouse interface, but it still has important differences from the real world, with its own advantages and disadvantages.

We observed that participants use a variety of two handed coordination. Some participants used two hands simultaneously, some used two hands in sync (hand over hand), some used coordinated hand-offs, and others used some combination of these. As a result, defining a group-by gesture requires some care because participants have different expectations about how grouping may be achieved when they first approach the Surface. In our particular implementation, participants sometimes had difficulty working with two hands independently and close together when our heuristic would make a group selection. We caution future designers of tabletop interfaces to consider this complexity in finding a good balance between physical metaphors and supporting gestures to invoke automation.

Multi-touch grouping turned out to be very useful. Many participants manipulated groups, and seemed to do so without thinking about it explicitly. Possibly the most valuable and common type of group manipulations were ephemeral operations such as the small open-handed grab and move. Massive group operations, such as moving large piles, also helped participants efficiently perform the clustering task. While our current implementation of group-select worked reasonably well as a baseline, we observed some difficulty with our hull system. We believe a better implementation of group select and increased user familiarity with multi-touch tabletop interfaces may bring user efficiency closer to what we observed in the Physical condition.

We have introduced a particular task that may be a useful benchmark for testing the efficiency and ergonomics of a particular type of basic tabletop interaction, but there is a great deal of scope for further studies. As was briefly mentioned in this paper, our study included a more challenging and creative task involving a layout of a network diagram. We intend to follow-up on this first exploration with an evaluation of a user-guided automatic layout interface that attempts to exploit the unique multi-touch capability of tabletop systems.

Acknowledgments. We would like to thank the participants of our user study for their participation and comments.

References

1. Forlines, C., Wigdor, D., Shen, C., Balakrishnan, R.: Direct-Touch vs. Mouse Input for Tabletop Displays. In: Proc. CHI 2007, pp. 647–656. ACM Press, New York (2007)
2. van Ham, F., Rogowitz, B.: Perceptual Organization in User-Generated Graph Layouts. IEEE Trans. Visualization and Computer Graphics (InfoVis 2008) 14(6), 1333–1339 (2008)
3. Hinckley, K., Baudisch, P., Ramos, G., Guimbretière, F.: Design and Analysis of Delimiters for Selection-action Pen Gesture Phrases in Scriboli. In: Proc. CHI 2005, pp. 451–460. ACM Press, New York (2005)
4. i2 – Analyst's Notebook, http://www.i2inc.com (accessed 29 January 2009)

5. Jong Jr., A.C., Landay, J.A., Rowe, L.A.: Implications for a Gesture Design Tool. In: Proc. CHI 1999, pp. 40–47. ACM Press, New York (1999)
6. Klemmer, S.R., Newman, M.W., Farrell, R., Bilezikjian, M., Landay, J.A.: The Designers' Outpost: A Tangible Interface for Collaborative Web Site Design. In: Proc. UIST 2001, pp. 1–10. ACM Press, New York (2001)
7. Leganchuk, A., Zhai, S., Buxton, W.: Manual and Cognitive Benefits of Two-Handed Input: An Experimental Study. ACM Transaction on Computer-Human Interaction 5(4), 326–359 (1998)
8. Microsoft Surface, `http://www.microsoft.com/surface`
9. Pedersen, E.R., McCall, K., Moran, T., Halasz, F.T.: An Electronic Whiteboard for Informal Workgroup Meetings. In: Proc. CHI 1993, pp. 391–398. ACM Press, New York (1993)
10. Proulx, P., Chien, L., Harper, R., Schroh, D., Kapler, T., Jonker, D., Wright, W.: nSpace and GeoTime: A VAST 2006 Case Study. IEEE Computer Graphics and Applications 27(5), 46–56 (2007)
11. Robinson, A.C.: Collaborative Synthesis of Visual Analytic Results. In: Proc. VAST 2008, pp. 61–74. IEEE Press, Los Alamitos (2008)
12. Stasko, J., Görg, C., Liu, Z.: Jigsaw: supporting investigative analysis through interactive visualization. Information Visualization 7(2), 118–132 (2008)
13. Thomas, J.J., Cook, K.A.: Illuminating the Path. IEEE Press, Los Alamitos (2005)
14. Tse, E., Greenberg, S., Shen, C., Forlines, C., Kodama, R.: Exploring true multi-user multimodal interaction over a digital table. In: Proc. DIS 2008, pp. 109–118. ACM Press, New York (2008)
15. Wilson, A.D., Izadi, S., Hilliges, O., Garcia-Mendoza, A., Kirk, D.: Bringing physics to the surface. In: Proc. UIST 2008, pp. 67–76. ACM Press, New York (2008)
16. Wise, J.A., Thomas, J.J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., Crow, V.: Visualizing the non-visual: spatial analysis and interaction with information from text documents. In: Proc. InfoVis 1995, pp. 51–58. IEEE Press, Los Alamitos (1995)
17. Wobbrock, J., Morris, M.R., Wilson, D.A.: User-Defined Gestures for Surface Computing. In: Proc. CHI 2009. ACM Press, New York (to appear, 2009)
18. Wu, M., Shen, C., Ryall, K., Forlines, C., Balakrishnan, R.: Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. In: Proc. TableTop 2006, pp. 185–192. IEEE Press, Los Alamitos (2006)

How Not to Become a Buffoon in Front of a Shop Window: A Solution Allowing Natural Head Movement for Interaction with a Public Display

Omar Mubin[1], Tatiana Lashina[2], and Evert van Loenen[2]

[1] Department of Industrial Design, Eindhoven University of Technology, The Netherlands
o.mubin@tue.nl
[2] User Experiences Group, Philips Research Eindhoven, The Netherlands
{tatiana.lashina,evert.van.loenen}@philips.com

Abstract. The user interaction solution described in this paper was developed in the context of an Intelligent Shop Window (ISW) with an aim to offer a user the interaction solution where system response would be triggered by naturally gazing at products. We have analyzed a possibility to realize such a user interaction solution using gaze tracking and concluded that remote calibration free eye tracking is still a subject of academic research, but that head tracking could be used instead. We argue that conventional use of head tracking requires conscious intentional head movements and thus does not fit into the context of applications such as the ISW. We further describe our experiment aimed to explore how head movements relate to eye movements when looking at objects in a shop window context. We show large variability in head movement and that per individual the gaze-head data could well be approximated with a straight line. Based on these results we propose a new solution that enables natural gaze interaction by means of head tracking.

Keywords: Gaze Interaction, Head Tracking, Augmented Reality Systems.

1 Introduction

Our current work originated from an exploration of new innovative interactive solutions for physical retail stores. In the last decade retail has undergone a significant transformation fuelled by the competition with on-line shops [1]. Retail brands grew into a blend of the old brick and mortar stores complemented by on-line services. This trend is also reflected in a shift towards experience, entertainment and themed retail, with the most known examples of NikeTown, and flagship stores of Apple, Nokia, and Prada. Physical retail was forced to appeal to shoppers in new ways to please the shopping motivations of modern consumers. In order to attract new customers to their shopping spaces retail frontrunners turn to different forms of interactive technologies. One that has been recently adopted by a number of leading stores is an interactive shop window display. These are interactive displays ranging from flat panel LCDs to holographic foil based rear projection screens [2, 3] integrated with a touch sensor like a capacitive touch screen [4] or an infrared optical sensor [5]. Leading brands like Ralph Lauren and Toyota have installed interactive shop window displays in their stores to make their shops accessible 24 hours a day [6, 7].

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 250–263, 2009.

To construct an elaborate view on the shop window context we conducted field studies in three downtown shopping centers and two shopping malls in Europe and United States. During these studies we observed and interviewed shoppers and retailers in order to understand the shop window from diverse standpoints. One of the results shows that the time most people spend at the shop window during shopping hours is very short and it varies from a short glance for duration of a second to a short stop for a minute or two. This time is too short to start browsing intricate information such as an electronic catalogue on the shop window. It has been demonstrated that due to the fact that public interactive displays are open to every onlooker it generally discourages people from trying to use them out of fear of doing something wrong or because it makes them look stupid [8]. One way to deal with both aspects of the shop window context mentioned above, is to use implicit interaction rather than explicit. Implicit interaction is one of the facets of augmented reality systems where the system can automatically trigger a particular reaction just by sensing and reasoning about user behavior. The advantage is that the user does not require any prior knowledge or training in order to interact with the system since no explicit intentional control actions are needed [9-11].

From the field studies we learned that the shop window is a visiting card of a shop designed to attract the right customers. That is why product items exhibited in the shop window are carefully selected eye-catchers that stand out from the rest of the assortment due to their novelty or originality. This motivated us to relate the content offered on the shop window display to the physical products exhibited inside the shop window. In our observation studies we, among other aspects, have observed how people attend to shop windows. Although we spotted people pointing at products, discussing products among each other and waiting for a partner at the shop window, mostly we saw people simply looking at products in the shop window. This triggered us to explore a possibility of detecting which product people look at in the shop window in order to implicitly trigger an interactive experience related to that product. We next describe the implemented concept where we have explored the possibilities offered by gaze interaction to shoppers at a shop window.

1.1 The Intelligent Shop Window Demonstrator

To explore the feasibility of the gaze interaction concept we have realized the demonstrator of the ISW in the shop area of our experimental facility. The shop area is part of a larger activity dedicated to realizing new innovations for retail and it is built to resemble a high end fashion store. In the ISW demonstrator, the shop window display is realized by using a holographic projection foil (i.e. Holo Screen [2]) that creates a semi-transparent display. Such display enables the user to shift easily the focus of visual attention between the products and the display (Figure 1). In the figure it is shown that a user is interacting with the shop window using touch based interaction. The window on the right is used for gaze interaction, where users can trigger selection of products with gaze. Details of how this is done are described next.

To build the gaze based implicit interaction for the shop window we needed a remote user independent gaze-tracker. The dilemma is that most commercially available gaze trackers are made for desktop use and thus operate on a short distance of about 60 cm. These systems require calibration and have poor tolerance to head motion.

Fig. 1. The Intelligent Shop Window (ISW) demonstrator in Philips Research Eindhoven

Remote user independent gaze tracking is the cutting edge of gaze interaction research with two main challenges: to eliminate the need of calibration per user and tolerate head motion [12-14]. Most advanced remote trackers do require a per-user calibration and only a few calibration free gaze trackers [15-18] have been shown to work in laboratory conditions but never in the field.

To start experimenting with gaze interaction in the context of the ISW we have acquired the Smart Eye 4-camera set-up [19] capable of detecting both head and eye gaze vectors. Smart Eye requires a calibration profile per user which is created by marking at least 12 feature points in at least 12 snapshots. Having installed the Smart Eye as part of the ISW we considerably stretched the boundaries of what the system could possibly deal with. In our set-up the cameras are placed behind the shop window and as a result there is a glass panel and a Holo Screen panel between the face of the user and the cameras. Part of the projected image reflects back into the cameras and could distort the tracking. Large distances between the person and the cameras of about 1.25m resulted in just a few pixels left for the pupil region in the image. Due to these difficulties, we could only achieve robust head tracking but not eye tracking. The experience with the Smart Eye triggered a question whether head tracking alone could be sufficient for a public interactive display such as the ISW.

Contrary to the challenges already mentioned for remote eye tracking, remote user independent head tracking is an affordable technology that has even been demonstrated in the context of interactive public displays [20-22]. These particular solutions focus primarily on the detection technology itself and not on its usability. In [20, 21, 23] head tracking is used in an explicit way, alike its traditional use in the field of assistive computing, i.e. head mice [24]. When head tracking is used in an explicit way the user needs to intentionally move her head in order to interact with the system. This is well accepted and even preferred in certain contexts [24], however, in a public context of a shop window people might feel that it is odd or unnatural to do due to the psychological barriers mentioned earlier [8]. Consequently, we defined our aim as to establish how head tracking could be used in an implicit way for interaction with the

ISW. To achieve this we had to figure out how head movements would relate to change of gaze in the context typical for a shop window. Therefore as an integral part of our paper we describe next the experiment that was carried out to determine the relationship between head and eye movements and the quantitative results that were found. Lastly we contemplate regarding the interaction solution that emerged from our results, which offers the tracking of gaze by employing head movement only.

2 Head Movement Experiment

2.1 Literature Overview and Our Approach

It has been widely accepted in literature that beyond a threshold of an angular viewing distance of 40 degrees head movements are mandatory [25, 26]. It has also been demonstrated that people develop different strategies as to what extent they involve head movement when looking at targets within the 40 degrees frontal view, which is referred to as head movement propensity [25]. All studies we reviewed demonstrated a large variability with respect to the head movement propensity. These differences gave birth to the classification of people as head movers and non-head movers. In [26] head movement was investigated for subjects whose head was free and a semicircular array of LEDs spaced 1 degree apart within a range of ±50 degrees was used to generate stimuli, e.g. illuminated LEDs. Again a great variability of head movement propensity was observed. This study reports an eye only region, which is a range of target eccentricities where head movement was not evoked by head non-movers. This region was observed to lie approximately within a range of ±20 degrees in the frontal view. It was revealed that outside the eye only region head movement linearly related to the amplitude of gaze shifts. Several studies [26-28] demonstrated an almost proportional relationship between gaze and head shifts within the ±20 degrees frontal view for head movers. It has been shown that the relationship between head movement and eye movement is influenced by experimental conditions [25, 26]. Therefore, we could not simply rely upon the results of the previous studies and had to experimentally show how head movements relate to gaze shifts in the context of the ISW.

Previous studies used very simple visual stimuli (i.e. lights) that suited the goals of their research [25, 26, 28] but it was disputable whether their results would hold in the context of the ISW. Our hypothesis was that the size and content of the target would have a bearing on how much head movement is carried out. As a result in our experiment we manipulated both product type and the size of the product. All the studies we found on head movement did not vary the distance between the user and the stimuli. Therefore we also included viewing distance as one of our independent variables. It was reported in head movement literature that vertical head movements have a smaller range than those of horizontal [29]. We did not find any estimate of that difference and therefore in our experimental design both horizontal and vertical conditions were included as independent variables.

2.2 Determining the Accuracy of the Smart Eye Tracking Device

Smart Eye claims a head pose accuracy (gaze vector originating from the nose in 3D) of a rotation of 0.5 degrees and a translation of 1 mm square [19]. In the context of the ISW, Smart Eye had to deal with challenging conditions, such as the glass and holographic foil, therefore it was imperative to estimate the precision of the Smart Eye in order to show whether it would be sufficient for the main experiment. In order to estimate the accuracy, we mounted an artificial head upon a movable tripod. A disk with an angular scale was mounted on top of the tripod and it allowed for up to one degree of radial movement. We also used a white sheet with equidistant targets marked upon it. In order to accurately orient the artificial head towards the targets, a laser pointer was attached on the top of the head, so that the laser beam would be parallel with the head orientation vector. The head could be moved vertically via adjusting the height of the tripod and horizontally by rotating it on top of the disk. Moreover, the artificial head was calibrated for the Smart Eye. We recorded Head Heading and Head Pitch generated by the Smart Eye tracking device. For each target we ensured the laser was aimed at that target and after the tripod and the disk were locked, data was recorded for 20 seconds. We performed the analysis of the data by computing an error distribution for each target. The error distribution was given by a distance matrix for each pose, which summarized the error differences. The analysis was based on the assumption that given that the head was not moving the readings would indicate no change over time. By normalizing the error distribution with a Gaussian function, we came to the conclusion that the Smart Eye had a precision error concentrated within 3 degrees. Hence we concluded that the precision of Smart Eye was acceptable to be used for the main experiment.

2.3 Method

Design. The aim of the experiment was to explore the relation between head movement and gaze changes, such that given head movement we could determine an approximation of the gaze. We recorded measurements for horizontal and vertical head movements (movement type) for different angular distances between the targets (angular distance), varied the distance between the subjects and the targets (viewing distance) and manipulated the size and the content of the targets (image type). The experiment was setup as a movement type (2) between subjects × screen viewing distance (3) × image type (3) × angular distance (5) within subjects design.

Participants. We recruited 21 participants in total. The gender profile and the allocation to the horizontal and vertical conditions are shown in Table 1.

Table 1. Number of participants per condition

	Participants	
	Male	**Female**
Horizontal	6	6
Vertical	5	4

Procedure. Participants were presented with a pair of targets simultaneously in the middle of a screen mounted inside the shop window. In order to notify the participants where to look first a counter was shown on the screen before the targets would appear. The counter appeared either on the left or the right of the screen (in the case of the horizontal condition), or on the top or the bottom (in the case of the vertical condition) along the center and it was placed at the position of the first target.

After displaying the counter, 3 intermittently appearing digits accompanied with 3 audio beeps, the counter would disappear and the two targets, T1 and T2, would appear simultaneously on the screen. We instructed the participants to look initially at the first target that would appear at the location of the counter and then at the second target (T2). The participants were specifically instructed not to look back and forth. We explicitly did not give any details as to how to look at targets rather we asked to look at the targets in the way they would naturally do. Both T1 and T2 were identical images of a particular object. The sequence of presentation of images was randomized in conjunction with the angular separation. Each pair of targets remained on the screen for 3 seconds. Between subsequent presentations of each pair of targets there was a gap of 5 seconds (2 seconds during which the screen was blank and the 3 seconds counter). The target presentation was repeated for a total of 3 (image type) X 5 (angular distance) X 6 (trials) = 90 times per each screen viewing distance. A typical experimental session with every participant lasted for slightly more than an hour.

In studies related to gaze interaction, time provided for fixation is an important factor. Fixation measurements in time normally for a scene with two targets or objects would consist of: the initial fixation on the first target (on average 750 ms); a scan time, being gaze shift from first target to first entry in the region of the next target (on average 100 ms); a final gaze fixation (on average 1000 ms) [30]. We had to be careful in allotting the time for each fixation, as it has been reported that longer fixations could encourage head movement [25]. We decided to increase slightly the window used to conduct fixation measurements up to 3 seconds, to ensure that all trials were successfully completed. The 3 second fixation period was confirmed in the pilot experiments. Each combination of within subject experimental variables (screen viewing distance, image type and angular distance) was tested in 6 trials. However, we manipulated the position of the counter in every set of 6 trials, such that for three trials the counter would appear in the left or top part of the screen and for the remaining three - in the right or bottom part, for the horizontal or vertical conditions accordingly. These left-right and top-down orientations of trials were randomized. Targets were presented in the centre of the screen with a distance separation that was computed upon the corresponding angular distance between the targets.

2.4 Materials

In order to display the stimuli a plasma screen (1.13 m by .65 m) was used which was suspended such that it was easily movable along a rail, hereby allowing us to manipulate its distance from the user. The position of every participant was measured prior to every viewing distance condition and the participants were instructed not to move during the condition (see Figure 2). The screen was mounted so that it would roughly lie in the center of the viewing fulcrum of a user standing in front of the ISW. For the vertical orientation condition, the same screen was rotated 90 degrees anticlockwise. All tracking was carried out by the Smart Eye device.

Fig. 2. Participants undergoing the experiment in both conditions: left-horizontal and right-vertical

2.5 Independent Variables

The image type factor had three levels conventionally coded as *small*, *medium* and *large*. *Small* corresponded to the image of an mp3 player, *medium* was the image of a perfume bottle and *large* - the image of a bag (see Figure 3). We used images of different products, instead of scaled versions of the same product to increase ecological validity and to prevent the participants from feeling bored or even withdraw from the experiment.

Fig. 3. Type of products: mp3 player, bottle and bag

The screen viewing distance factor was also defined according to the three levels: *small*, *medium* and *large*. The observation studies mentioned in the introduction provided an approximate estimate of typical viewing distances: 0.5m to view smaller products and 1.5m for larger products. Due to a restriction of our cameras set-up, we could not test the distance of 0.5 m. Therefore we tested the three distances of 1.0 m, 1.25 m, and 1.5 m for *small*, *medium* and *large* conditions respectively. The distance was measured from the centre of the feet of the participants to the display screen. In order to counter balance any effects of fatigue or boredom over the three viewing distances, the order of the presentation of the screen viewing distances was varied between subjects. In total we had 6 possible orders, where each order was executed for two participants, where possible. Between each of the three sessions, participants were able to take a break, while the viewing distance was adjusted.

From literature [26] we learned that a ±20 degrees window was identified as a so-called eye only region and head movements beyond this range would become mandatory. This range is also typical for viewing products in a shop window context. We chose the following five angular distances between images, expressed in degrees: 10, 15, 20, 25 and 30. Since we could not reliably track gaze with the Smart Eye we could only rely on our subjects shifting their gaze in a way that corresponded to the angular distances we had defined. To ensure this we carefully instructed the participants on the sequence they had to follow when looking at the targets.

2.6 Measurements

We carried out several pilot studies in order to test our experimental setup and check whether the data output corresponded to what we expected to happen. The Smart Eye outputs numerous parameters per frame. The most relevant parameters for our case were the angular direction of the head pose vector (expressed in radians) in both horizontal and vertical directions. They are termed as Head Heading and Head Pitch respectively. In order to convert the head angular measurements into head movement shifts we first plotted the Head Heading data as shown in Figure 4. As it can be seen there are three distinct regions. These three regions represent the fixations on the two targets (horizontal regions 1 and 3), and the transition between the two fixations (central region 2). In order to obtain the estimate of the head movement we were interested in the distance between the two horizontal parts of the curve.

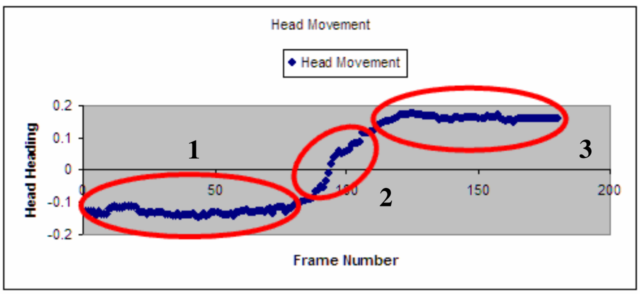


Fig. 4. Head Heading plot over a trial for a left to right head movement

To compute the head movement shifts, we used a simple k-means clustering algorithm, which was executed offline in MATLAB, after each experiment. By dividing the data of a trial into three clusters, we could subtract the centroids of the two most populated clusters to get the resulting head movement for each and every trial. The final value for head movement incurred for a particular combination of angular distance, image type and viewing distance was the median of the 6 trials. The median was used to account for outliers, if and when the tracking was not flawless.

3 Results

We shall present our results in a threefold manner. We evaluate the two between subject orientation conditions separately (horizontal and vertical) and conclude with a final comparison of both movement types. All statistical analysis was done on the median head movement values per each combination of independent variables.

3.1 Horizontal Head Movement: ANOVA Main Effects

The average horizontal head movement per participant across all the experimental conditions is visualized in Figure 5. Our first research question was whether the viewing distance, image type and the angular distance had an effect on the horizontal head movement. We performed a 3 by 3 by 5 repeated measures within-subjects analysis of variance (ANOVA) on the data from 12 subjects with gender and the screen viewing distance order as between subject factors. The repeated-measures ANOVA revealed significant difference in head movement for different angular distance conditions ($\underline{F}$ (4, 16) = 33.04, $\underline{p} < .001$). The image type did not reveal a significant influence head movement ($\underline{F}$ (2, 8) = 1.63, $\underline{p} < .26$). The same was true for the viewing distance ($\underline{F}$ (2, 8) = .863, $\underline{p} < .46$). Surprisingly, gender appeared to have an effect ($\underline{F}$ (1, 4) = 10.88, $\underline{p} < .03$). Females were observed to carry out almost twice as much horizontal head movement as males. Moreover, we found a significant interaction effect between gender and angular distance, ($\underline{F}$ (4, 16) = 5.06, $\underline{p} < .008$). Screen viewing distance order did not have an effect ($\underline{F}$ (5, 4) = 4.315, $\underline{p} < .09$).

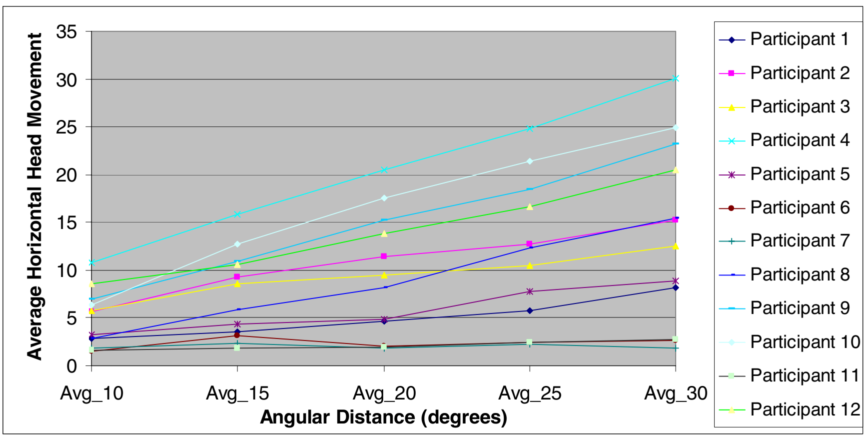


Fig. 5. Average horizontal head movement per participant

3.2 Vertical Head Movement: ANOVA Main Effects

The average vertical head movement per participant across all the experimental conditions is visualized in Figure 6. Similar to the horizontal movement condition we carried out a within-subjects repeated measures ANOVA. The results of the repeated-measures ANOVA turned out to be similar to the results that we had computed previously, for the horizontal movement condition. We found out that different angular distance conditions were significantly different with respect to the amount of vertical head movement ($\underline{F}$ (4, 28) = 5.69, $\underline{p} < .002$). The screen viewing distance did not have a significant effect ($\underline{F}$ (2, 14) = .390, $\underline{p} < .68$). No effect was found for image type ($\underline{F}$ (2, 14) = .12, $\underline{p} < .89$) as well. However, contrary to our results for horizontal head movement, gender did not have a significant effect on vertical head movement ($\underline{F}$ (1, 7) = .011, $\underline{p} < .92$). Screen viewing distance order did not show an effect either

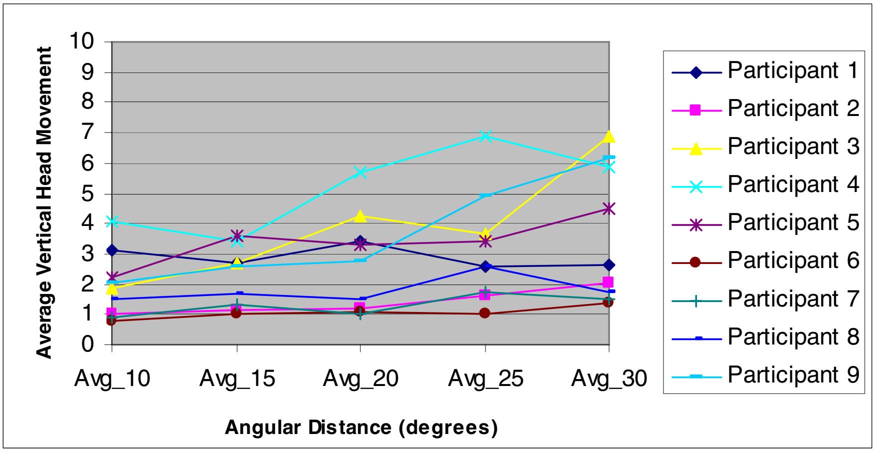


Fig. 6. Average vertical head movement per participant

(F (5, 1) = 25.14, p < 0.150). There was no interaction effect found for the vertical condition. Prior to conducting the experiment we suspected the height of a participant to have an effect on the vertical head movement. This was mainly due to the possibility that the starting point of fixation would vary from person to person due to the height, which could in turn influence the net vertical head movement carried out. Therefore the height of a user was also taken into consideration as a between subject factor in the ANOVA analysis. However, the ANOVA results did not reveal any significant effect of the height ($\underline{F}$ (1, 6) = .289, $\underline{p}$ < .61).

3.3 Comparison between Horizontal and Vertical Head Movements

As stated previously, in earlier research it has been suggested that the vertical head movements has a relatively smaller range than the horizontal. We did not come across any studies that would quantitatively compare the two. In our experiment we collected head movement data in both horizontal and vertical conditions. To estimate the difference we compared the overall average head movement among the horizontal and vertical conditions. A two samples independent t-test was executed to compare the overall means for horizontal and vertical head movement conditions. The difference turned out to be significant ($\underline{t}$ (12.54) = 3.6, $\underline{p}$ = .003). Secondly, to estimate the relative difference we compared the median head movement in the horizontal condition (Mdn_H=9.13) with the median head movement in the vertical condition (Mdn_V=2.88). We concluded that the vertical head movement was on average 3.17 times less than the horizontal.

4 Discussion

In both horizontal and vertical conditions we did not observe any effect of the product type and the viewing distance on head movement. These results are important since previous studies did not investigate the effects of these two factors. For the shop window interactive applications these results mean that within the typical range of viewing distances head movement that accompanies natural gaze behavior is not effected by the distance towards the products or even by product type. With regard to the screen viewing distance we tested the range of distances typical for shop window situations (1, 1.25, 1.5 m) but the range could be different in other contexts. A limitation we had is that the smaller distance of 0.5 m could not be tested due to the physical arrangement of the apparatus. The viewing distance effect should be further investigated including the 0.5m distance.

In previous research it has been demonstrated that people could be categorized as head movers and head non-movers [28]. It has also been shown that any individual could be positioned on a continuum between extreme head movers and head non-movers [26]. Our results are in line with previous research and we observed a large variability in head movement propensity. Our results also show that both horizontal and vertical head movement had a positive increase for larger angular distances between products. Upon visualization of the results we showed that per individual data could be well approximated with a straight line. Similar linear dependency has been suggested previously in [26, 28] but was not investigated intricately.

Gender proved to be an interesting factor with regard to horizontal head movement. It is generally known that females demonstrate a larger variability for certain cognitive tasks that are experimentally related to natural hormonal changes. In contrast, gender has not been reported to influence head movement in earlier studies. One possibility could be that the content of the images or the context of the shop window was more appealing for females. Or it could be that the case that the gender effect was observed due to a small sample size. The gender effect needs a separate study to be verified. The results obtained triggered ideas for bringing them into practice for the ISW and next we provide a description of a new gaze interaction solution for the shop window.

5 User Interaction Solution

In line with previous research [25, 26, 28] we have observed a large variability in the head movement propensity (Figure 5 and Figure 6). It can be concluded that only a small fraction of people align their heads with their gaze targets. In other words in case of conventional head tracking as described in [20, 21], where head headings are translated into gaze headings identically for all users, the majority of people would have to intentionally exaggerate their head movement to be able to interact with such systems. In order to control such a system the user needs to be trained beforehand in order to understand how the system is controlled.

Another key result is that for every participant head movement observations for different angular target distances approximate well upon a straight line. To deduce this linear head-eye relationship it is sufficient to know at least one single head displacement between two widely spaced products. Using this single displacement and a priori statistics collected on head movement behavior it is possible to select the best fit line that most closely approximates the head movement of the individual. As we have demonstrated experimentally vertical head movement was on average 3.17 times less than horizontal. Therefore the vertical component for head movement can be calculated by simply dividing the horizontal by 3.17. Here we make an assumption that the amount of head movement for an individual is proportionally equal in both horizontal and vertical conditions. Since we tested the two movement types between subjects our afore-mentioned claim would need to be substantiated.

Ideally the proposed system calibration should occur in the background, hidden from the user. To achieve this we propose a method of attracting the user's attention to the targets in such a way that it would not be immediately obvious to the user. In visual perception it has been demonstrated that a dynamic visual effect on a static background is among the most suitable ways to attract human attention [31]. In our design we suggest to use a dynamic colored light effect to highlight one after the other two widely spaced products in a shop window. In this way the user's attention could be drawn first to the first product and then to the second product. In parallel to attracting the user's attention the system measures the corresponding head vector displacement and calculates the approximate gain factor for translating the head movement measurements into the gaze vector.

The interaction solution as described above enables natural and untrained interaction for public interactive displays where instructing users would neither be possible nor desirable. In case the user is aware that the system is controlled by head movement, the proposed interaction solution offers a benefit of deploying natural head movement without the need of intentional or exaggerated head movement.

6 Future Research and Conclusion

Currently we are in the process of implementing the described interaction solution as part of the ISW in order to be able to test the feasibility and usability in a laboratory setting. As we progress with our demonstrator we need to address a number of technical challenges. One of them is the difficulty the system will have with extreme head non-movers. It still needs to be determined what the minimum head movement angle should be in order to be sufficient for the system to distinguish between gazing at different targets. A gender effect that was observed should be definitely further investigated. Head movement gender differences might lead to new gaze interaction solutions for retail. The interaction solution as described above should be applicable to contexts beyond the ISW, e.g. product displays inside a shop, museum exhibition stands, and other public interactive displays. How our interaction solution would fit these new contexts needs further exploration.

In this study, we aimed to find a method to robustly estimate a user's gaze direction when she is looking at objects in a shopping window, without having to build a profile of the user's face in advance as is typically required for state of the art gaze trackers. User independent head tracking technology is already available, but it does not provide a gaze direction. To solve this dilemma, we explored the relation between head and gaze movement when viewing products of typical sizes at typical distances and at relative positions found in shop windows. For horizontal directions, we found that the amount of head movement for a given gaze shift differed greatly between participants. However, for each individual, the relation between head shifts and angular gaze shifts was found to be close to linear for the angular ranges studied

(10-30 degrees). No significant influence of object size or object distance was found. However, for the vertical direction head movements were about 3 times smaller than for horizontal movements. From these observations we derived a gaze tracking method, such that a user's gaze direction is estimated from her head movement. This interaction solution allows users to look at objects naturally, without the need to move the head intentionally which would exaggerate the movement amplitude.

References

1. Harvard Design School Guide to Shopping. Harvard University Graduate School of Design
2. DNP Holo Screen holographic rear projection screen, `http://www.en.dnp.dk/`
3. Pronova HoloPro screen, `http://www.gb-pronova.de/en/index/`
4. 3M MicroTouch™ ClearTek™ II Capacitive Touch Screens, `http://solutions.3m.co.uk/wps/portal/3M/en_GB/TouchScreens/`
5. NextWindow optical touch screen technology, `http://www.nextwindow.com/benefits/touchscreen_technology.html`
6. Ralph Lauren kicks off interactive e-commerce at Rugby.com, on Internet Retailer, `http://www.internetretailer.com/article.asp?id=27601`
7. Interactive Toyota ShowRoom on YDreams, `http://www.ydreams.com/ydreams_2005/index.php?page=372`
8. Brignull, H., Rogers, Y.: Enticing People to Interact with Large Public Displays in Public Spaces. In: Proc. INTERACT 2003, pp. 17–24 (2003)
9. Bellotti, V., Back, M., Edwards, W., Grinter, R., Henderson, A., Lopes, C.: Making sense of sensing systems: five questions for designers and researchers. In: Conference on Human Factors in Computing Systems (CHI 2002), pp. 415–422. ACM Press, New York (2002)
10. Dietz, P., Raskar, R., Booth, S., van Baar, J., Wittenburg, K., Knep, B.: Multi-projectors and implicit interaction in persuasive public displays. In: AVI: Working Conference on Advanced Visual Inter-face, pp. 209–217. ACM Press, New York (2004)
11. Sukaviriya, N., Podlaseck, M., Kjeldsen, R., Levas, A., Pingali, G., Pinhanez, C.: Embedding Interactions in a Retail Store Environment: The Design and Lessons Learned. In: INTERACT 2003. LNCS, pp. 519–526. Springer, Heidelberg (2003)
12. Duchowski, A.: A Breadth-First Survey of Eye-Tracking Applications. Behavior Research Methods Instruments and Computers 34, 455–470 (2002)
13. Duchowski, A.: Eye Tracking Methodology: Theory and Practice. Springer, Heidelberg (2003)
14. Morimoto, C., Mimica, M.: Eye gaze tracking techniques for interactive applications. Computer Vision and Image Understanding 98, 4–24 (2005)
15. Morimoto, C., Amir, A., Flickner, M.: Free head motion eye gaze tracking without calibration. In: Proc. of Conference on Human Factors in Computing Systems (CHI), pp. 586–587. ACM Press, New York (2002)
16. Morimoto, C., Amir, A., Flickner, M.: Detecting Eye Position and Gaze from a Single Camera and 2 Light Sources. In: Proc. of 16th International Conference on Pattern Recognition, pp. 314–317 (2002)
17. Shih, S., Wu, Y., Liu, J.: A Calibration-Free Gaze Tracking Technique. In: Proc. of the conference on Pattern Recognition, pp. 201–204 (2000)
18. Yoo, D., Kim, J., Lee, B., Chung, M.: Non-contact eye gaze tracking system by mapping of cornealreflections. In: Proc of AFGR 2002, pp. 94–99 (2002)
19. Smart Eye gaze tracker, `http://www.smarteye.se/smarteyepro.html`

20. Chang, C., Chung, P., Yeh, Y., Yang, J.: An Intelligent Bulletin Board System with Real-time Vision-based Interaction Using Head Pose Estimation. In: Proc. of the 18th International Conference on Pattern Recognition, Washington DC, USA, pp. 1140–1143. IEEE Computer Society Press, Los Alamitos (2006)
21. Liu, X., Krahnstoever, N.O., Yu, T., Tu, P.H.: What are customers looking at? In: Proc. of IEEE International Conference on Advanced Video and Signal-Based Surveillance (2007)
22. Nakanishi, Y., Fujii, T., Kiatjima, K., Sato, Y., Koike, H.: Vision-Based Face Tracking System for Large Displays. In: International Conference on Ubiquitous Computing. LNCS, pp. 152–159. Springer, Heidelberg (2002)
23. Matsumoto, Y., Zelinsky, A.: An algorithm for real-time stereo vision implementation of headpose and gaze direction measurement. In: Proc. of 2000 IEEE conference on Automatic Face and Gesture Recognition, pp. 499–504 (2000)
24. Bates, R., Istance, H.: Why are eye mice unpopular? A detailed comparison of head and eye controlled assistive technology pointing devices. Universal Access in the Information Society 2, 280–290 (2003)
25. Fuller, J.: Head movement propensity. Experimental Brain Research 92, 152–164 (1992)
26. Stahl, J.: Amplitude of human head movements associated with horizontal saccades. Experimental Brain Research 126, 41–54 (1999)
27. Kuzuoka, H.: Spatial workspace collaboration: a SharedView video support system for remote collaboration capability. In: SIGCHI conference on Human factors in computing systems, pp. 533–540. ACM Press, New York (1992)
28. Bard, C., Fleury, M., Paillard, J.: Different patterns in aiming accuracy for head-movers and non-head movers. In: The head-neck sensory motor system, pp. 582–586. Oxford University Press, Oxford (1992)
29. Pelz, J., Hayhoe, M., Loeber, R.: The coordination of eye, head, and hand movements in a natural task. Experimental Brain Research 139, 266–277 (2001)
30. Henderson, J.: Human gaze control during real-world scene perception. Trends in Cognitive Sciences 7, 498–504 (2003)
31. Sutcliffe, A.: Multimedia and Virtual Reality: Designing Multisensory User Interfaces. Lawrence Erlbaum Associates, Mahwah (2003)

Chucking: A One-Handed Document Sharing Technique

Nabeel Hassan, Md. Mahfuzur Rahman, Pourang Irani, and Peter Graham

Computer Science Department, University of Manitoba
Winnipeg, R3T 2N2, Canada
nhassan@obsglobal.com, {mahfuz,irani,pgraham}@cs.umanitoba.ca

Abstract. Usage patterns of private mobile devices are constantly evolving. For example, researchers have recently found that mobile users prefer using their devices with only one hand. Furthermore, current hardware in these devices reduces the need for a stylus and instead relies on finger input. However, current interactive techniques, such as those used for sharing documents between private and public devices have not taken advantage of these recent developments. For example a popular technique, Flick for sharing documents between devices relies on pen and stylus use and has not been adapted to support one-handed interaction. In this paper, we present Chucking, a gesture-based technique for sharing documents between private mobile devices and public displays. Chucking is based on the natural human gesture used for throwing or passing objects. We present the various design parameters that make Chucking an effective document sharing technique. In a document positioning task, we evaluated Chucking against Flicking. Our results show that under certain contexts users were more accurate and effective with Chucking. Participants also preferred Chucking as it maps closely the type of interaction one naturally performs to share objects. We also introduce extensions to Chucking, such as Chuck-Back, Chuck-and-Rotate, and Chuck-and-Place that constitute a suite of techniques that facilitate a large range of document sharing interactions between private mobile devices and public displays.

Keywords: Chucking, Flicking, public-to-private document sharing, multi-document environments (MDEs).

1 Introduction

We are witnessing the introduction of public displays in numerous environments, such as in schools, airports, museums, and shopping centers. The recent proliferation of such displays has led to the establishment of multi-display environments (MDEs) [1,2,3,14] in which several private devices (PDAs, cell-phones, etc) can be used to interact with information available on public displays. Recently, there has been growing interest in determining ways to bridge the interactions that couple together private and public displays [12,22,25]. To this end, one common application that has emerged is that of sharing documents between devices. This includes scenarios such as an instructor sharing photographs from a field trip with a classroom, a group of users sharing their private documents in a boardroom during a meeting or organization members placing information on a bulletin board like display.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 264–278, 2009.

Researchers have developed and studied document sharing techniques across multiple devices and platforms [5,12,21]. While such techniques have shown significant advantages in specific usage contexts, they require the use of a stylus or pen to facilitate the interaction, a feature that is becoming less common on platforms such as the iPhone and smart phones. Furthermore, current designs do not easily support one-handed interactions, a mode of operation that is highly favored when using mobile handheld devices [9,11]. Generally, for the task of sharing documents from a private device to a public display, no current method is able to satisfy all of the following design goals:

- *Stylus independence:* on multiple devices, styli operation is non existent or limited. Document sharing should not be inhibited by the lack of external styli.
- *One handedness:* studies reveal that 74% of mobile device users use one hand when interacting with their devices [9,10]. New mobile applications should support the prevalent use of one-handed operations.
- *Natural and fluid:* techniques that map to natural ways of working are easier to learn so any new design should be as natural as possible.
- *Position independence:* one-handed interaction suffers from a non-uniform distribution of thumb reach. Limitations of thumb reach should not prevent a technique from operating uniformly well across different conditions.

We propose Chucking, a one-handed gesture-based interaction for sharing documents from a private mobile device to a public display (Figure 1). Chucking was designed to alleviate some of the limitations of existing techniques for one-handed use on devices that do not have a stylus. In a document positioning task, users were more accurate with Chucking than one-handed Flicking. Furthermore, participants preferred Chucking to Flicking and found the technique natural to learn and use. We also introduce a suite of techniques based on Chucking, including Chuck-Back for pulling documents from the public display to the private device, Chuck-and-Rotate to rotate a chucked document to accommodate orientation issues on tabletops, and Chuck-and-Place, a technique to allow for more accurate document placement.

(a) (b) (c)

Fig. 1. From (a) a steady state, (b) the user Chucks a document to the public display (c) with an extension of the forearm or wrist. The gesture is analogous to dealing cards like a dealer.

2 Related Work

We present the related work on one-handed interactions, on sharing documents in multi-display environments, and on tilt-based interactions.

2.1 One-Handed Interaction

One-handed interaction is a highly popular method for interacting with mobile devices. Karlson and Berderson, conducted several in-situ observations which concluded that 74% of mobile users employ one hand when interacting with their cellular devices [10]. The same observations and web surveys suggested that PDA users were inclined to more often use two hands, but users of such devices also expressed significant interest in using one hand if possible, suggesting the need for better tools to support one-handed interactions.

Applens and LaunchTile were two of the earliest systems to support one-handed thumb use on PDAs and cell-phones [11]. With Applens, users were provided a simple gesture set, mimicking the up-down-left-right keys, to navigate a grid of data values. LaunchTile provided access into the tabular data view by allowing users to press on soft buttons associated with an area of the grid. In a user study, users performed correct gestures 87% of the time, suggesting that simple thumb gestures are memorable. However, they also found that users were reluctant to use gestures and preferred tapping. Their results also showed that error rates were also influenced by the direction of the gestures and the number of gestures available, suggesting a limit on the number of different gestures one should design.

Further investigation by Karlson et al [10] on the biomechanical limitations of one-handed thumb use has revealed that users do not interact with all areas of a device with equal facility. Instead, user grip, hand size, device ergonomics, and finger dexterity can influence movement and thumb reach. For example, right handed users have limited thumb movement in the NorthEast-SouthWest direction. Additionally, regions of the device away from the right edge are more difficult to reach. These findings supported the development of Thumbspace [9]. To facilitate thumb reach with Thumbspace, users customize and shrink the entire workspace into a 'thumb-reachable' bounded box. In an extensive study, users performed better at selecting targets that would be further to reach, using Thumbspace than other techniques. Used in conjunction with Shift [23], a technique for high-precision target selection with the finger, Thumbspace resulted in higher accuracy than using either technique alone.

Overall, the results of these studies suggest that while there is sufficient evidence of one-handed use with mobile devices, we do not have a broad range of techniques to support such usage. Furthermore, the design of one-handed interactions needs to be concerned with the size of object, the placement and position of control items, and the range of allowable gestures with the thumb. We considered these limitations in the design of Chucking.

2.2 Document Sharing Techniques

Researchers have proposed a number of techniques to support direct document sharing between private and public devices [21,27,12,1,2]. Pick-and-drop [21] is a technique that allows a user to tap on an object with a digital pen to 'pick it up' and then to 'drop' the object in another location on the workspace. However, pick-and-drop does not support very well cases when the user is interacting with a very large display unless the user can directly reach the desired object. To resolve this 'out-of-reach' dilemma, researchers have proposed proxy based techniques to allow the user to interact with distant content. In the vacuum filter [2] and drag-and-pop [1] techniques, users pull in proxies of remote content, with a brief gesture of the stylus. In this manner, users can either select distant objects or drag-and-drop items without carrying out large arm movements or displacing themselves physically. However, with both the vacuum filter and drag-and-pop approaches, users are required to be in close proximity to the large display, a situation that is inappropriate in certain contexts such as boardroom meetings.

Researchers have also proposed document sharing techniques that do not require strict proximity with the display [19]. Most techniques in this group have investigated the use of human gestures for resolving issues with distance [3,5,27]. The use of human gestures in interactive systems has a natural appeal as designers believe that such gestures are analogous to how we interact with physical objects in the real world. Additionally, gestures can be efficient if they are designed to integrate both command and operand in a single motion, to conserve space on devices, and to reduce the need for buttons and menus [9].

Toss-it is a gesture-based technique to facilitate data transfer between mobile devices [27]. With this technique, users can 'toss' information from one PDA to another with a swinging action, analogous to pitching a softball. Based on the 'strength' of the action and the layout of mobile devices in the environment, toss-it selects the device that is most suited to receive the data. Toss-it inspired the design of Chuck. However, toss-it allows users only a limited number of gestures. Furthermore, it has not been designed to work with large public displays, thus issues such as positioning or orientation of documents are not easily resolved using toss-it. Finally, toss-it does not exploit the full range of gestures possible by the human hand.

Researchers have also proposed techniques that are slightly more natural than toss-it. Geiβler's throw technique [3] requires the user to make a short stroke over a document, in the direction opposite to the intended target, followed by a long stroke in the direction of the target. Wu et al. [26] describe a 'flick and catch' technique, in which an object is 'thrown' once it is dragged up to a certain speed. Reetz et al. [19] demonstrated the benefits of Flicking as a method for passing documents over large surfaces. Flicking was designed to mimic the action of sliding documents over a table, and closely resembles the push-and-throw model designed by Hascoet [5]. Flicking was found to be much faster than other document passing techniques for tabletop systems [13] but incurred a noticeable cost in terms of accuracy. Superflick [19] was designed to improve the accuracy of flick by allowing the user to adjust the position of the object on the shared display. Superflick performed with equal accuracy and as efficiently as radar, a technique that provides an overview of the public space on the user's private device [19]. While radar is a fairly robust technique, the overview space it covers would be impractical on private devices such as mobile displays, where issues such as occlusion and small targets would make radar unusable.

Flicking has numerous advantages for various interactions. However this technique has not been evaluated under conditions of mobility where the usage is one-handed and on devices without a stylus. Furthermore, most of the document sharing techniques discussed above are not designed for bidirectional transfer (i.e. back and forth). We present variants of chucking that allow for bidirectional transfer, for accurate positioning and for rotating objects on horizontal public displays (section 7).

2.3 Tilt-Based Techniques

Chucking builds on tilt-based techniques. One of the earliest uses of tilt input was proposed by Rekimoto [20] for invoking an analog input stream where tilt could be used to invoke menus, interact with scroll bars, or to browse maps. One particularly appealing feature of such an interaction, as noted by Reikimoto, was the use of only one hand to both, hold and operate the device. Since Rekimoto's proposal, a significant number of tilt-based techniques have emerged. Recent work in tilt interfaces can be grouped into rate-based or position-based tilting. Hinckley et al. [6] demonstrated the use of accelerometer enabled tilting for automatic screen orientation and scrolling applications, a now common feature on commercial devices.

Numerous systems have defined a fixed mapping from tilt position to a function in the workspace. Oakley and Park [17] described a tilt-based system with tactile augmentation for menu navigation. Each menu item was selectable by tilting the device at a fixed angle. TiltText [24] allows text entry by tilting the device to the correct position to enter a letter. Results have shown that this form of interaction can improve text entry speed but at the cost of errors. A study by Oakley and O'Mondrain [16] has shown that position-based mapping is more accurate than rate-based mapping, presumably due to the feedback obtained based on the position of the arm. None of the above described systems have fully explored tilt for one-handed document sharing.

3 Design Framework

We propose a framework for chuck-based interactions. This framework highlights five primary factors that influence performance with chucking: gesture commands, mapping function, feedback, invocation and simultaneous touch input.

3.1 Gesture Commands

Chucking is a one-handed document sharing interaction. To invoke a Chuck, the user selects an item with the thumb or other input (e.g. jog dial) and then extends the forearm or wrist as in chucking cards on a table. The use of a hand gesture alleviates the need for a stylus or pen. To recognize the gesture, we use the values from tilt sensors, a common feature on many devices such as the iPhone or smartphones. With 2D tilt sensors, we are capable of allowing users a more flexible range of gestures that are natural to the concept of throwing to share documents [5].

3.2 Mapping Function

In Chucking we use a mapping function to position the document in a given location on the remote display. Existing proposals for tilt-based interactions have suggested a mapping using either the rate of tilt or the angle of tilt to control a virtual cursor. In a rate-based tilt design, the speed of the cursor is mapped to the degree of tilt. Designers have reported that rate-based systems are difficult to control and do not provide the level of precision possible with position-based control. In a study, Oakley and O'Modharin [16] found that users were more efficient and more accurate controlling items on a menu with position-control than rate-control tilting. They observed that a primary benefit of position-based tilt control is the lack of reliance on any hidden virtual model. Additionally, position-based tilt reinforces feedback with the physical orientation of the hand. We designed Chucking based on position-control mapping instead of rate-based mapping. Several variants of position-based mappings exist. An absolute position mapping captures the absolute tilt value and uses this to identify the direction of a user's chuck. Alternatively, we can also use a relative position mapping, which maps the amount of tilt to the position of the document on the remote display. We applied a mapping in which shorter differences in orientation are mapped to closer locations on the remote display, while larger movements are mapped to locations that are further away from the user.

3.3 Feedback

A feedback loop is a necessary component of many interactive techniques. In the case of tilt, the user can infer a certain degree of feedback from the position of the hand. However, such a feedback mechanism is insufficient if the interaction technique necessitates a large number of unique tilt gestures. Visual feedback is a common feedback method for most tilt-based techniques [16,17,20,24]. With visual feedback, the tilt techniques are restricted to the range of motion such that the feedback signal is not inhibited. Auditory feedback can also complement visual feedback for tilting [17]. However, this form of feedback is limited based on the context of its use and the specific techniques employed. Oakley and Park [17] presented the case for motion based interaction without any visual feedback but with limited tactile feedback. To evaluate if their systems could be used eyes-free, they gave participants very minimal training with their menu system. Their results showed that performance improved with visual feedback and accuracy was unaffected in the eyes-free condition. Based on prior work, complete visual feedback, in which clear visual indication suggests when the transfer has occurred, would benefit most Chucking based interactions. However, augmenting this form of feedback with audio also enhances the interaction and demarcates the completion of the task. We use both visual and auditory feedback. Auditory feedback consisted of a 'click' sound when the transfer is successful.

3.4 Invocation

Since chucking is based on tilt interactions, the system needs to disambiguate typical non-chuck gestures, with those that are intentionally created for the purpose of sharing documents. One method for invoking the gesture is to simply click anywhere on the device, and then to start the gesture. Alternatively on devices that are not touch-based, the user can press a button or the jog dial to select the item to transfer.

3.5 Sequential Touch Input

Since Chucking relieves the fingers from performing any major work, we can take advantage of this and combine the use of fingers in the interaction. For document positioning two possible configurations are available. By embedding a touch mechanism with Chucking, users can place their finger in one of several locations before sending the document (i.e. in conjunction with the velocity vector) to fine tune adjustments such as in Superflick [19]. Alternatively, the finger can be used for orienting the document on a horizontal display. The use of a finger position on the private display can suggest the orientation the document may take once it is placed on the horizontal surface. We discuss these details in section 7 when presenting Chuck-and-Rotate, and Chuck-and-Place.

4 Data Capture

We first captured data points, consisting of angular tilt in the X and Y dimensions of the 2D tilt sensor, to identify distinct sets of gestures for Chucking. We used a Dell AximX30 PDA with a TiltControl accelerometer attached to it and polled the X and Y tilt angles of the TiltControl at 40 msec intervals. The application was written in C#.NET. Four right-handed volunteers performed a set of gestures. All participants were university students; three males and one female (average age of 23.25 years). All had experience using PDAs but none had any exposure to the Chucking technique.

4.1 Method for Capturing the Unique Gestures

Participants were asked to perform the Chucking gesture as if they would be tossing an object in one of nine different locations on a grid: Right-High, Right-Medium, Right-Low, Center-High, Center-Medium, Center-Low, Left-High, Left-Medium and Left-Low. We limited the data capture to only nine locations as pilot studies revealed significant overlap of tilt values for anything larger than a 3×3 grid. Additionally, the basic Chucking technique is suitable for sending documents in a general region, and we provide refined control for accurate positioning with Chuck-and-Place (section 7).

After an initial demonstration, the participants were asked to chuck an imaginary item on the PDA for a total of five trials for each condition. The experiment collected a total 180 trials as follows:

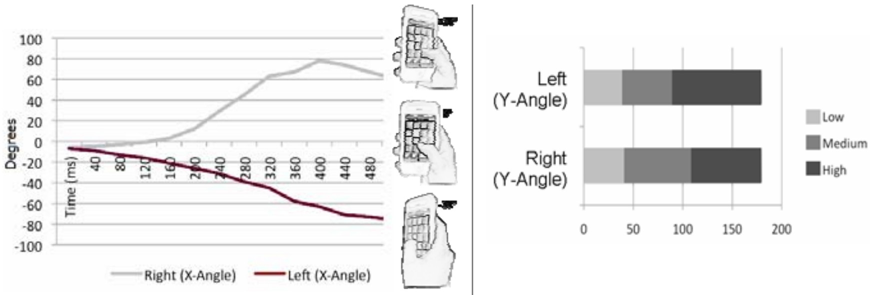
4 participants × 5 trials per gesture × 9 gesture locations = 180 trials.

The application recorded sample points from the moment the users pressed the 'Capture' button with the thumb to when they released it after completing the gesture. The user repeated the same gesture for a total of five readings and then the user was asked to perform another five gestures for another location on the grid.

4.2 Data Analysis

We used the X and Y angular readings to find the average relative displacements that created unique gestures. We found that the distinct cutoff points in the X movement were unambiguous. However, average tilt displacements in the Y direction varied and standard deviations revealed some overlap among the set of High/Medium/Low gesture locations.

Based on the samples collected we found that for right sided gestures, the average X angular points was clearly positive for gestures toward the right and negative for left gestures (Figure 2.a). To identify distinct relative movement for High, Medium or



(a) (b)

Fig. 2. (a) As can be seen from the above sample of X-Angle values, the X-Angle average is positive for Right gestures (+30° average) and negative for Left gestures (-30° average); (b) High, Medium and Low cutoff points for right/left sides have distinct ranges and are somewhat different due to differences in wrist ulnar/radial deviation for left and right sides [276]

Low positions, we inspected the Y angular values. Where unlike the right-left gestures that revealed sufficiently distinct averages on the X-axis of tilt, the distinct ranges for the relative angular movements in the Y-axis varied based on whether the user was Chucking to the Left or to the Right (Figure 2.b).

4.3 Determining the Gestures

X and Y angular points reveal unique patterns which help distinguish what gesture the user is performing with Chuck. For example we know that if the average of X is greater than zero, a right-sided gesture is being carried out. To break it down further, we use the distance in Y to determine if Chucking is to be mapped to a High, Medium or a Low gesture. The distance was calculated as: $\sqrt{(\Omega_2 - \Omega_1)^2}$ (where **Ω** is the Y-angle).

This form of relative mapping of X and Y angular positions was found to be easier to learn and less error prone as opposed to having an absolute mapping where gestures had to fall within a fixed range of angles to be recognized successfully.

5 Experiment and Results

We carried out an experiment to assess the limitations and performance of Chucking. We evaluated Chucking against Flick, with a target positioning task adapted from [14]. We alsocollected subjective ratings from the participants.

5.1 Hardware Configuration

We used the same PDA as described in section 4. The public display was simulated using a 52 inch monitor connected to a PC, to which the users transferred the targets. The PDA and PC communicated via a wireless connection and all the data was collected on the PC.

5.2 Performance Measures

The experimental software recorded number of errors, number of attempts, failure rate, and completion time as dependent variables. We recorded an *error* on any given trial in which the user was unable to move the PDA object to the *correct* target location. The trial ended only when the user selected the correct target, so multiple errors or multiple *attempts* were possible for each trial. For each trial we recorded the number of attempts used in correctly transferring the object. After five attempts, we marked the trial as a *failed trial*, and the participant then started the next trial. Trial completion time is defined as the total time taken for the user to either Flick or Chuck the object successfully. Since arm movements take longer and have a wider range-of-motion than finger movements, completion times with Flicking should be lower. However, we expected error and failures rates to be lower with Chucking as the user can engage in wider ranges of movements with the arm, than with the thumb.

5.3 Participants

Ten participants (8 males and 2 females) between the ages of 20 and 30 were recruited from a local university. All subjects had previous experience with graphical interfaces and were right-handed. None of the users were color blind, thus allowing the users to clearly see the red target objects. Furthermore, all ten participants were familiar with tilt based interactions, either from using the Wii™ or the iPhone™.

5.4 Task and Stimuli

To test whether participants could accurately and effectively position an object from the PDA onto the large display, we devised a task that consisted of placing a PDA object onto a target location on the large surface. The target location on the large surface was in one of several positions in a grid. Targets on the PDA were positioned in one or two locations, either at the center or toward the edge of the PDA. The choice of placement is representative of conditions in which the user would necessitate reaching one of the PDA locations with the thumb when operating with one-hand. Furthermore, this simulates the common case of having objects in an image browsing application, or any other application in which a list of items typically appear as thumbnails or icons. The object on the target position on the large surface appeared in red. If the user moved the PDA target to an inaccurate surface position, that position would get highlighted in yellow. When the object landed accurately on the surface target, the target changed color to green, and the user was then presented with the next trial. A trial would not be completed unless the user accurately moved the document to the desired position on the large surface. Users who failed the first time were given five attempts, after which the trial was marked as failed. The position to transfer to was randomly selected on the large surface as well as the target object on the PDA.

5.5 Procedure and Design

The experiment used a within-participants factorial design. The factors were:

- Technique: Flick, Chuck.
- Location (of object on PDA): Center, Edge
- Grid Size: 2×2, 3×2.

The order of presentation first controlled for technique. Half the participants performed the experiment first with Flick and then with Chuck. Levels of all the other factors were presented randomly. We explained the task and participants were given ample time to practice the techniques with various parameters of the other independent variables. The experiment consisted of two blocks with each block comprising ten trials per condition. With 10 participants, 2 Techniques, 2 Locations, 2 Grid Sizes (4 or 6 positions), the system recorded a total of (10×2×2×4×2×5) + (10×2×2×6×2×5) = 4000 trials. The experiment took approximately 45 minutes per participant, including training time.

6 Results

We recorded a total of 4000 trials. We present our results in terms of number of error trials, number of attempts, number of failures, and completion times.

6.1 Error Rate

We recorded an error on any given trial in which the user was unable to Chuck the PDA object to the correct location on the remote display. We used the univariate ANOVA test and Tamhane post-hoc pair-wise tests (unequal variances) for all our analyses with subjects as random factors.

There was a main significant effect of *Technique* ($F_{1,9}$=38.295, $p < 0.001$) and of *Grid Size* ($F_{1,9}$=71.303, $p < 0.001$) on error rate. Surprisingly, we did not find any main effects for *Location* ($F_{1,9}$=0.242, $p = 0.634$) on error rate. We found no significant interaction effect between these factors ($F_{1,9}$=0.698, $p = 0.425$, for *Technique* × *Grid Size*), ($F_{1,9}$=3.649, $p = 0.088$, for *Technique* × *Location*) , ($F_{1,9}$=0.013, $p = 0.911$, for *Grid Size* × *Location*). From the 4000 experimental trials (split in half for each technique), users performed an error on 342 trials with Chuck (over 2000 trials), yielding an error rate of 17.1% and 686 errors with Flick (over 2000 trials), resulting in an error rate of 34.3%. Figure 3.a shows the average error rate for each Technique and Location by Grid Size.

6.2 Number of Attempts

For each error trial, users could retransfer the target up to a limit of five attempts. After five attempts, we marked the trial as a *failed trial*, and the participant then started the next trial. Since users were asked to retry after an error, we also recorded the number of attempts taken to position a target accurately, and analyzed this dependent variable. There was a main significant effect of *Technique* ($F_{1,9}$=59.395, $p < 0.001$) and *Grid Size* ($F_{1,9}$=16.675, $p < 0.01$) on number of attempts. We did not find any main effects for *Location* ($F_{1,9}$=2.097, $p = 0.182$) on number of attempts. We found no significant interaction effect between *Technique* × *Grid Size* ($F_{1,9}$=0.001, $p = 0.921$) and *Grid Size* × *Location* ($F_{1,9}$=.114, $p = 0.743$). We found an interaction effect for *Technique* × *Location* ($F_{1,9}$=6.565, $p = 0.031$). Overall participants performed an average of 1.29 (s.d. .79) attempts with Chuck and 1.75 (s.d. 1.3) attempts with Flick. Figure 3.b shows the average number of attempts for each Technique and Location by Grid Size.

6.3 Failure Rate

After five attempts, we ended the trial and recorded it as a failure. There was a main significant effect of *Technique* ($F_{1,9}=14.865$, $p = 0.004$). We did not find any main effect for *Grid Size* ($F_{1,9}=1.186$, $p = 0.304$) or *Location* ($F_{1,9}=3.154$, $p = 0.109$) on the average number of failures. We found no significant interaction effect between *Technique* × *Grid Size* ($F_{1,9}=0.966$, $p = 0.351$), *Grid Size* × *Location* ($F_{1,9}=2.111$, $p = 0.180$) and *Technique* × *Location* ($F_{1,9}=2.727$, $p = 0.133$). Overall participants failed on 20 trials with Chuck (for a failure rate of 1%) and on 88 trials with Flick (resulting in a failure rate of 4.4%). Figure 3.c shows the average failure rate for each Technique and Location by Grid Size.

6.4 Completion Times

We measured the completion time for each successful trial. With Flick the completion time was the time from the moment the user placed his/her finger on the display up to the lift-off of the Flick. With Chuck, completion time elapsed from touching the target to the time to lift-off the finger to signal the completion of the gesture.

There was a main significant effect of *Technique* ($F_{1,9}=29.121$, $p < 0.001$) and of *Grid Size* ($F_{1,9}=14.285$, $p < 0.001$) on completion time but no main effects for *Location* ($F_{1,9}=0.412$, $p = 0.537$). We found no significant interaction effect between *Technique* × *Grid Size* ($F_{1,9}=4.967$, $p = 0.053$), *Grid Size* × *Location* ($F_{1,9}=.176$, $p = 0.685$), or *Technique* × *Location* ($F_{1,9}=0.534$, $p = 0.484$). On average it took participants 777 msecs to complete a trial with Chuck and 352 msecs with Flick. Figure 3.d shows the average completion time for each Technique and Location by Grid Size. It

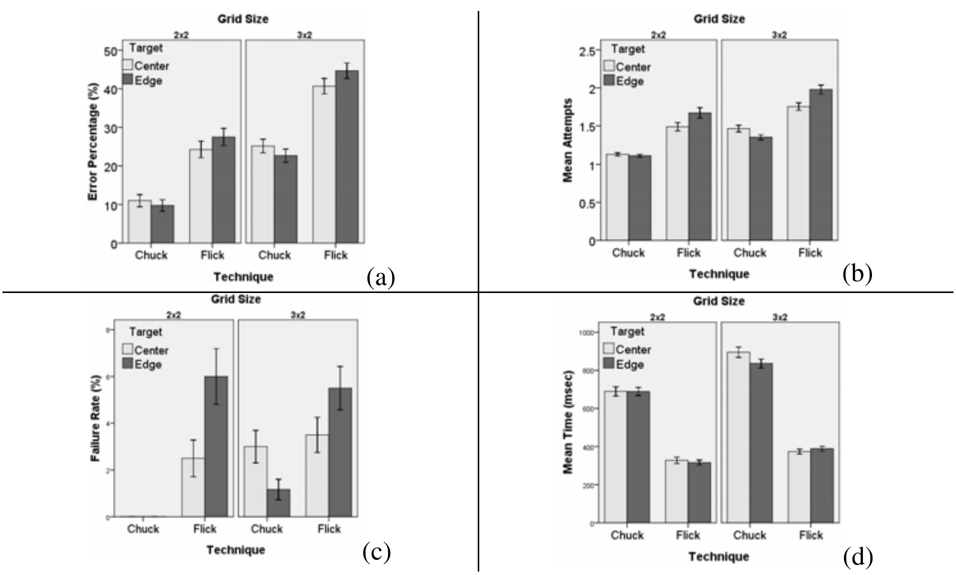


Fig. 3. (a) Average error rate, (b) Average number of attempts, (c) Average failure rate (d) Average completion time; grouped by Target Location and Grid Size. (+/-1 s.e.)

is not surprising that flicking took users less time to complete. The length of the gesture is significantly less with flick and finger movements are known to take much less time than full arm movements [4].

7 Extensions of Chucking

We designed three extensions to Chucking to create a suite of techniques for coupling private devices with public displays. These are primarily focused on achieving some of the primary tasks for document sharing, including document retrieval, precise document placement and document rotation on the public display. Very few if any analogies to these techniques exist with current systems (e.g. Toss-it [27] or Flick [13]). We demonstrate a complete set of private-public coupling interactions with Chuck-based metaphors.

7.1 Chuck-Back, Chuck-and-Place, Chuck-and-Rotate

While researchers have given significant attention to moving documents from the private device onto the public display, less attention has been placed on the opposite transfer. In *Chuck-Back*, users are able to retrieve objects from the public display and place them on the private device. To invoke Chuck-Back, users press the jog dial on the device. This places a cursor on the public display. The user can control the movement of the cursor by tilting the device forward, backward or side-to-side. Upon hovering over an object, the user can then initiate a pullback by making a motion in the direction opposite that of chucking (a flexion of the forearm). When the user releases the jog-dial, the intended public document is placed on the private device.

Chucking can also be extended to facilitate accurate object positioning. For experimental purposes, we restricted our document sharing task to only a few segments of the available space on the remote display. Similar to Superflick [14], *Chuck-and-Place* allows users to accurately position the document anywhere in the xy-plane, regardless of a grid. Unlike Superflick, we allow for fluid and accurate positioning by allowing the user to perform slight wrist movements within a given time period after transferring the document. Forward and backward tilt, moves the object in the y-plane, while side-to-side tilt moves the object in the x-plane.

Researchers have studied specific techniques to facilitate the orientation of documents on public spaces, such as on tabletops. Rotate-and-Translate (RNT) [14], allows the user to position a document and at the same time rotate it in the correct orientation once passed over to another side of the table. Similarly, *Chuck-and-Rotate* allows the user to freely rotate the document either in a clockwise or counter-clockwise manner using the thumb once the transfer is complete. We map the rotation of the thumb directly onto document rotation such that a thumb motion on the private device also rotates the active object on the remote display.

8 Discussion

We now provide a general discussion of our results and describe some potential limitations of Chucking.

8.1 Performance and Limitations of Chucking

Our results reveal that users are highly accurate with Chucking when the technique is used for transferring objects to general areas of a display. In particular, Chucking alleviates the common problems associated with thumb reach in one-handed interaction. As a result, this technique can reduce steps that would normally be required to bring objects closer to the thumb. Our results, although not statistically significant, show that target location of the object on the PDA was easier to select with Chucking when the object was more accessible (i.e. at the center).

We also observe that performance with either technique was better with a smaller size grid than a larger grid. This suggests that there is a threshold grid size beyond which performance will start to degrade. This is generally expected since there is only a limited range-of-motion for the wrist and forearm that will result in distinct gestures [14]. Chucking is particularly useful in cases where devices such as a cell phone or an iPod do not provide support for pen input. Instead, many of these devices use tilt sensors so Chucking would adapt well to these environments. Additionally, researchers have found that for brief interactions user prefer to not remove the stylus on a device and instead to use their fingers or hands. Chucking would be particularly beneficial for these types of situations. Finally, our study was only conducted with right-handed users. As such the mappings we used are specific to this group of individuals. Generally, ranges-of-motion for the wrist and forearm are symmetrical across both hands [14]. Thus, with minor modifications to our application, Chucking could work for left-handed users. We will investigate the necessary changes to make Chucking adapt to handedness.

8.2 Recommendations

Based on the results and observations we provide the following recommendations for creating one-handed techniques for document sharing:

- One-handed tilt gestures for document sharing are effective for mobile devices, particularly on those devices that lack pen input but provide built-in tilt sensors.
- Performing hand gestures can relieve the fingers for other tasks that may be required during the interaction, such as to rotate an object in a given direction.
- Hand gestures appear natural to users and can be learned with little effort if developed with the appropriate mappings, such as the use of relative tilt input.

9 Conclusion

We introduce Chucking, a one-handed tilt based technique that uses natural metaphors to assist users to share documents from a private device to a public display. In a study we show that users are more accurate with Chucking than one-handed Flicking. Chucking also provides the opportunity to map other tasks (such as orienting a document during the transfer) using the added input dimension of fingers such as the thumb, during the interaction. Any finger based technique for this type of task is inhibited by the range of coverage that the finger has on the device and by the edges of the device. We also introduce extensions to Chucking to create a suite of techniques that facilitate a range of interactions between mobile devices and public displays.

Additional work is needed to investigate thoroughly the simultaneous use of fingers and arm movements in one-handed interaction, the achievable levels of precision possible with Chucking, and the use of Chucking for other tasks, such as map navigation or document browsing.

References

1. Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., Zierlinger, A.: Drag-and-pop and drag-and-pick: techniques for accessing remote screen content on touch- and pen-operated systems. In: Proc. of Interact 2003, pp. 57–64 (2003)
2. Bezerianos, A., Balakrishnan, R.: The vacuum: facilitating the manipulation of distant objects. In: Proc. of CHI 2005, pp. 361–370 (2005)
3. Geißler, J.: Shuffle, throw or take it! working efficiently with an interactive wall. In: Proc. of CHI 1998 Extended Abstracts, pp. 265–266 (1998)
4. Grandjean, E.: Fitting the task to the man: an ergonomic approach, p. 372. Taylor & Francis, Abington (1969)
5. Hascoët, M.: Throwing models for large displays. In: Proc. of British HCI 2003, pp. 73–77 (2003)
6. Hinckley, K., Ramos, G., Guimbretière, F., Baudisch, P., Smith, M.: Stitching: pen gestures that span multiple displays. In: Proc. on Advanced Visual Interfaces (AVI 2004), pp. 23–31 (2004)
7. Hinrichs, U., Carpendale, S., Scott, S.D.: Evaluating the effects of fluid interface components on tabletop collaboration. In: Proc. on Advanced Visual Interfaces (AVI 2006), pp. 27–34 (2006)
8. Kruger, R., Carpendale, S., Scott, S.D., Tang, A.: Fluid integration of rotation and translation. In: Proc. of CHI 2005, pp. 601–610 (2005)
9. Karlson, A., Bederson, B.: One-handed touchscreen input for legacy applications. In: Proc. CHI 2008, pp. 1399–1408 (2008)
10. Karlson, A., Bederson, B., Contreras-Vidal, J.: Understanding Single-Handed Mobile Device Interaction. Tech. Report HCIL 2006 (2006)
11. Karlson, A., Bederson, B., SanGiovanni, J.: Applens and launchtile: two designs for one-handed thumb use on small devices. In: Proc. CHI 2005, pp. 201–210 (2005)
12. Maunder, A., Marsden, G., Harper, R.: SnapAndGrab: accessing and sharing contextual multi-media content using bluetooth enabled camera phones and large situated displays. In: Proc. CHI Extended Abstracts (CHI 2008), pp. 2319–2324 (2008)
13. Nacenta, M., Aliakseyeu, D., Subramanian, S., Gutwin, C.: A comparison of techniques for multi-display reaching. In: Proc. of CHI 2005, pp. 371–380 (2005)
14. Nacenta, M., Sallam, S., Champoux, B., Subramanian, S., Gutwin, C.: Perspective cursor: perspective-based interaction for multi-display environments. In: Proc. of CHI 2006, pp. 289–298 (2006)
15. Nacenta, M., Sakurai, S., Yamaguchi, T., Miki, Y., Itoh, Y., Kitamura, Y., Subramanian, S., Gutwin, C.: E-conic: a perspective-aware interface for multi-display environments. In: Proc. of UIST 2007, pp. 279–288 (2007)
16. Oakley, I., O'Modhrain, M.: Tilt to scroll: Evaluating a motion based vibrotactile mobile interface. In: Proc. WHC 2005, pp. 40–49 (2005)
17. Oakley, I., Park, J.: A motion-based marking menu system. In: Proc. CHI 2007 Extended Abstracts, pp. 2597–2602 (2007)

18. Pering, T., Anokwa, Y., Want, R.: Gesture connect: facilitating tangible interaction with a flick of the wrist. In: Proc. of Tangible and Embedded Interaction 2007, pp. 259–262 (2007)
19. Reetz, A., Gutwin, C., Stach, T., Nacenta, M., Subramanian, S.: Superflick: a natural and efficient technique for long-distance object placement on digital tables. In: Proc. Graphics Interface 2006, pp. 163–170 (2006)
20. Rekimoto, J.: Tilting operations for small screen interfaces. In: Proc. UIST 1996, pp. 167–168 (1996)
21. Rekimoto, J.: Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. In: Proc. UIST 2007, pp. 31–39 (1997)
22. Swindells, C., Inkpen, K., Dill, J., Tory, M.: That one there! Pointing to establish device identity. In: Proc. of UIST 2002, pp. 151–160 (2002)
23. Vogel, D., Baudisch, P.: Shift: a technique for operating pen-based interfaces using touch. In: Proc. of CHI 2007, pp. 657–666 (2007)
24. Wigdor, D., Balakrishnan, R.: TiltText: Using tilt for text input to mobile phones. In: Proc. of UIST 2003, pp. 81–90 (2003)
25. Wilson, A., Sarin, R.: BlueTable: connecting wireless mobile devices on interactive surfaces using vision-based handshaking. In: Proc. of Graphics Interface 2007, pp. 119–125 (2007)
26. Wu, M., Balakrishnan, R.: Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In: Proc. of UIST 2003, pp. 193–202 (2003)
27. Yatani, K., Tamura, K., Hiroki, K., Sugimoto, M., Hashizume, H.: Toss-it: intuitive information transfer techniques for mobile devices. In: Proc. CHI Extended Abstracts 2005, pp. 1881–1884 (2005)

This Just In! Your Life in the Newspaper

Bruno Antunes, Tiago Guerreiro, and Daniel Gonçalves

Universidade Técnica de Lisboa
Av. Rovisco Pais, 49
1050-001 Lisboa, Portugal
bruno.r.t.antunes@ist.utl.pt, tjvg@immi.inesc-id.pt,
daniel.goncalves@inesc-id.pt

Abstract. It is not uncommon for computer users to work on several things at once. However, to the computer, all documents, emails and applications are considered equal, regardless of why they were created or used. Little support is provided when trying to recall important information about a particular project or subject at a later time. What is more, there is no effective way to help users review their past activities to identify when a particular subject was of importance, what were their concerns at a given moment in the past, or simply review their activities during a period of time at a glance. In this paper we describe PersonalNews, a system in which users are presented with a personal newspaper, in which the news articles describe the subjects they were concerned with in a given period of time. Those articles are automatically generated from the users' documents, grouped according to their subject and analyzed for relevant passages describing them. We show that PersonalNews is able to recognize the subjects and projects the users were involved in, and even help them recall some they had forgotten about. Also, it can be used effectively to help retrieve documents on particular subjects.

Keywords: Personal Information Management, Personal Document Retrieval, Newspaper Metaphor, Information Visualization.

1 Introduction

Nowadays, many aspects of our daily lives require the use of a computer. From tasks at the workplace, to eGovernment initiatives, computers increasingly pervade our lives. This results in most users having to handle several different projects at the same time, shifting their attention from one to another, and producing a wealth of information, some of it unrelated, that might need to be retrieved at a later time.

Computers have no way of knowing when the users' work context changes. Instead, all documents, emails, appointments and other personal information are treated the same regardless of their subject or why they were handled or produced. Most current filesystems are the direct descendents of solutions invented when the amount of information to be handled was small, and its possible uses limited. As such, apart from their location in a hierarchy, filename and a set of dates, documents have little relevant semantics associated to them, making their organization difficult. Indeed, it has long ago been found that other more personal and semantically rich features, such

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 279–292, 2009.

as a document's subject, purpose or appearance, are better criteria for their organization than the hints provided by filesystems. This is clear in the case of email, in which email clients are often overloaded as document-organization tools, even if not originally designed to do so. This happens because email messages, unlike documents, have associated to them a wide range of contextual information (sender, subject, etc.) that makes managing them easier [15]. What is more, information relating to a single project or activity can be spread throughout different filesystems locations or even applications, with nothing to link it as a coherent whole. This Fragmentation Problem [1] further hinders the handling of personal information.

All this causes problems when users try to review their activities to find some relevant information. There is no simple way to visualize all information grouped into meaningful sets, directly derived from the users' interests and projects. Such an overview of the users' lives would help them to more easily recognize when was a certain project addressed, what was it about, and what documents resulted from it. Thus, such visualization would help users to find relevant information directly but also indirectly, by providing them with clues that can be used in traditional organization methods, such as dates and keywords.

In this paper, we describe PersonalNews, a system in which a newspaper metaphor is used to help users visualize their interests and projects, at any given period of time, at a glance. It automatically infers the different subjects that concerned the users by looking at their documents. Then, relevant passages of those documents are extracted and manipulated to create news articles that describe those subjects. Those articles are then laid out on a newspaper page according to their importance. By looking at this personal newspaper, the users are able to immediately recall relevant aspects of their lives and find relevant information for reference or re-use.

In the following section, we will describe some information visualization research that, like ours, tried to provide an overall view of personal information. We'll then describe the PersonalNews interface and, in the next section, after a description of the overall system architecture, we'll discuss how we managed to infer the different subjects the users worked on and extract relevant excerpts from documents. Afterwards we'll show how those excerpts were used to create news articles. The user evaluation of PersonalNews is described next. Finally, we'll conclude, pointing to potentially interesting future research in the area.

2 Related Work

In the last years we have witnessed many developments in personal information visualization. One of the first studies in the area was Forgot-me-not [8]. This article describes a PDA-like device where personal information is displayed, trying to help the memory of its user. For example, it allows remembering where to find a document or the name of a friend. While limited by technology existing at the time, but nevertheless interesting to us as it allows the users' lives to be visualized at a glance.

Lifestreams [5] organizes personal documents as a time-ordered stream that functions as a diary of user's life. It uses a simple organizational metaphor (document streams). The tail of the user's stream contains documents from the past. Moving away from the tail and toward the front, we find more recent documents. Its disadvantage is that it only focuses on documents and gives time a central role, while other, more semantically rich features (subject, purpose, etc.) are neglected.

LifeLines [10] is a general-purpose technique for visualizing personal history record summaries. They provide a complete visualization environment offering overviews, zooming, filtering and details on demand. An overview is always visible while providing easy access to the details. Its problem is that it handles personal data in limited fairly structured contexts, such as medical or court records.

Another interesting application is FacetMap [13], an interactive visualization system guided by queries. It uses faceted search where the topics used to organize the information (time, author, etc.) can be used to filter search results. User tests showed the interface to be confusing when dealing with several facets.

Stuff I've Seen [4] is a system with a built-in search engine that can index all the information that a user saw over a period of time. This information can refer to Web pages, emails and documents, among others. Search results are presented with the help of a timeline, annotated by personal landmarks (pictures, tasks, etc.) and public events (reports, vacations, holidays, etc.). This system is limited to simple, direct searches providing no in-depth semantic analysis of the available information.

More recently, Themail [14] creates a visualization of the information preserved in its users' email files. It shows the relation between the users and their email contacts, across the time, highlighting the most important keywords in the messages exchanged with them. Still, it is limited to the information contained in email messages and treats all the messages similarly, not taking into account their relative importance.

All these works provide a way to visualize personal information. However they ignore some important aspects considered by PersonalNews, such as presenting information allowing the immediate perception of semantically relevant patterns. Also, the information is shown in different ways according to its relative importance, making this evident to the eyes of the user, and helping guide its exploration.

3 The PersonalNews Interface

The PersonalNews interface was implemented as a Web Application whose front-end is a web page with the traditional look of a periodical. A locally-installed dedicated web server allows it to run locally and access the users' documents. However, it would also run remotely, if necessary. This makes our solution amenable to be used in tandem with the increasingly common web-office applications and remote document storage solutions, in which the users' documents no longer reside in their hard drives. The use of HTML and CSS gives us the versatility to format the different news in heterogeneous, dynamic ways. Furthermore, using this technology makes it possible for anyone who has ever surfed the web to use our application.

In the generation the newspaper page we used different font types, sizes and weights, to give it an appearance closer to that of an actual newspaper. Also, we resorted to other graphical elements such as horizontal and vertical lines to separate the articles, whenever relevant, as if the newspaper had sections. Also, when indexing the users' documents, we extract images in them. Those images will illustrate the articles, helping users to recognize their subject with help from their visual memories.

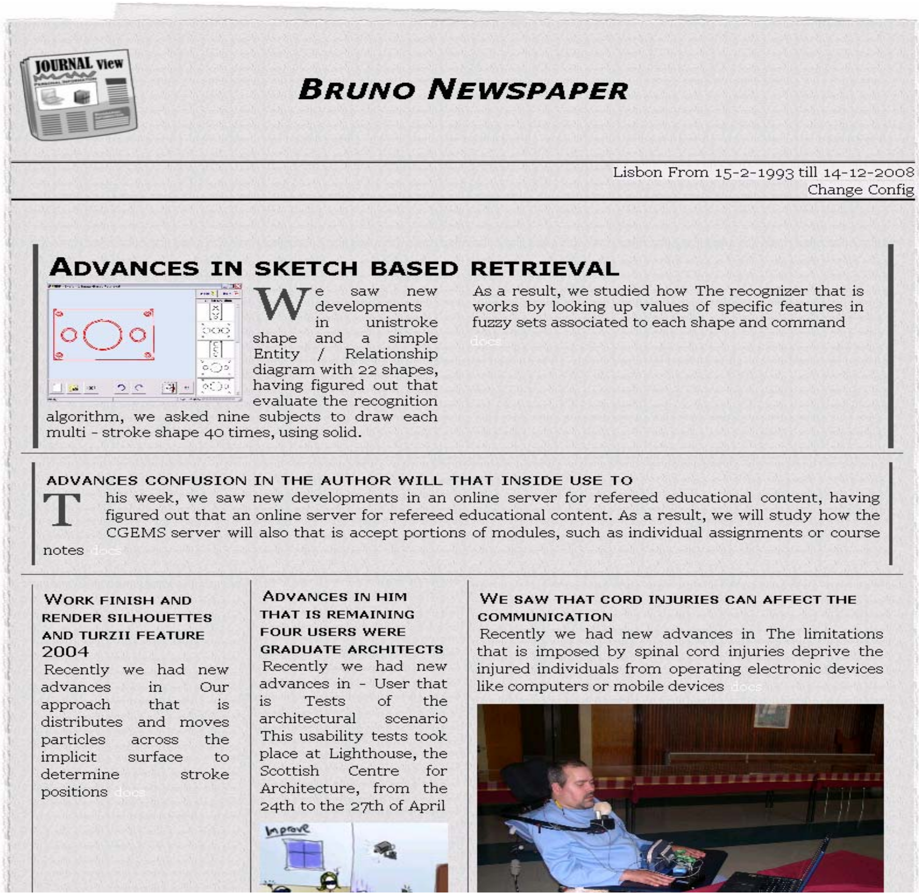


Fig. 1. The PersonalNews Interface

When the users launch the application for the first time, it asks them to configure the application, most importantly, to give a name and the location of the documents to index. The users can also control the newspaper generation with other parameters. However, we provide good enough defaults, as described in the next section.

The main application page shows the personal newspaper edition for a selected time period (Figure 1). At the top is the newspaper heading, including its title and the dates which it refers to. Below appear the news articles, placed in the page according to their relative importance (as inferred from the number of personal documents underlying each article). More important articles are shown, as for real newspapers, with bigger headings and occupying a larger area. Each article has, at its bottom, a link which the users can follow to reach a list documents that were used to create the article. This gives PersonalNews the potential to be used as a document-retrieval tool.

To select a time span the interface provides a browseable, zoomable timeline (Figure 2). It displays previously created newspaper editions, which are cached to

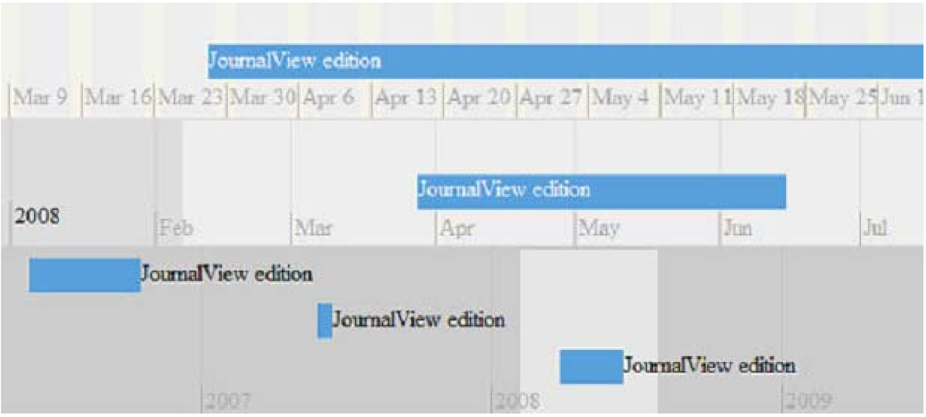


Fig. 2. Newspaper timeline

make re-visiting them more efficient. To create an edition reporting on a new time period it is enough to select a time interval by direct manipulation of the timeline.

3.1 Newspaper Layout

The placement of the articles on the page is automatically determined based on the number of news to be displayed in the newspaper (depending on the number of relevant subjects in the time-period) and their importance (based on the number of documents that gave rise to them). An algorithm of successive subdivision of space is used to accomplish this. The page is first divided in two vertically, and the topmost part is reserved for the article of bigger importance. The remaining of the page then is divided to accommodate the news of the next importance levels. If each importance level has up to three articles, they will be placed side-by-side, in different columns. Otherwise, three will be placed side-by-side and the remaining in the next vertical division, and so on. The size reserved for each sub-division depends on the importance of the news to place there, relatively to the others. This ensures that more important news (corresponding to more important subjects) are displayed more pre-eminently and closer to the top.

The images extracted from the users' documents to illustrate the news are placed close to them. The decision of whether or not to use an image is governed by two criteria: the template used to create the article (as described later on this paper) might suggest its use; or they will be used to produce a more aesthetically pleasing result.

4 Creating Personal News Articles

In order to create the news articles, it is necessary to analyze the documents and the information therein. First it is necessary to identify the documents' subjects and group them accordingly. Then, documents in the same group are analyzed, resulting in a set of excerpts that meaningfully represent the group's subject. Using those excerpts, and taking into account their morphology, it is possible to create news articles.

4.1 Inferring Document Subjects

Given a set of personal documents, in order to infer which subjects those documents relate to it was necessary to group them according to their content. Since it is impossible to know beforehand what subjects might arise, some kind of unsupervised technique is necessary. Hence, chose to do this by resorting to a clustering algorithm.

Clustering Documents. One of the better known clustering algorithms is k-means, and its variants [16]. However, it requires the desired number of clusters to be know beforehand. It is impossible to know this as it will vary greatly from user to user and time-span to time-span. There are strategies in which different cluster numbers are tried and the best result is chosen. However, they are not well suited for an interactive system, since they are time-consuming and computationally intensive.

For those reasons, we chose the QT-Clust algorithm [6]. It does not require the number of clusters to be known beforehand. Instead it needs a clustering radius. While this value might be hard to assess for the general case, it is still easier to do so than to arbitrarily predetermine the number of subjects. Furthermore, this radius can be used to fine-tune the granularity with which document subjects are considered. A large radius might result in all HCI papers to be placed in the same cluster, while a smaller one could separate between usability, user models, and other sub-topics.

For QT-Clust to work, it is necessary to have some kind of measure of how closely related documents are. We chose as its basis sets of relevant keywords extracted from the documents using the tf-idf algorithm [12]. This gives us keywords that appear frequently in a document but rarely in others, thus discriminating it. Stopwords were removed and the remaining stemmed, using the Porter Stemmer [11], before tf-idf.

The most relevant keywords for each document aren't enough to adequately group them. Related documents might not share enough keywords to be deemed similar. For instance, a document might refer the "clustering of documents" and another to the "grouping of documents". We needed to, somehow, abstract from the keywords into the underlying subjects. To accomplish this, we used Latent Semantic Analysis.

Latent Semantic Analysis. Latent Semantic Analysis (LSA) [3] abstracts from keyword frequencies to semantic descriptions of the documents' subjects and requires no a priori knowledge of those subjects. It is, thus, of special interest to us as it is able to go beyond keywords appearing in documents when assessing their semantics. LSA is based on a matrix operation called Singular Value Decomposition that identifies identifying less relevant keywords whose influence can be discarded, thus removing much "noise" and in effect reducing the semantic dimensionality of the data. The use of tf-idf to choose only the most relevant keywords in the documents allows us to control the computational complexity of LSA. The matrix resulting from LSA can be seen as giving us the coordinates of documents in the semantic space (the vector of the weighs given to each term for each document). The Euclidean distance between them establishes how closely related documents are, a necessary input for QT-Clust.

Join Clusters. We found that, while the clustering algorithm produced good results, it tended to generate a number of clusters larger than the number of actual subjects in the documents. A closer inspection of the results showed that several of those clusters were nevertheless very close together in the semantic space. Increasing the clustering radius, however, quickly led to the creation of too general clusters. Instead, we opted for merging clusters whose centers are closer than a pre-determined threshold. This maintains semantically related documents in the same groups.

Tuning the Parameters. In this process, several parameters are relevant: the *clustering radius*; the *join distance* used to merge nearby clusters; the *dimension reduction* of in LSA; and the *number of keywords* used. It will be hard for users to know which values produce good results. Hence, we performed a set of tests to estimate which values function well in an average or general situation.

We performed our tests resorting to two different document sets, pre-classified into different subjects. The first test set, TS1, was composed of 116 scientific papers produced in our research group[1], in eight different but related areas: *eLearning*, *Accessibility*, *Mobile*, *CG*, *Improve*, *Multimedia Information Retrieval*, *Sketch-Based* and *Narrative Based Document Retrieval*. The second test-set, TS2, is composed of papers downloaded from the IEEE and ACM digital libraries, on eleven subjects: *Collaborative Tagging*, *Consensus Algorithm*, *Data Mining*, *Embedded Computation*, *Intelligent Environment*, *LSA*, *Network Protocols*, *Parallel Computation*, *Speech Synthesis*, *Use cases* and *Sketch Based Interfaces*.

To assess the quality of the results, we used two different measures: the percentage of correctly grouped documents and the Silhouette Coefficient (SC) [7] that gives us an estimate of how independent the clusters are from each other. Values between 0.7 and 1.0 indicate an excellent cluster separation, values between 0.5 and 0.7 a medium separation, and lower values show that the clustering process yielded poor results.

In terms of the *clustering radius*, we performed the clustering with all possible values between 0 and 1.0, with a 0.01 step (this step was used for all tests). The results, for both test sets, were similar, and are exemplified, for TS2, in Figure 3. As can be seen, the best classification is found for a radius between 0.20 and 0.25. In terms of SC, the best values (around 0.65, indicating an average cluster separation) also appear in that interval, for both test-sets. We see that a good value for the radius is 0.22.

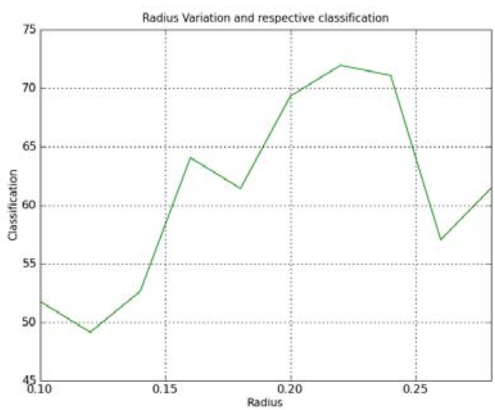


Fig. 3. Classification % vs Clustering Radius

[1] Visualization and Intelligent Multimodal Interfaces group of INESC-ID, http://www.vimmi.inesc-id.pt

As the SC values indicate, there are several close-by clusters that might be merged together. To control this, we studied the best values for the join distance. We found that the SC value improves significantly in the 0.2-0.7 interval. Of course, as the SC improves, the number of clusters reduces, as does the percentage of correctly classified documents. This value starts worsening at 0.3, so we conclude that a good value for the join distance is in the 0.2-0.3 range. This makes sense, as it is comparable to the clustering radius. By choosing a similar value we are merging together clusters that the algorithm might have grouped in the first place if not for some quirk in its working. We also tested both variables at once. The results corroborate our previous findings.

In terms of the number of keywords used to infer the documents' subjects, we found that the value that produces the best results is between 20 and 30. The results were similar but not uniform across test-sets. This was to be expected, as the keywords vary greatly according to the domain. For TS2 there were two peak classifications, at 20, 50 and 80 keywords. The values for those numbers were, however, very similar, allowing us to assume a good compromise at 20 keywords.

Finally, regarding the dimension reduction in LSA, we saw, for both test-sets, that reducing the dimension by more than 10 produces a sharp drop in classification quality. As the time it takes to process the data increases with the dimension reduction, we chose to fixate this parameter at 1.

We conclude that the best values for the different parameters are: *clustering radius=0.22*; *join distance=0.22*; *dimension reduction=1*; and *number of keywords=20*. We used those values in all the user tests described below.

We also looked at the actual documents placed in each cluster. For the most part, all clusters strongly relate with a particular subject. They either contain documents of a single subject, or most of them are. The results were better, as expected, for TS2, as its subjects were much more independent than those of TS1. Indeed, while the documents have been classified into subjects, it is not impossible for them to mention things of others. In TS2, for instance, several Multimedia Retrieval documents were classified as Sketch-Based Interaction. A closer look at the documents showed them to describe a sketch-based retrieval tool. Hence, while the results aren't numerically perfect, in semantic terms they are more than adequate. It can be argued that the algorithm, in some cases, performs a classification which is better than the one done by hand. Furthermore, when the sets of documents are further processed in order to create the news articles, this some of the wrongfully classified documents will be filtered out, as will be shown below.

Excerpt Extraction. From sets of semantically-related documents, we wished to create news articles succinctly describing those subjects. To do so, we extracted relevant excerpts from the documents.

As a starting point, we considered, for each cluster, the keywords previously extracted from the documents in the cluster using the tf-idf algorithm. We took into account the possibility of applying the algorithm again to only the documents in the cluster, but this would have yielded worse results. The existing keywords describe the document in terms of how unique it is when compared with all of the users' documents. A new set of keywords would compare the documents inside the cluster, producing keywords that stand out in it. If, for instance, a cluster has 95% of documents on a particular subject and the remaining 5% were erroneously placed there, the most likely outcome would be for keywords actually describing the clusters' subject not to be chosen (as they are bound to be very common inside it), and keywords regarding the 5% misses to be selected, thus polluting the results.

Given the keywords, we repeat the following algorithm for all documents in the cluster: after removing the stopwords and stemming the document, we look for a keyword in the text. If we find it, we look for other keywords inside a two word radius from it. We repeat this process if new keywords are found inside the radius. At the end, we will have identified an excerpt in which several keywords occur and that is likely to be relevant to the document's subject.

If when a keyword is identified no others are found inside the two-word radius, this radius is increased by one (up to a limit of five). While close-by keywords are preferable, we found it better to have poorer excerpts than none. Even so, we defined a quality metric for the excerpts: the number of keywords in the excerpt divided by the radius used to obtain it. Excerpts with closer keywords will be better, as it is more likely they are semantically related. The method described above offers no guarantees regarding whether entire phrases are extracted, or that the extracts are understandable in their entirety. Some post-processing is done to minimize these problems.

It would be extremely difficult to accurately use a full-fledged parser, such as a chart parser, to try to understand if the phrases are complete and make sense. Such a parser requires a well-defined grammar and, as one of our premises is that the excerpts most likely *won't* be syntactically correct, such an approach would fail. Less constrained parsers such as a chunk parser might do better, but their unstructured nature could make it difficult to properly suggest missing sentence elements. Thus, we chose to define a set of simplification / correction heuristics that detect the most common problems and correct them.

Most notably, looking for commas and periods offers a good way to understand how far a phrase should extend to make sense. For instance, if an excerpt's last word are preceded by a period, it will most likely belong to another sentence and can be removed. Similarly, if the excerpt ends just one word before a period or comma, that word needs, most likely, to be included in the excerpt. We took care to recognize special cases, such as commas separating enumerated list elements. Pronouns such as "that" and "which" usually separate the main part of a sentence from a relative clause that, while clarifying some meaning, is often not central to the overall subject of the sentence. Such cases, when detected, are subject to a similar treatment.

Creating News Articles. In the possession of a set of meaningful excerpts describing the subject of a cluster of documents, it is now necessary to use those excerpts to create readable texts describing that subject which can be used as a news article. This is not simply a matter of juxtaposing several excerpts, as it would yield incoherent texts. Furthermore, we can take advantage of the article creation process to select the excerpts with greater quality, and include additional information in the articles, gleaned from other sources, if the need arises.

We created a set of news article templates, in which different patterns for possible articles are defined, with blank areas to be filled with document excerpts. Our templates are expressive enough to accommodate a wide range of situations. This is done by allowing the specification of alternate texts to be used depending on properties of the excerpts. For instance, different texts can be chosen depending on the tense of the verb in the excerpt chosen to fill in the following blank. Also, it is possible to specify, in the excerpt declaration itself, what properties it should have, such as its size, and whether it begins with a noun phrase, verb phrase, etc. All this expressivity allows us to build rich templates that adapt to the wide range of possible excerpts that might arise. Whenever an excerpt is needed to fill in a blank, all excerpts for the cluster that satisfy the constraints specified in the template are considered. The better classified will be chosen. In case of a tie, the excerpt is chosen randomly.

Even with all the versatility we've just described, sometimes the border between the text in the template and that of the excerpts is a little awkward, making the article hard to understand. So, as a final step in the article generation phase, the text resulting from the template is post-processed to improve its quality.

The first step in that process is the use of heuristics to polish the text. For instance, when a noun from the template immediately precedes a verb from an excerpt, it is often the case that inserting *that* or *that is* improves the results. This would transform "We implemented improvements increased the retrieval rate" into "We implemented improvements *that* increased the retrieval rate"

Even after these heuristics are applied, sometimes there are situations still to be solved. So, the second step is to compare the sentences' structures to that of structures known to be valid English sentences, correcting them where necessary. We created a list of possible sentence structures by analyzing all sentences in 18000 news articles from Reuters. A chunk parser was used to find the category of the different words in those sentences. From this data we created a list of sentence structures, ordered by frequency. Whenever a news article needs to be polished, its own words are classified according to their type, and then the Levenshtein distance [9] is used to select structures that more closely match it. If any words are missing, they will be chosen among those in the documents from where the excerpts were taken by looking at where they originally were located and their grammatical type. This yields imperfect but readable texts. One example of a (short) news article created by PersonalNews is:

Work in mesh quality and mesh accuracy

Recently we had new advances in partitioning techniques that are the most popular methods for rendering implicit surfaces, by creating a polygonal mesh.

5 Evaluation

To prove the adequacy of our approach and the quality of the personal newspaper, we performed a set of user tests. A total of 18 users tested PersonalNews. Their own personal documents were used. To do so, we visited them at home or working place and, after installing the system on their computers, indexed their documents. As finding users for which large numbers of English language documents accurately describe their interests is not easy, in Portugal, we focused on college students, for which this is true: most, if not all, books, papers, and other materials are provided in English to those students. Each user was asked to perform four different tasks:

1. *Setting-up and launching the application:* the user must run PersonalNews for the first time, configure it and let it index all relevant documents;

2. *Identify subjects from different time spans:* the user will start by choosing a particular week, and then writing down all subjects he or she remembers working on at that time, classifying each according to its perceived relevance from 1 (less relevant) to 5. Then PersonalNews is asked to create a newspaper edition for the particular time period and the subjects of the news articles are identified by the user and compared to those previously recorded.

3. *Search for a particular subject:* the user starts by choosing a subject and then requests PersonalNews to create an edition for a time when it is thought that the subject might have been of interest. The articles are then scanned for news about that subject.

4. *Search for a document:* after choosing a particular document, the user tries to employ PersonalNews to find that document by generating an edition in which one of the news might have been created using it.

While Task 1's main purpose was to introduce the users to the system, the other three allow us to analyze its *completeness* by checking if all subjects found relevant by the user are identified and result in the creation of news articles in which the subject is identifiable. Also, we look at how PersonalNews can help the users *remember subjects* they might have forgotten about. In this capacity, the system will act as a proxy memory for the user. Finally, we evaluate how PersonalNews can work as a *personal document retrieval tool*, to which users can resort to find their documents, based on their subject. Right now, PersonalNews is able to handle most common document types. This includes plain text, HTML, pdf, ps, Microsoft Word, etc. but also Excel worksheets, Powerpoint presentations, etc. It resorts on a plugin-based system we developed that is capable from handling all those formats. Other formats could be added with ease. The only constraint is that PersonalNews works with text, so there must be some text to be extracted somewhere in the document.

Task 1 – Acquaint users with PersonalNews. The installation was, in all cases, straightforward. The users had only to provide a folder where their personal documents were stored. (often "My Documents"). All users used the default values for the parameters.

Task 2 – Evaluating Completeness. The results for this task are summarized in Table 1. For a week-long interval, the users remembered between 1 and 8 different relevant subjects (avg=3.67, st.dev=2.11). Of those, an average of 3 were identified by the

Table 1. Completeness Results

	Week			Month			Year		
	$\bar{x}$	σ	%	$\bar{x}$	σ	%	$\bar{x}$	σ	%
num. remembered	3.67	2.11		3.78	2.05		4.33	2.44	
% rem. found PN	3	1.49	**81%**	3.27	1.53	**86.7%**	3.78	1.85	**87.2%**
# found PN only	1.72	1.07		1.56	0.98		1.67	0.77	
Import. rememb.	3.45	1.17		3.44	1.08		3.73	1.02	
Imp.not found PN	2.5	1		2.56	1.01		2.5	1.08	

application (sd.dev=1.49). This represents a success rate of 81.81%. Furthermore, the application found up to 4 subjects (avg=1.72, st.dev=1.07) not previously remembered by the users, nearly half of the ones they recalled alone.

Similar results can be found for monthly and yearly periods. As expected, the number of subjects remembered by the users is greater and we found them to be broader in scope for longer periods of time. The percentage of subjects found by the application grows from 81.81% (week) to 86.7% (month) and 87.2% (year). Broader and more important subjects are more easily inferred by PersonalNews. It is also important to note that while the number of retrieved subjects is around 81%-87%, most of those not found by PersonalNews are the least relevant for the user. Figure 3 depicts this clearly: all subjects of importance 5, and nearly all of 4, are found.

We conclude that PersonalNews is able to find most of the subjects the users find relevant, especially the most important.

Task 3 – PersonalNews as a Memory Proxy. When asked to find a news article from a subject of the users' liking (decided before using the system), we found that this was possible 88.9% of times. Only in two situations there wasn't a relevant article in the newspaper edition generated by the user.

Task 4 – Retrieving Documents. After choosing a particular document, the users then tried to find it using PersonalNews, which entailed creating a newspaper edition for the right period of time, identifying the appropriate article in it, checking the list of documents that were used to create it, and finding the target document in that list. All users managed to do so, for a success rate of 100%. Ten users did it on their first attempt, six had to check the document lists for two articles, and two had to create two newspaper editions to do so.

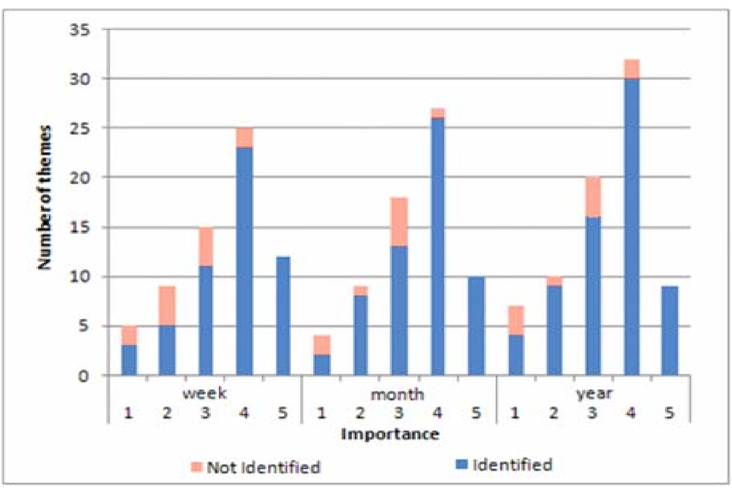


Fig. 4. Subjects found by personal news, by importance

5.1 Discussion

PersonalNews identifies the vast majority of subjects that users remember and, of those, all they find more relevant. Furthermore, it can even help users remember forgotten information. Also, we see that the users have little trouble understanding the news articles and their subjects. Thus, while is clear that some of the sentences used in the news articles are not perfect, they are nevertheless readable and their contents easily understood. We can state that PersonalNews provides an effective way for users to have a meaningful overview of their lives and interests.

After performing the aforementioned four tasks, the users were then asked to fill in a satisfaction questionnaire. Each question was classified with a score ranging between 1 (Bad) to 5 (Excellent). The users were, mostly, pleased with PersonalNews. Noteworthy are the users' opinions regarding the completeness of the application, (avg=4.17, stdev=0.86), corroborating the results obtained from the four tasks. Users also found it easy to retrieve documents (avg=4.56, stdev=0.92) and, overall, they think the system is useful (avg=4.17, stdev=0.38). To be improved is article legibility (avg=3.67, stdev=0.68), probably due to a less than perfect sentence quality (avg=3.11, stdev=0.58). The users found it adequate but not excellent.

6 Conclusions

One of the problems computers users nowadays face is to deal with the growing amounts of personal information in their computers. It is hard to to remember and retrieve relevant information from a particular point in time

PersonalNews provides the users with a personal newspaper that gives, at a glance, to have an idea of their interests for a particular time span. It can be used for memory elicitation, as well as a document retrieval. However, we've also seen how the production of subject summaries, in the guise of newspaper articles, might hinder those tasks. Unfortunately, a full-fledged natural language understanding approach might be unfeasible given the unrestricted nature of personal documents. This suggests that, when creating the summaries, we take more care in extracting full syntactic units from the documents, rather than trying to create our own. This should be addressed in future versions of this research. Also, it would be interesting to resort to other information sources, such as emails, instant messaging logs or datebooks to create richer news articles. Finally, new ways to lay out the articles would merit further study, moving to other technologies that, unlike HTML+CSS, give us better positioning metrics control.

References

1. Bergman, O., Beyth-Marom, R., Nachmias, R.: The project fragmentation problem in personal information management. In: Proceedings CHI 2006, pp. 271–274. ACM Press, Montreal (2006)
2. Brooks, C., Montanez, N.: An Analysis of the Effectiveness of Tagging in Blogs. In: Proc. AAAI Conf. 2005 AAAI Spring Symposium on Computational Approaches to Analyzing Weblogs. AAAI Press, Menlo Park (2006)

3. Deerwester, S., Dumais, S., Furnas, G., Landauer, T., Harshman, R.: Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science 41(6), 391–407 (1990)
4. Dumais, S., Cutrell, E., Cadriz, J.J., Jancke, G., Sarin, R., Robbins, C.D.: Stuff I've seen: A System for personal information retrieval and re-use. In: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 72–79. ACM Press, New York (2003)
5. Fertig, S., Freeman, E., Gelernter, D.: Lifestreams: an alternative to the desktop metaphor. In: Conference companion on Human factors in computing systems, pp. 410–411. ACM Press, New York (1996)
6. Heyer, L.J., Kruglyak, S., Yooseph, S.: Exploring expression data: identification and analysis of coexpressed genes. Genome Res. 9(11), 1106–1115 (1999)
7. Kaufman, L., Rousseeuw, P.J.: Finding Groups in Data: An Introduction to Cluster Analysis. Wiley Series In Probability and Statistics. Wiley Interscience, Hoboken (2005)
8. Lamming, M., Flynn, M.: "Forget-me-not" Intimate computing in Support of Human Memory. In: Proceedings FRIEND21 Symposium on Next Generation Human Interfaces (1994)
9. Levenshtein, V.: Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics Doklady 10, 707–710 (1966)
10. Plaisant, C., Milash, B., Rose, A., Widoff, S., Shneiderman, B.: LifeLines: visualizing personal histories. In: Proceedings of the SIGCHI conference on Human factors in computing systems: common ground, pp. 221–227. ACM Press, New York (1996)
11. Porter, M.F.: An Algorithm for Suffix Stripping. Program 14(3), 130–137 (1980)
12. Salton, G., McGill, M.J.: Introduction to modern information retrieval. McGraw-Hill, New York (1983)
13. Smith, G., Czerwinski, M., Mayers, B., Robbins, D., Robertson, G., Tan, D.S.: FacetMap: A Scalable Search and Browse Visualization. IEEE Transactions on visualization and computer graphics 12(5), 797–804 (2006)
14. Viégas, F.B., Golder, S., Donath, J.: Visualizing Email Content: Portraying Relationships form Conversational Histories. In: CHI 2006: Proceedings of the SIGCHI conference on Human Factors in computing systems, Montreal, Canada, pp. 979–988. ACM Press, New York (2006)
15. Whittaker, S., Sidner, C.: Email overload exploring personal information management of email. In: Conference proceedings on Human factors in computing systems, pp. 276–283. ACM Press, New York (1996)
16. Xu, R., Wunsch, D.: Survey of clustering algorithms. IEEE Transactions on neural networks 16(3), 645–678 (2005)

Instruction, Feedback and Biometrics: The User Interface for Fingerprint Authentication Systems

Chris Riley, Graham Johnson, Heather McCracken, and Ahmed Al-Saffar

Advanced Technology & Research
NCR Labs
Dundee
United Kingdom
cr230046@ncr.com

Abstract. Biometric authentication is the process of establishing an individual's identity through measurable characteristics of their behaviour, anatomy or physiology. Biometric technologies, such as fingerprint systems, are increasingly being used in a diverse range of contexts from immigration control, to banking and personal computing. As is often the case with emerging technologies, the usability aspects of system design have received less attention than technical aspects. Fingerprint systems pose a number of challenges for users and past research has identified issues with correct finger placement, system feedback and instruction. This paper describes the development of an interface for fingerprint systems using an iterative, participative design approach. During this process, several different methods for the presentation of instruction and feedback were identified. The different types of instruction and feedback were tested in a study involving 82 participants. The results showed that feedback had a statistically significant effect on overall system performance, but instruction did not. The design recommendations emerging from this study, and the use of participatory design in this context, are discussed.

Keywords: Biometrics, Fingerprint, Instruction, Feedback.

1 Introduction

As information and communication technologies (ICT) become ever more pervasive in modern life, the security of these systems has become an increasingly important issue. Authenticating legitimate users of computing systems is a necessary process with a number of unique challenges. User authentication falls into three different categories; knowledge-based authentication, token-based authentication and biometrics [1]. Knowledge-based authentication, such as passwords and personal identification numbers (PINs) rely on non-obvious information to confirm the legitimacy of an individual. Token-based authentication relies on the presence of a physical object to authenticate users. In contrast, biometric authentication technologies measure physical, behavioural or anatomical characteristics of the user to verify identity. The attraction of using biometrics is that the characteristics used to authenticate people cannot be lost, forgotten or readily stolen [2]. Biometrics have the potential to confirm the

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 293–305, 2009.

presence of the actual user, rather than just their password or identity token and are therefore seen as more secure than other forms of authentication.

Biometric authentication technology is beginning to mature and the technology is finding application in both commercial and public sector environments. The International Biometrics Group predicts that the biometrics market will see steady growth and will double in size over the next 5 years [3]. There are a number of trends that underscore this increasing uptake of biometrics. Firstly, there is an international trend towards secure user identification. There are now several large-scale, public facing implementations of biometric systems, including the US-VISIT scheme and the proposed identity card scheme in the United Kingdom. Secondly, the increase of computer security incidents and the need to safeguard information will contribute to an increased usage of biometric technology [4, 5]. Finally, biometric technology is often described as a positive development for the public at large [6]. According to some authors, the benefits of biometrics will eventually lead to the technology being used in almost every application that requires personal authentication [7].

1.1 Usability of Biometric Authentication Systems

Biometrics may be described as the future of user authentication, but there are a number of issues associated with the use of biometrics. The process of biometric authentication involves two stages; an enrolment or registration stage and an authentication stage. During enrolment a biometric sample is associated with an individuals' identity. Identification or verification is the process of matching a second biometric sample with the enrolment sample to verify an individual's identity [8]. The process of automated identity verification through biometrics is often not transparent to users though and most people have little or no familiarity with the technology. Obtaining a high quality enrolment sample is key to ensuring overall system performance, as a poor quality enrolment has a detrimental effect on all subsequent authentications. However, enrolment is the first time most people will have ever used a biometric system, so ensuring quality enrolment is a challenge for all biometric systems. Furthermore, like many other user authentication systems there is a negative relationship between the security and usability of biometric systems [9]. If an implementation of biometric technology is to be successful, both the performance and usability of the system must be carefully considered.

Other usability issues have been identified with specific biometric technologies. Most of the user-centric research on biometrics to date has centered on fingerprint systems, as these systems are the most commonly used biometric [3]. A number of usability issues have been identified with fingerprint systems. For instance past research [10] has found that that finger placement issues arose as users had difficulty placing their finger in a consistent manner. Difficulties in placing the correct part of the finger on the sensing surface, and applying the correct amount of pressure, were also described as problems with fingerprint systems. A lack of system feedback and a lack of instruction were further issues identified when different fingerprint systems were evaluated [10, 11]. They argued that the design of the systems needed to be improved to facilitate image acquisition. Usability problems with biometric systems have significant ramifications, as people are unlikely to tolerate being mistakenly denied access to their place of work, computer or funds.

1.2 The Biometric Interface

The interface plays an integral role in the usability, or otherwise, of any interactive system and biometrics are no different. Recent examples include the importance of interface design for web tasks [12] and the effect of interface design on peoples' ability to use encryption tools [13]. The design of the interface for fingerprint systems has received an increased amount of attention over the previous few years. Broadly speaking, this research has assessed the instruction and feedback provided to people as they use biometric systems. These issues will be discussed in turn.

Previous research conducted within the financial industry has found that different types of instruction have a significant impact on the overall performance of fingerprint systems. In our experience, verbal instruction from an experienced operator helps people give the highest quality enrolment and is superior to other forms of instruction. Other research has investigated modes of instruction, comparing the effectiveness of face-to-face, video and graphical instructions on the use of a fingerprint system [14]. It was found that pictorial instruction performed significantly worse than either face-to-face or video instruction. Graphical and pictorial instruction for biometric systems has started to be investigated by others [15, 16], however it is unclear how these approaches were evaluated. The small number of studies published to date suggest that face-to-face instruction best facilitates the enrolment process. Verbal instruction is often not a practical approach however, as many biometric systems are used in unattended environments such as automatic teller machines (ATMs) and access to secure physical locations [7]. There is a need to understand how to effectively deliver instruction for biometric systems without face-to-face communication. The information that should be contained in instructions for fingerprint systems is also not well understood.

The feedback a system provides is a second essential aspect of biometric authentication. In addition to providing information about when to place and remove the finger from the sensor, feedback about finger placement is necessary to facilitate the image acquisition process. The position of the finger on the sensor, pressure, finger movement and skin wettedness all affect the image acquisition process, but it is difficult to relay all of this information without overloading users. Feedback in the form of a biometric sample quality measure has been investigated and the effect of image acquisition assessed [17]. The effect of this 'quality gauge' feedback over several weeks and it was found that peoples' performance with the fingerprint sensor improved overtime. This is not an optimal approach to presenting feedback however, as no specific information about finger placement is provided. It was argued that people are not readily able to view an image of their fingerprint on screen and assess whether it is a high or a low quality image [17]. A well designed fingerprint system should provide feedback that is immediately understandable the first time it is used, though it remains unclear which type of feedback is most helpful for users.

We set out to design a graphical user interface for a fingerprint system that would include instructional and feedback aspects that would allow people to use the system effectively without assistance. A participatory design process was used to develop this interface. The development and testing of this interface are described below.

2 System Development

A graphical user interface was developed to support a commercially available, capacitance fingerprint sensor. An image of the sensor can be seen in figure 1 below. An iterative participative approach was taken when developing the user interface and both experienced and novice users where brought into the design process. Firstly, a review of existing instruction and graphical interfaces for biometric systems was carried out and six people were invited to use the fingerprint sensor with a legacy interface. Participants were then asked to volunteer ideas or ways to improve the instruction, interface or overall design of the system. A number of these interface design ideas were then explored using low fidelity paper prototypes. These low fidelity prototypes were then evaluated with a further six people, including 2 individuals from the first group of participants. Numerous issues with the interface were discovered during this design and evaluation process, such as a lack of understanding about the difference between enrolment and verification and the need for graphical instruction. A combination of the designs that were viewed most favourably by participants, and the approaches the designers believed to be most appropriate, were then developed in functional prototypes. Images of the low and high fidelity designs can be seen below in Figure 1.

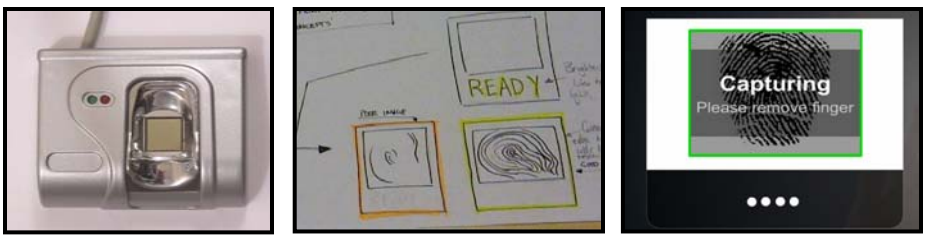


Fig. 1. The fingerprint sensor used in this evaluation (left). Low and high fidelity design prototypes.

During the design process, two different types of instructional lead-through were developed into fully functional prototypes. Both instructional approaches were designed to assist people with correct finger placement when using the fingerprint sensor. The first focused on which part of the finger should be placed on the sensor, as placing the tip of the finger on the sensor is a common behaviour that leads to poor quality images. The second method focused on using tactile features of the sensor as a finger placement guide. The fingerprint system used had a defined ridge at the lower edge of the sensing surface, designed to sit under the crease of the distal interphalangeal joint, when a finger is placed on the sensor. The tactile instruction emphasised this ridge and encouraged people to use it to assist finger alignment. A decision was made to make the lead-through animated, as the work of Theofanos et. al. [14] suggests that static pictorial instruction is less effective than video instruction. Each instructional video was 20 seconds in duration. Still images from both types of instructional lead-through can be seen below in figure 2 below.

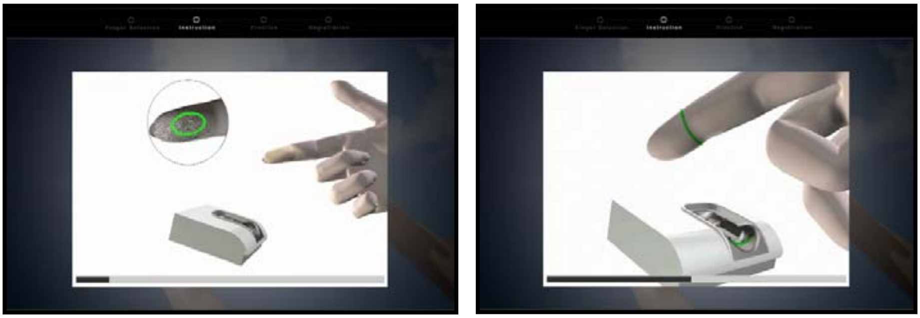


Fig. 2. Images of the lead-through focusing on finger placement (left) and tactile cues (right)

Two different approaches to the presentation of feedback were also identified during the development process. Displaying an image of the user's fingerprint on-screen emerged as one way to assist users during image acquisition. A graphical representation of the quality of the fingerprint image also received positive feedback from people involved in the design process. The image quality feedback used here was similar to the quality feedback used by Theofanos et. al. [17]. A third feedback approach was also developed into a functional prototype, a combination of pictorial and quality feedback. Here, a pictorial image of the users fingerprint is displayed on screen along with an associated measure of quality for that image. These three types of feedback can be seen in figure 3 below, and will be referred to as *pictorial, quality* and *combined* feedback for clarity. All feedback was displayed during image acquisition in near real time. The quality feedback was based an algorithm specific measure of the quality of a fingerprint image.

Fig. 3. Images of different feedback approaches: *Pictorial* image of users fingerprint (left), image *quality* (centre) and *combined* pictorial plus quality (right)

3 System Evaluation

The functional prototypes underwent a further evaluation to determine which type of lead-through and which type of feedback are the most appropriate for commercial applications of fingerprint technology. A no instruction condition was also tested to understand how helpful the instructional graphics actually were. The different

Table 1. Number of participants who used each type of instruction and feedback. In total 9 different prototype interfaces were tested.

	Pictorial Feedback	Quality Feedback	Combined Feedback	Total
Finger Placement	9	9	9	**27**
Tactile Instruction	9	9	9	**27**
No Instruction	9	9	9	**27**
Total	**27**	**27**	**27**	**82**

instruction and feedback approaches were counterbalanced to produce nine different evaluation interfaces. The structure of the evaluation and the number of participants who used each interface can be seen in table 1 below. Each participant used only a single instruction/ feedback combination in an independent groups design. A repeated measures design was not used as assessing the impact of instruction would be problematic if learning affects were present. The functional prototypes were tested with a group of people who had not been involved in the design process.

During the evaluation participants were asked to follow the instructions on screen, enrol and subsequently verify their identity using the fingerprint system. Four images were captured by the system during the enrolment process. Participants were then asked to use the system five times to verify their identity. A number of dependent measures were recorded during system use including image acquisition time, a measure of image quality and a matching score. The fingerprint system was tested with the default image acquisition and matching settings. During system usage the experimenter provided no assistance and only stepped in if the participant became stuck or experienced significant difficulty. After using the fingerprint system, participants completed a questionnaire designed to collect subjective information about their experience using the system. Finally, a brief semi-structured interview concluded the evaluation.

4 Results

A total of 82 people took part in this evaluation, with 27 people experiencing each of the instruction and feedback approaches. Participants ranged in age from 18 to 62 years with a mean age of 26.1 years. 51 of the participants were female and 30 were male, with gender information not recorded for one participant. All participants were recruited from a local university and were a mixture of students and staff. Participation in this study was voluntary, though people were rewarded for taking part.

4.1 Instruction

The different types of instruction did not have a significant effect on the overall performance of the fingerprint system. One way ANOVA tests revealed that there was no

Table 2. Summary of performance metrics across the different instruction conditions. Measures of time are given in seconds. Standard deviations in parenthesis.

	Mean enrolment quality	Mean enrolment time	No. of failures to enrol	Mean verification quality	Mean verification time	Mean matching score	No. of false rejection
Finger placement	248.0 (18.9)	24.0 (19.5)	0	247.6 (18.6)	4.1 (3.2)	457.3 (229.0)	7
Tactile Instruction	247.4 (13.7)	21.9 (17.8)	1	245.7 (21.8)	5.2 (3.1)	423.4 (162.2)	7
No Instruction	237.8 (40.2)	36.7 (44.2)	2	247.1 (23.7)	5.0 (4.9)	472.9 (229.5)	7

significant effect of instruction on any of the enrolment or verification metrics recorded by the system. Table 2 summarizes the performance of the three different types of instruction.

4.2 Feedback

The different methods of feedback affected the overall performance of the fingerprint system. There was a significant difference in average image quality during enrolment across the three feedback conditions. Assumptions of parametric testing were not met, so non-parametric tests were used. A Kruskal-Wallis test revealed that there was a significant effect of feedback ($H(2) = 8.45$, $p < .05$) on image quality. Bonferroni corrected post hoc testing revealed that there was a difference between the *pictorial* and *quality* feedback approaches ($U = 242$, $p < .0167$), but no other differences. Figure 4 below shows the differences in quality scores for the three types of feedback. It should also be noted that all three failures to enrol occurred where participants did not have pictorial feedback, though this data is not suitable for hypothesis testing.

There was also a significant difference in the average matching scores across the feedback conditions as revealed by a one way independent ANOVA ($F(2,77) = 4.97$,

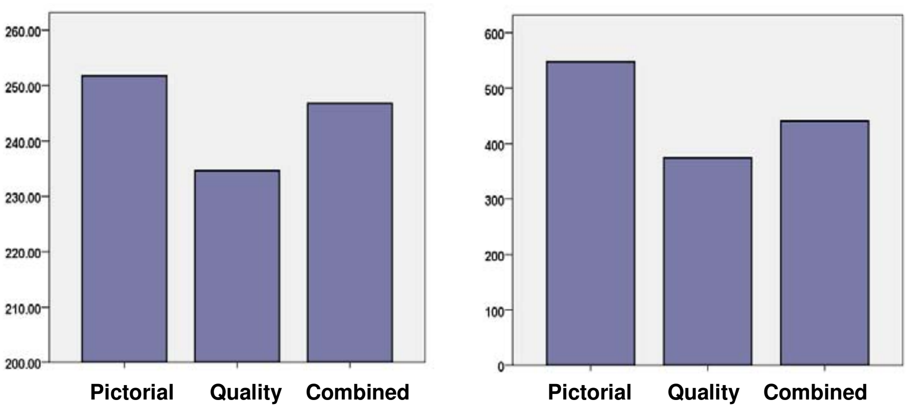


Fig. 4. Graphs of average image quality during enrolment (left) and average verification score (right) across the different feedback conditions

Table 3. Summary of performance metrics across the different feedback conditions. Measures of time are given in seconds. Standard deviations in parenthesis.

	Mean Enrolment Quality	Mean Enrolment time	No. of failures to enrol	Mean verification quality	Mean verification time	Mean matching score	No. of false rejection
Pictorial Feedback	**251.8 (9.8)**	25.8 (28.1)	0	249.1 (19.2)	4.9 (3.6)	**547.5 (216.5)**	3
Quality Feedback	**234.6 (37.1)**	25.9 (24.0)	3	245.9 (16.6)	4.3 (3.1)	**374.5 (184.2)**	13
Combined Feedback	**246.7 (24.3)**	31.3 (38.0)	0	245.2 (26.5)	5.1 (5.6)	**440.5 (191.8)**	3

$p < .01$). The matching score is a statistics reflecting the similarity of verification images to the enrolment template and is used to determine the match/ no match result. Bonferroni corrected post hoc testing revealed that the *pictorial* feedback preformed significantly better than the *quality* feedback approach ($t(48) = 2.73$, $p < .0167$). There were no other differences in matching score between the different types of feedback. The mean matching scores are illustrated in figure 4 below. Table 3 below summarizes the performance statistics across the three types of feedback. Interaction effects between instruction and feedback were also tested in a two way ANOVA, for all dependant variables, but none were found to be significant.

4.3 User Perception of the System

After using the fingerprint system participants completed a short questionnaire. Questions about the ease of use, speed, security, acceptability, aesthetics, privacy impact and clarity of feedback and instructions were included. A number of questions about participants' wiliness to use biometric systems in the future were also included. Figure 5 below shows participants' average ratings of system ease of use, privacy

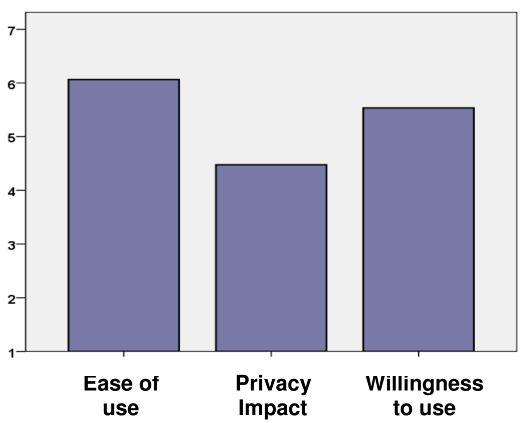


Fig. 5. Participants' mean ratings of system ease of use, privacy impact and willingness to use biometrics across all participants. Questions used a rating scale where 1 was negative and 7 was positive.

impact of biometrics and willingness to use biometrics across all feedback and instruction conditions. Overall, participants had a positive perception of the fingerprint system and the mean scores for all questions were towards the positive end of the scale. Kruskal-Wallis tests revealed that participants' perception of the fingerprint system was not affected by the different interface designs. There were no differences in participants' answers for any question, across the three types of instruction. There were also no significantly differences in participants' opinion towards the system across the different types of feedback.

4.4 Demographic Effects

Tests were performed to identify any demographic effects present in the results. There were no significant effects of age, gender, height, handedness or previous experience with biometrics on system performance. Age and gender were also compared with results from all questionnaire responses. In all 20 tests were carried out and one relationship was statistically significant. One test would be expected to return a significant result at the .05 level (type I error) over this number of tests however, so this no effects are reported.

5 Discussion

5.1 Effect of Instruction

The instruction used to support use of the fingerprint system did not have an effect on overall system performance in this study. It was though that there would be considerable differences in the way the system was used with different levels of instruction. Based on the work of Coventry [2] and Sasse [18] this would have seemed to be a reasonable prediction to make. There were no differences between the two types of graphical lead-through and no difference between the presence and absence of instruction seen in any of the dependant measures recorded. This result would seem to be at odds with the previous work by Theofanos et. al. [14] who found clear differences between instruction models. This study compared the semantic content of instructions rather than the way the instructions were delivered, so the results are not directly comparable. However, the lack of difference between the presence and absence of instruction suggests that the instruction had very little effect.

One possible explanation of the lack of effect of instruction, is that participants did not understand the graphical instructions. After using the fingerprint system participants were asked about the clarity of instruction and feedback. Almost all participants described the instruction as clear or easy to understand. Some participants demonstrated they understood the instruction and comments such as the following were not uncommon:

"It was telling me to line my finger up with that ridge so I did that…" A second explanation that is more consistent with participants' comments, is that people expected the system to be easy to use and so were less inclined to attend to the instructions. For instance, one participant tried to swipe their finger across the sensor despite the animated lead-through showing a finger being placed on the sensor. It is not unusual for people to ignore instructions, and Sasse [19] has described situations where

users have ignored instructions for other authentication and security systems. Comments from participants such as:

"The video at the start was too long" were also common. Overall, the transaction was comparatively simple; participants had to enrol and validate with only one finger, so the level of instruction may have exceeded the complexity of the task. If the different instructions had been tested in a more challenging environment, differences between instruction types may have emerged. This study however, where most participants did not experience significant problems using the system, was not sensitive enough to detect any difference between the presence and absence of instruction.

5.2 Effect of Feedback

The different types of feedback presented to users during the image acquisition had a clearer effect on the performance of the fingerprint system. Feedback had a significant effect on the quality of images captured during enrolment and the overall verification performance was also different. Feedback based on the *quality* of the fingerprint image resulted in a lower average matching score than *pictorial* feedback of the users own fingerprint. Lower matching scores mean that users are more likely to be falsely denied access to the system. All cases of participants failing to enrol with the system also occurred when people had no pictorial feedback. From these results it seems reasonable to conclude that the *pictorial* feedback is preferable to feedback relating to the *quality* of the biometric sample.

Feedback about the quality of a submitted image has been shown to improve users' interaction with fingerprint systems in the past [17]. This study did not test *quality* feedback against no feedback at all, but it did show that *pictorial* feedback is more helpful to users than information about *quality* only. An assertion that underpinned Theofanos et. al. [17] study was that normal users were not good at visually interpreting an image of their fingerprint onscreen and adjusting their behaviour. Participants in this study described the *pictorial* feedback positively and it appeared that people were able to interpret the *pictorial* feedback and adjust their behaviour accordingly during image acquisition. Inconsistent finger placement between enrolment and verification is one of the main reasons for false rejection, and *pictorial* feedback seemed to make people more aware of this issue.

The fingerprint technology used in this study and the Theofanos et. al. [17] study was not the same. The sensor used here was a small, direct contact, silicon sensor with a sensing area of 14mm x 14mm. A larger, optical sensor was used in past research [17], which could capture images from multiple fingers at a time. Though *Pictorial* feedback proved to be useful in this evaluation, it is likely that larger fingerprint sensors are more tolerant of inconsistent finger placement and *pictorial* feedback may not be as useful in this situation. Although participants' comments suggest *pictorial* feedback was also useful when determining how pressure affected the image quality, so *pictorial* feedback may still be useful for larger optical fingerprint systems.

At the start of this study it was hypothesised that combining *pictorial* and *quality* style feedback would be the easiest for people to interpret, and would consequently lead to better performance. This was not the case however, as there was no statistically significant difference between *pictorial* feedback and the *combined* feedback approach. Across most measures, there was a trend of *pictorial* feedback performing better than the *combined* feedback. A possible explanation for this result is that the *combined* feedback approach was too busy or complicated for people to usefully interpret. Having both an image of the fingerprint and a measure of quality on screen, each updating several times a second, could have been too much for people to attend to. Displaying two types of feedback may have caused participants to divide their attention between the two information sources, with a corresponding deterioration in performance. If this was the case a different design, with the two information sources more closely integrated may not have suffered the same drop in performance. Alternatively, the *combined* feedback could have been too information rich for people to use. Our aim of making the fingerprint system as usable and accessible as possible may have resulted in an overcomplicated interface, relative to the task. The image acquisition process is short, typically lasting only a few seconds, so any feedback presented to people must be simple and easy to understand.

5.3 User Perception of Biometrics

In general the people who took part in this study had a positive opinion towards the fingerprint system. People rated the system as easy to use and described themselves as willing to use biometrics again in the future, however some people in this study were concerned about the privacy impact of biometrics. This question about privacy received the lowest rating overall and the mean score was just above the scale midpoint. This result is consistent with previous research which has identified privacy concerns about biometric technology [20, 21]. It is worth noting that peoples' opinion of the fingerprint system were not affected by the interface they used; neither instruction nor feedback had an effect on participants' ratings of biometrics. This suggests that it is difficult to influence peoples' opinion towards biometrics by altering the design of the interface.

5.4 Participative Design Process

A participative design approach was taken when developing this interface. People were brought into the design process at several stages and a number of different lead-though and feedback approaches were discussed. In general, it proved difficult for people to articulate and describe the problems they encountered when using the fingerprint system, or to provide suggestions for improvement. Most of the people we talked to thought that fingerprint systems would be easy to use and this perception persisted throughout discussions of system design. This interface was essentially designed to support a single behaviour – placing one's finger on a fingerprint sensor. This would seem to be a simple task and the difference in the biomechanical movement between good and poor finger placement is very small. In our experience, it was difficult to engage people in discussion about this particular issue. People were happy to volunteer their thoughts and feelings about the applications for, and suitability of biometric systems, but it proved difficult to engage people when discussing this narrow aspect of interface design. The participative design process yielded richer information for wider issues such as the acceptability of biometrics, rather than issues like instruction and interface design.

6 Conclusions

Designing usable biometric systems is a challenging task. Ensuring that people can use systems effectively on first use remains an issue for biometric authentication systems. The results of this evaluation show that the design of the interface is essential element of usable fingerprint systems and the way feedback is presented effects overall system performance. Based on the results of this study we make the following design recommendations for fingerprint systems:

- Displaying pictorial feedback is beneficial for people using a small fingerprint sensor.
- Provide a graduated level of assistance to users. Instruction proved unnecessary for most people, but is helpful for those who do experience difficulties. Provide more instruction and guidance if people struggle during the image acquisition process.
- Keep the information on screen to a minimum. For most, the interaction with a fingerprint sensor is very brief, so people do not have time to process large amounts of information.
- Ensuring a high quality enrolment is an essential aspect of biometric authentication system design.

The challenges inherent in implementing biometric technology are one reason why biometric systems have not received wider uptake despite their reputed advantages. Designing a biometric system that can be used by the general public, without providing assistance, would have a significant benefit and would increase the number of applications and contexts were biometrics could usefully be used.

A usable interface for biometric systems is only part of the issue however. Many people are genuinely concerned about data security and the use of biometric systems, and these concerns must be addressed if any implementation of biometrics is to be successful. These issues are much broader than the user interface, and further research should address how to effectively convey information about data storage, data access rights and security policies to the people who will use biometrics. Making biometric systems usable is an essential element of system design. However, biometrics must also be acceptable for the people who use them and this issue has not received the attention it deserves.

Acknowledgements. We would like to thank Jamie Shek for his invaluable help during the design stages of this project. We would also like to thank Andrea Szymkowiak and Jim Bown from the University of Abertay Dundee for their support of this research. Finally, we are indebted to the Scottish Executive and Technology Strategy Board U.K. for their sponsorship of this project through the Knowledge Transfer Partnership scheme.

References

1. Renaud, K.: Evaluating Authentication Mechanisms. In: Cranor, L.F., Garfinkel, S. (eds.) Security and Usability. O'Reilly, Sebastopol (2005)
2. Coventry, L.: Usable Biometrics. In: Cranor, L.F., Garfinkel, S. (eds.) Security and Usability. O'Reilly, Sebastopol (2005)

3. International Biometrics Group (IBG): Biometrics Market and Industry Report 2009-2014 (2009), `http://www.biometricgroup.com/reports/public/market_report.html`
4. Chandra, A., Calderon, T.: Challenges and Constraints to the Diffusion of Biometrics in Information Systems. Communications of the ACM 48(12), 101–106 (2005)
5. Maple, C., Norrington, P.: The Usability and Practicality of Biometric Authentication in the Workplace. In: First International Conference on Availability, Reliability and Security. IEEE Computer Society, Los Alamitos (2006)
6. Celent.: Biometric Technologies: Are We There Yet? Celent Communications, Boston (2006)
7. Jain, A., Hong, L., Pankanti, S.: Biometric Identification. Communications of the ACM 43(2), 91–98 (2000)
8. Ashbourn, J.: Biometrics: Advanced Identity Management. Springer, Heidelberg (2000)
9. Patrick, A.: Usability and acceptability of biometric security systems. In: NATO Workshop on Enhancing Information Systems Security Through Biometrics, Ottawa, Canada (2004)
10. Coventry, L., DeAngeli, A., Johnson, G.I.: Biometric Verification at a Self Service Interface. In: Contemporary Ergonomics 2003, pp. 247–252 (2003b)
11. Coventry, L., Angeli, A.D., Johnson, G.I.: Honest It's Me! Self Service Verification. In: The Proceedings of the CHI Conference on Human Factors in Computing Systems, Workshop on Human-Computer Interaction and Security Systems, Fort Lauderdale, Florida (2003a)
12. Miller, R. C., Chou, V. H., Bernstein, M., Little, G., Kleek, M. V., Karger, D., et al.: Inky: A Sloppy Command Line for the Web with Rich Visual Feedback. In: Symposium on User Interface Software and Technology, Monterey (2000)
13. Garfinkel, S.L., Miller, R.C.: Johnny 2: A User Test of Key Continuity Management with S/MIME and Outlook Express. In: Symposium on Usable privacy and security (SOUPS), Pittsburgh (2005)
14. Theofanos, M., Stanton, B., Orandi, S., Micheals, R., Zhang, N.-F.: Ten-Print Fingerprint Capture: Effect of Instructional Modes on User Performance. In: Human Factors and Ergonomics Society 51st Annual Meeting (2007)
15. Hoffman, P.: Visualizing the Use of Biometric Systems: Tackling Symbolism, Context and Interpretation. In: International Workshop on Usability and Biometrics, Washington (2008)
16. Ormiston, G.: Addressing the Worldwide Biometric Enrolment Challenge: Guidance Without Words. In: International Workshop on Usability and Biometrics, Washington (2008)
17. Theofanos, M., Micheals, R., Scholtz, J., Morse, E., May, P.: Does Habituation Affect Fingerprint Quality? In: The CHI Conference of Human Factors in Computing Systems, Montréal, Canada (2006)
18. Sasse, A.: Red-Eye Blink, Bendy Shuffle, and the Yuck Factor: A User Experience of Biometric Airport Systems. IEEE Security & Privacy, 78–81 (2007)
19. Sasse, A.: Usability and trust in information systems. Cyber Trust & Crime Prevention Project. University College London, London (2004)
20. BioSec.: Report on results of first phase usability testing and guidelines for developers. BioSec Consortium (2004)
21. Toledano, D.T., Pozo, R.F., Trapote, A.H., Gomez, L.H.: Usability evaluation of multimodal biometric verification systems. Interacting with Computers 18, 1101–1122 (2006)

Measurement of Olfactory Characteristics for Two Kinds of Scent in a Single Breath

Kaori Ohtsu[1], Junta Sato[1], Yuichi Bannai[2], and Kenichi Okada[1]

[1] Faculty of Science and Technology, Keio University
Hiyoshi 3-14-1, Kohoku-ku, Yokohama, Kanagawa, 223-8522 Japan
{ohtsu,sato,okada}@mos.ics.keio.ac.jp
[2] Global Environment Promotion Headquarters, Canon Inc.
Shimomaruko 3-30-2, Ohta-ku, Tokyo, 146-8501 Japan
bannai.yuichi@canon.co.jp

Abstract. This study describes a presentation technique of scent designed for users to recognize multiple scents during a very short time period. We measured the olfactory characteristics of subjects when two kinds of scents were presented in a single breath. We defined and measured the minimum ejection interval in which subjects could discriminate the two individually emitted pulses of scent, which we term "separable detection threshold", and the minimum ejection interval in which they could specify both kinds of scents, "separable recognition threshold". Further, "response time" and "duration of scent perception" were measured. As a result, we found the duration of scent perception and the separable recognition threshold were positively correlated. Knowledge of this olfactory characteristic brings us closer to being able to provide a greater sense of realism in multimedia environments, by describing more than one object by scent at the same time as the objects are seen on screen.

Keywords: Olfactory Information, Olfactory Display, Pulse Ejection, Olfactory Characteristics.

1 Introduction

Information and communication via computers tends to be limited to visual information and audio information. However, in the real world, humans gather external information via the five senses of sight, hearing, touch, smell and taste, allowing them to react appropriately to local circumstances. Accordingly, the conveyance of such information and its communication via the five senses has lately attracted much attention [1]. Olfactory information presented to match visual information has been shown to have a dramatic effect on deepening the viewer's understanding of the image content [2]. In addition, in the case of viewers watching a movie accompanied by scent emission, the relation between the scent and the viewer's feelings has been analyzed by measuring the viewer's brain waves and estimating his or her psychological state [3]. The presentation of olfactory information is considered to be an effective means for enhancing the sense of reality, similarly to the use of three-dimensional images and sounds [4]. Therefore, trials on the transmission of olfactory information together

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 306–318, 2009.

with that of audio/visual information are currently being conducted in the field of multimedia.

In order to transmit scent together with other media, it is necessary to control the presentation of scent in accordance with changes in the images/sounds over time. However, existing techniques have not yet overcome the problem of emitting too much odor over a continuous period, which causes various problems such as olfactory adaptation and the lingering of odors in the air, making it difficult to synchronize scents with the ever changing images and sounds. In efforts to resolve these problems, we have focused on reducing the amount of scent emitted by using pulse ejection for a very short period of time [5]. In the viewing of movie scenes especially, it is thought that the sense of realism is increased by delivering scent adapted to the specific images being viewed. Frequently, multiple objects are shown at the same time in scenes, but a technique for presenting multiple scents during a very short time period has not been reported to date.

During measurement of olfactory characteristics using a pulse ejection technique which was designed to avoid scent remaining in the vicinity of the user and thus prevent olfactory adaptation, we found that humans can recognize a switch between two kinds of scents in a single breath. Therefore, in this study, in efforts to develop a presentation technique of multiple scents in a single breath, we measure the olfactory characteristics when two kinds of scents are presented in a single breath, and report the results here.

2 Related Work

Trials on the transmission of olfactory information together with audio/visual information are ongoing. Work first started in the 1950s when Heilig developed Sensorama [6], the first virtual reality (VR) system that presented olfactory information together with audio/visual information. The recently developed virtual space system, Friend Park [7], provides users with an increased sense of reality by generating the aroma of a virtual object or environment, where the aroma is defined as the area in which a scent can be perceived. Kaye's article [8] describes some systems that add scent to web content, and computer controlled olfactory displays such as iSmell [9] and Osmooze [10] are utilized in these systems. Another type of display, the air cannon olfactory display that generates toroidal vortices of scent in order to present it in restricted space, has also been proposed [11].

In research on the transmission of olfactory information, Nakamoto et al. [12] designed a smell synthesis device that presents the scent of a virtual object remotely. The system analyzes the smell to be transmitted and presents the analyzed data as the composition ratio of the scent elements. On the receiver side, a feedback control changes the ratio of the scent elements owned by the receiver in order to reproduce the target scent. However, the system can not yet handle arbitrary scents.

A wearable olfactory display with a position sensor has also been developed [13]. By controlling the density of odor molecules, it can present the spatiality of olfaction in an outdoor environment. The olfactory information transmitting system consists of the aforementioned display, a sensing system using three gas sensors, and matching database. The user can experience a real sense of smell through the system by translating the olfactory information obtained.

AROMA [14] tries to introduce the olfactory modality as a potential alternative to the visual and auditory modalities for messaging notifications. Experimental findings indicate that while the olfactory modality was less effective in delivering notifications than the other modalities, it had a less disruptive effect on user engagement in the primary task.

The addition of a scent to image media such as movies has been proposed by a number of researchers. Okada et al. [3] measured the viewer's mental state by his or her brainwaves, and analyzed the relation between the scent and the viewer's feelings while watching. A movie that adds olfactory information to the visual/audio information has been created, but since the synthetic scent did not consistently accord with the image and the scent was not deodorized, the movie could not be widely distributed.

3 Characteristics of Olfaction

A fragrance substance is a compound that stimulates the olfactory cells in the nasal cavity. Fragrance substances can be inorganic substances such as hydrogen sulfide and ammonia, although most are organic compounds. It is said that of the approximately two million kinds of organic compounds in existence, about four hundred thousand of these have an odor [15]. However, humans perceive and recognize about five thousand scents routinely. The characteristics of human olfaction are now briefly described.

3.1 Olfactory Threshold

The olfactory threshold is the value used as a standard to express the strength and weakness of a scent. Three kinds of values are generally used for the olfactory threshold: the detection threshold, the recognition threshold, and the differential threshold [16], usually expressed in units of mol (concentration) and mass percentage. However, because the olfactory threshold is a measure of the lowest olfactory stimulus intensity at which an individual can perceive scent, this value does not reflect the intensity (strength and weakness) of the scent perceived.

- *Detection threshold*: the smallest density at which scent can be detected and where the user does not need to recognize the kind of smell.
- *Recognition threshold*: the smallest density at which the kind of scent can be recognized, and its value reflects the ability of the user to express quality and characteristics of the scent.
- *Differential threshold*: the density at which the user can distinguish the strength of a scent, where its value reflects the ability of the user to detect changes in the stimulus and to quantify the change.

  Generally such changes are expressed as the % change of stimulation quantity of the original. In the case of olfaction, it differs with different kinds of scent, but is in the range of about 13-33%.

3.2 Adaptation

Adaptation is the phenomenon where sensory nerve activity is decreased by continuous stimulation by odor molecules. Adaptation itself and the speed of recovery from adaptation differ according to the kind of scent. Adaptation is gradually strengthened over time, but is restored for a short time (3-5 minutes) by eliminating the scent.

In addition, there are various patterns of adaptation, influenced by the kind of scent and recognition factors.

4 Olfactory Characteristics in Relation to Two Kinds of Scent in a Single Breath

Previously, while testing a pulse ejection method that ensures scent does not remain in the vicinity of the user, we found that humans can recognize a switch between two kinds of scents in a single breath. Therefore, the purpose of this study was to develop a presentation technique that enables users to recognize multiple scents during a very short time period by measuring the olfactory characteristics of subjects.

Using the pulse ejection technique, the ejection time interval between two pulse ejections was about 1 second. It is known that a typical person breathes about 12 times/min and inhales for an average of 2 sec at rest [17]. This means that a third pulse ejection is not presented during inspiration in most cases, so many people cannot perceive the scent presented. Therefore, we limited the scents presented in a single breath to two kinds, and measured the time characteristics of olfaction. Figure 1 depicts the two pulse ejections of scent individually emitted in a single breath, where the ejection time interval of two pulse ejections is defined as T_i.

When scent is emitted as depicted in Figure 1, humans cannot discriminate the two individually emitted pulses of scent when the interval T_i is small. Naturally, this applies in the case of the emission of two different scents, one at each pulse. Thus, we defined the minimum ejection time interval in which the subject could discriminate the two individually emitted pulses of different kinds of scent as the "separable detection threshold" and the minimum ejection time interval in which the subject could specify both kinds of scents as the "separable recognition threshold", and measured them.

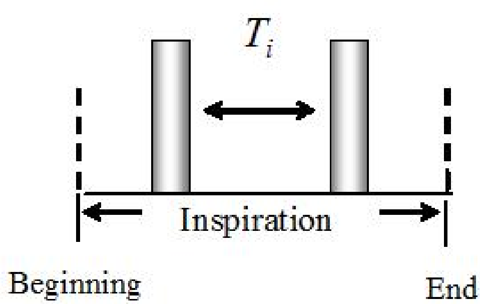


Fig. 1. Ejection time interval of two pulse ejections "T_i"

The separable detection threshold of scent was defined on the basis of the separable detection threshold of pain. When touching two nearby points on the skin with a sharp tipped object, at a certain distance we perceive the two points at just one location and beyond this point we perceive them at separate locations. This minimum distance to perceive pain at two locations is known as the separable detection threshold of pain.

When this separable detection threshold is defined as “the threshold at which a subject can perceive two concomitant stimuli of pain (via the sense of touch) as separate stimuli”, the separable detection threshold in regard to other senses can be also defined. The separable detection threshold is classified into spatial distance and temporal distance: the above separable detection threshold of pain is spatial while that of smell is temporal. Therefore, we defined the separable detection threshold in regard to the sense of smell as “the time interval in which a subject can discriminate two individually ejected pulses of different scents”.

During measurements of the separable detection threshold and separable recognition threshold in regard to the sense of smell, 100-msec, stable pulse ejections of the scents were presented. Measurements were then made to determine the “response time” and “duration of scent perception”, which were defined respectively as follows: the time period from when the olfactory display started pulse ejection until when the subject started to perceive the scent, and the time period from when the subject started to perceive the emitted scent until when he or she no longer perceived it.

5 Olfactory Display and Performance Evaluation

5.1 Olfactory Display

We developed an olfactory display in conjunction with Canon Inc. Figure 2 shows the prototype olfactory display. This display is ink-jet in order to produce a jet which is broken into droplets from the small hole in the ink tank. Ejection control is possible for a unit of 100 msec.

Figure 3 is a plain view of the olfactory display. The display can set up 3 scent ejection heads. Since each head can store one large tank and 3 small tanks, the display can present, in total, 12 kinds of scents utilizing 3 large tanks and 9 small tanks. There are 256 minute holes in the head connected to the large tank and 128 in the head connected to the small tank. The display can emit scent from multiple holes for a period of 100-msec, and the ejection concentration can be varied using between 0 to 255 holes for the large tank and 0 to 127 holes for the small tank. We denote the average ejection quantity from each minute hole as the “unit average ejection quantity (UAEQ)”, and the number of minute holes emitting during a 100-msec ejection period as “the number of simultaneous ejections (NSE)”.

For all measurements, we used only the small tanks. The scents used were lavender, lemon and peppermint. The UAEQ from the small tanks is 4.7 pl for lavender scent, 3.7 pl for lemon scent and 4.7 pl for peppermint scent. Examination at Canon Inc. confirmed the quantity to be approximately constant without depending on the residual quantity of ink. In addition, the user can set the number of ejection times from one hole in the 100-msec period in the range of 1 to 150 times, which we denote the

Fig. 2. Olfactory display

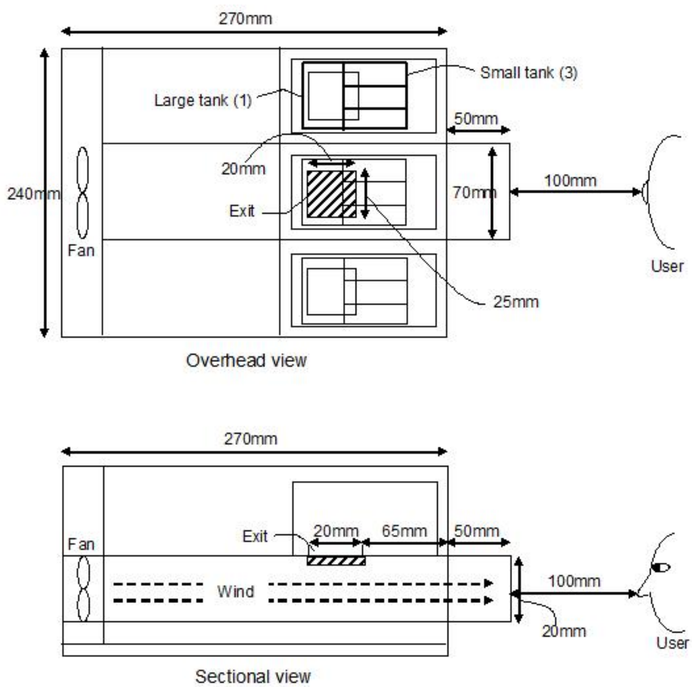


Fig. 3. Plain view of the olfactory display

"volume". In this study, the volume was set to 150 times without variation. Therefore, the ejection quantity in the 100-msec period (EQ) is calculated as follows.

$$\text{EQ (pl)} = 4.7 \text{ or } 3.7 \text{ (pl: UAEQ)} \times 0\text{-}127 \text{ (NSE)} \times 150 \text{ (times: Volume)} \quad (1)$$

The scent is diluted by 5% with ethanol and water.

$$\text{Scent quantity (pl)} = \text{EQ (pl)} \times 0.05 \quad (2)$$

Fig. 4. Use of the olfactory display

The display is equipped with a fan and there are 10 phases of wind velocity control in the range of 0.8 m/sec to 1.8 m/sec. The scent presentation hole is a rectangle of 2 cm height and 24 cm width.

Figure 4 is a photograph showing the use of the olfactory display. The user places the chin on the chin rest, fixing the distance from the olfactory ejection point to the nose.

5.2 Continuance Time of Pulse Ejection

If more than the required amount of scent is emitted, it remains in the air and can cause adaptation. Therefore, it is desirable that the continuance time of pulse ejection is the shortest for which an individual can sense the scent without adaptation. As the shortest ejection continuance time of the olfactory display is 100-msec and all 20 subjects we recruited could recognize pulse ejection of 100-msec, we set the pulse ejection to 100-msec in this study.

5.3 Recognition Threshold

Measurements to determine the recognition threshold were conducted using 100-msec pulse ejections of three kinds of scent (lavender, lemon and peppermint) with 6 subjects.

Measurement was undertaken for between 5 and 40 simultaneous ejections which determine ejection concentration, in decrements of 5. Using the paired comparison method [18], we measured the recognition threshold of each scent. The olfactory display presented scented and unscented ejections to each subject, and we instructed the subject to indicate which of the two was the scented one and the kind of the scent. Ejection concentration was decreased until the subject could no longer recognize the scent.

Table 1 shows the measurement results of the recognition threshold. The maximum recognition threshold of 6 subjects was an ejection concentration of 30 NSE for three scents. Thus, all subjects could recognize a scent when the ejection concentration is set at more than 30. Therefore, NSE in the following measurement was set to 30.

Table 1. Recognition threshold (in 127 phases)

Kind of scent	Average	Maximum	Minimum
Lavender	23.3	30	15
Lemon	19.2	30	10
Peppermint	17.5	10	5

6 Measurement of the Temporal Characteristics of Olfaction

6.1 Response Time and Duration of Scent Perception

We measured the olfactory characteristics of pulse ejection with regard to "response time" and "duration of scent perception" in a single breath, with 24 subjects. Olfactory ejection was presented at the beginning of inspiration for each subject. In order to measure "response time", subjects were instructed to click a mouse when they perceived the scent. For each subject we measured the time from starting the ejection of scent to the mouse click three times for each scent, and recorded the average value for each scent. In order to measure "duration of scent perception", subjects were instructed to click a mouse when they could no longer perceive a scent. Similarly to the measurement of response time, for each subject we measured the time from starting the ejection of scent to the mouse click three times for each scent and recorded the average value for each scent. We then calculated the response time and the duration of scent perception as follows.

$$\text{Response time} = \text{Average time to beginning to perceive the scent} \quad (3)$$

$$\begin{aligned}\text{Duration of scent perception} = &\ \text{Average time to no longer perceiving the scent} \\ &- \text{Average time to the beginning to perceive the scent}\end{aligned} \quad (4)$$

Table 2 shows the measurement results for the response time and the duration of scent perception. In order to examine the effect of differences among scents and among individuals on the response time and the duration of scent perception, the results were analyzed using a two-way ANOVA (kind of scent, individual). The results for response time showed no significant differences for the interaction of the two factors ($F(46)=1.45$, $P>0.05$) or for the main effect of kind of scent ($F(2)=3.06$, $P>0.01$). However, a significant difference was found for the main effect of individual ($F(23)=1.60$, $P<0.01$); however, the individual difference was small since the average

Table 2. Response time and duration of scent perception

Kind of scent	Response time		Duration of scent perception	
	Average (sec)	Standard deviation	Average (sec)	Standard deviation
Lavender	1.08	0.14	1.26	0.48
Lemon	1.09	0.12	1.21	0.46
Peppermint	1.04	0.13	1.26	0.48
Average	1.07	0.13	1.24	0.47

of the standard deviation was as small as 0.13 (variation coefficient 12.0%). The results for duration of scent perception showed no significant differences for the main effect of kind of scent ($F(2)=3.20$, $P>0.05$). A significant different was found for the main effect of individual ($F(23)=1.77$, $P<0.01$) , which was large since the average of the standard deviation was as large as 0.47 (variation coefficient 37.7%).

6.2 Separable Detection Threshold and Separable Recognition Threshold

We measured the “separable detection threshold” and the “separable recognition threshold” with 24 subjects in the same manner as described above. “Scent scenarios” were presented in a single breath. The scent scenarios prepared were six scent patterns (ejection of different combinations of two of the total three scents) and six dummy patterns (three patterns where the scent of the two pulse ejections was the same and three patterns where the number of pulse ejections was only one). The measurement range of the interval between the two pulse ejections was set between 0.2 and 2.0 sec by 0.2 sec.

Measurement was performed as follows. A scent scenario with a given ejection interval was randomly chosen from 12 scenarios, and the scenario was executed by synchronizing the first pulse ejection with the beginning of inspiration. When each scent scenario was completed, subjects answered two questions: *“1. How many times did you notice the scent?”* and *“2. What was the kind of the scent?”* The first question was an open question and we judged that the subject could separate and detect the smells when answering “two times”, and the second question was a closed question. When both questions were answered correctly, we judged that the subject could separate and recognize the two scents presented. Measurement was repeated for all six of the scent patterns while changing the ejection interval. The measurement was finished when the result was stabilized (i.e. when the subject answered correctly more than twice for scent presentation at the ejection interval that was considered the shortest one) and we then recorded the shortest ejection interval at which each subject could detect as well as recognize the two distinct kinds of scent.

Table 3 shows the measurement results of the separable detection and separable recognition thresholds. To examine the effects of the differences among scent scenarios and individuals on the separable detection threshold, the results were analyzed

Table 3. Separable detection threshold and separable recognition threshold

Scent scenario		Separable detection threshold		Separable recognition threshold	
First	Second	Average (sec)	Standard deviation	Average (sec)	Standard deviation
Lavender	Lemon	0.78	0.41	1.02	0.49
Lavender	Peppermint	0.61	0.34	0.94	0.45
Lemon	Lavender	0.87	0.36	1.09	0.35
Lemon	Peppermint	0.60	0.29	0.81	0.46
Peppermint	Lavender	0.83	0.43	1.04	0.42
Peppermint	Lemon	0.80	0.36	0.95	0.35
Average		0.75	0.37	0.98	0.42

using a two-way ANOVA (kind of scent, individual). There was no significant difference for the main effect of kind of scent (F(2)=2.29, P>0.01). However, a significant difference was found for the main effect of individual (F(23)=1.62, P<0.01), which was large since the average of the standard deviation was as large as 0.37 (variation coefficient 48.9%).

In the case of the separable recognition threshold, some subjects could not separate and recognize two scents even if the ejection interval was set to the maximum time of 2.0 sec. Therefore, Table 3 shows the value determined excluding data from those subjects. Since the number of data items was not equal, the results were analyzed using a one-way ANOVA (kind of scent, individual). For the separable recognition threshold, there was no significant difference for kind of scent (F(2)=2.29, P>0.05), but a significant difference for individual (F(23)=1.63, P<0.01), which was large since average of the standard deviation was 0.42 (variation coefficient 43.1%).

6.3 Discussion

This study revealed that the kind of scent used had no influence on the measurement results obtained. On the other hand, differences were found in the measurement results among individuals, particularly with regard to duration of scent perception and separable recognition threshold. Therefore, we examined the correlation of these two values. A weak correlation (0.56) was found between duration of scent perception and separable recognition threshold. Figure 5 is the correlation diagram showing individual values and an approximate curve. There are some outliers, but a positively correlation is seen.

To consider this aspect in detail, we divided the 24 subjects into 3 groups of 8 based on the results of duration of scent perception: the short group, medium group, and long group. Figure 6 shows the duration of scent perception histogram for the 3 groups.

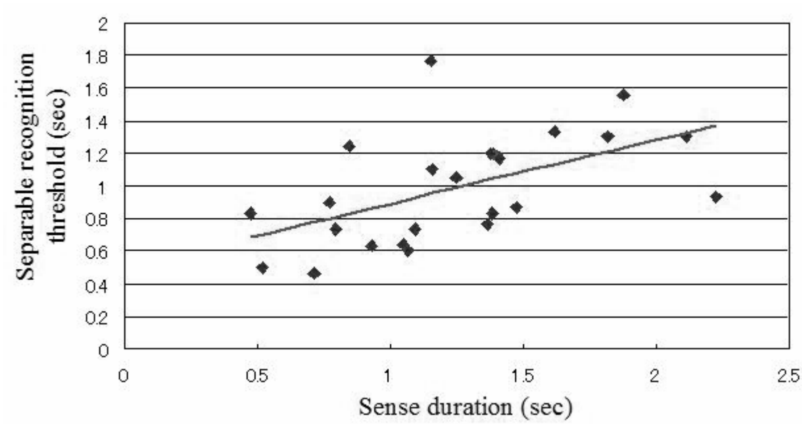


Fig. 5. Correlation diagram between duration of scent perception and separable recognition threshold

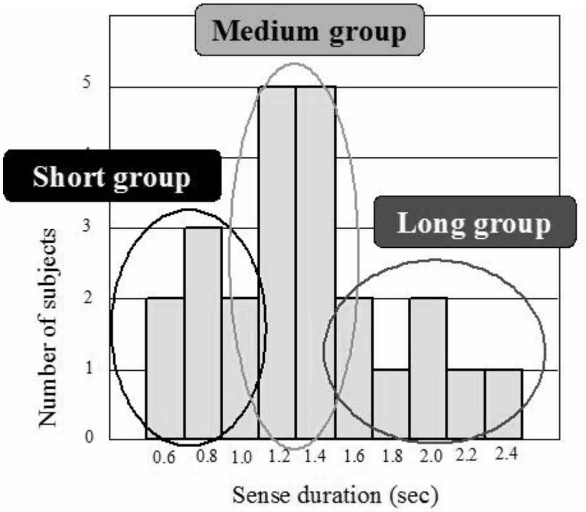


Fig. 6. Duration of scent perception histogram

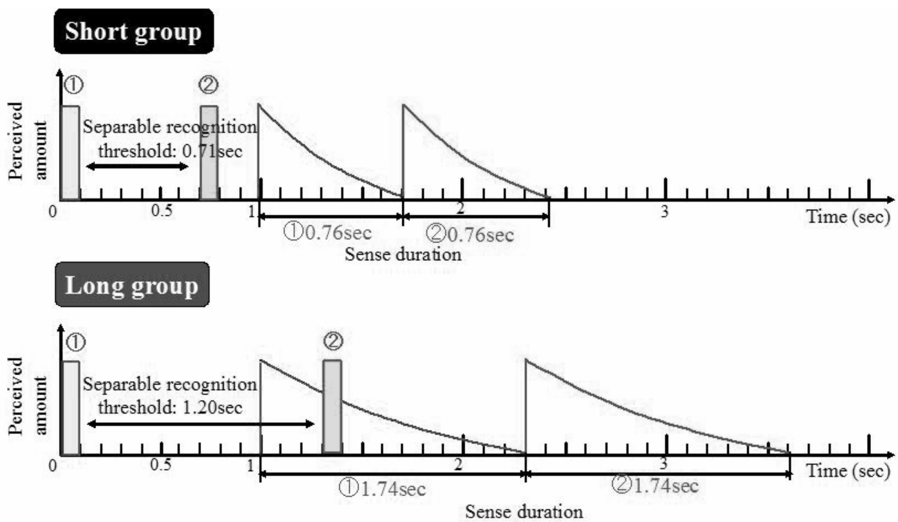


Fig. 7. Sensory perception differences for each group

Figure 7 shows the sensory perception results for subjects in the short and long groups. The horizontal axis represents time and the vertical axis represents the amount of scent perceived. The gray bars represent pulse ejections and the curve represents the change in the amount of scent perceived. This curve is drawn on the basis of results estimated by an olfactory model reported in an earlier study [19]. Overall, there were little individual differences in response time, and all subjects began to perceive the scent around 1 sec after pulse ejection started. Furthermore, as shown by the second curve in each graph, the subjects began to perceive the second scent just after they stopped perceiving the first scent.

There were several subjects for whom the two curves overlapped or were separated only slightly. The interval time between the two pulse ejections is the separable recognition threshold, and it appears that when pulse ejections were presented at an interval shorter than the subject's separable recognition threshold, he or she could not distinguish the two scents because the two curves overlapped greatly.

7 Conclusion

In order to realize a scent presentation technique that enables users to recognize multiple scents during a very short time period, we measured olfactory characteristics when two kinds of scents were presented by pulse ejection in a single breath. We measured "response time" and "duration of scent perception" in relation to the pulse ejections of scent. We were able to define and measure the "separable detection threshold" and the "separable recognition threshold" that was the interval between two pulse ejections. As a result, we found large individual differences in the duration of scent perception and the separable recognition threshold; however, these values were positively correlated. In general, it is said that the presentation technique could not handle well some large individual differences in olfactory characteristics, but the temporal characteristics of smell measured in this study did show a correlation among individuals.

By advancing the measurement of olfactory characteristic in future, we will be able to use scent easily. For example, olfactory displays could be customized in accordance with such characteristics. In addition, by setting the interval of pulse ejections based on the separable recognition threshold of the long group, all users should be able to perceive two kinds of scent in a single breath. Such technology brings us closer to being able to provide a greater sense of realism in multimedia environments by describing more than one object by scent at the same time as the objects are seen on screen.

Acknowledgment. This work was supported by SCOPE (Strategic Information and Communications R&D Promotion Programme) of the Ministry of Internal Affairs and Communications, Japan.

References

1. Kim, J.-D., Kim, D.-J., Han, D.-W., Byun, H.-G., Ham, Y.-K., Jung, W.-S., Park, J.-S., Oh, S.-K.: A Proposal Representation, Digital Coding and Clustering of Odor Information. Computational Intelligence and Security 2006 1, 872–877 (2006)
2. Tomono, A., Yamamoto, S., Utsunomiya, M., Ikei, D., Yanagida, Y., Hosaka, K.: Effect that the Image Media with Scent Gives to Contents Understanding. In: Human Interface Symposium 2004, pp. 249–254 (2004)
3. Okada, K., Aiba, S.: Toward the Actualization of Broadcasting Service with Smell Information. Institute of Image information and Television Engineering of Japan Technical Report (in Japanese) 27(64), 31–34 (2003)

4. Hirose, M., Tanigawa, T.: Information and Communication of Odor. Aroma Science Series 21 Wearable Olfactory Display (in Japanese), pp. 60–76. Fragrance Journal Ltd. (2007)
5. Kadowaki, A., Sato, J., Bannai, Y., Okada, K.: Presentation Technique of Scent to Avoid Olfactory Adaptation. In: ICAT 2007, pp. 97–104 (2007)
6. Retrofuture: Sensorama's pre-virtual reality, http://www.retrofuture.com/sensorama.html
7. Shigeno, H., Honda, S., Osawa, T., Nagano, Y., Okada, K., Matsushita, Y.: A Virtual Space Expressing the Scent and Wind -A Virtual Space System "Friend Park". Journal of Information Processing Society of Japan (in Japanese) 42(7), 1922–1932 (2001)
8. Kaye, J.: Making Scents: aromatic output for HCI. Interactions 11(1), 48–61 (2004)
9. Edge Review: DigiScent Ismell, http://www.edgereview.com/ataglance.cfm?category=Edge&ID=136
10. http://www.osmooze.com/
11. Yanagida, Y., Noma, H., Tetsutani, N., Tomono, A.: An Unencumbering, Localized Olfactory Display. In: CHI 2003 Extended abstracts, pp. 988–989 (2003)
12. Nakamoto, T., Nakahira, Y., Hiramatsu, H., Moriizumi, T.: Odor Recorder Using Active Odor Sensing System. Sensors and Actuators B 76, 465–469 (2001)
13. Yokoyama, S., Tanikawa, T., Hirota, K., Hirose, M.: Olfactory Field Simulation Using Wearable Olfactory Display. Trans. of Virtual Reality Society of Japan (in Japanese) 9(3), 265–274 (2004), http://www.cyber.rcast.utokyo.ac.jp/project/nioie.html
14. Bodnar, A., Corbett, R., Nekrasovski, D.: AROMA: Ambient Awareness through Olfaction in a Messaging Application. In: ICMI 2004, pp. 183–190 (2004)
15. Kawasaki T., Nakajima M., Tonoike M.: Aroma Science Series 21 Characteristics and Analytical Estimation of Smell Material (in Japanese). Fragrance Journal Ltd. (2003)
16. Term and Commentary of Odor (in Japanese), Odor Control Association of Japan (2003)
17. Tanaka, K., Kakizaki, H.: Theory and Technology of Breath Exercise Therapy (in Japanese), pp. 70–71. Medical View Co., Ltd. (2003)
18. Odor Simplified Measurement Guidebook 2005 (in Japanese), Japan Association for Odor Environment (2005)
19. Kadowaki, A., Sato, J., Bannai, Y., Okada, K.: The Response Model of Human Sense of Smell during Pulse Emission (in Japanese). Japan Association for Odor Environment 39(1), 36–43 (2008)

Keyboard before Head Tracking Depresses User Success in Remote Camera Control

Dingyun Zhu[1,2], Tom Gedeon[2], and Ken Taylor[1]

[1] CSIRO ICT Centre, Acton, Canberra, ACT 0200, Australia
[2] School of Computer Science, College of Engineering and Computer Science, The Australian National University, Canberra, ACT 0200, Australia
dingyun.zhu@csiro.au, tom.gedeon@anu.edu.au, ken.taylor@csiro.au

Abstract. In remote mining, operators of complex machinery have more tasks or devices to control than they have hands. For example, operating a rock breaker requires two handed joystick control to position and fire the jackhammer, leaving the camera control to either automatic control or require the operator to switch between controls. We modelled such a teleoperated setting by performing experiments using a simple physical game analogue, being a half size table soccer game with two handles. The complex camera angles of the mining application were modelled by obscuring the direct view of the play area and the use of a Pan-Tilt-Zoom (PTZ) camera. The camera control was via either a keyboard or via head tracking using two different sets of head gestures called "head motion" and "head flicking" for turning camera motion on/off. Our results show that the head motion control was able to provide a comparable performance to using a keyboard, while head flicking was significantly worse. In addition, the sequence of use of the three control methods is highly significant. It appears that use of the keyboard first depresses successful use of the head tracking methods, with significantly better results when one of the head tracking methods was used first. Analysis of the qualitative survey data collected supports that the worst (by performance) method was disliked by participants. Surprisingly, use of that worst method as the first control method significantly enhanced performance using the other two control methods.

Keywords: Head Tracking, Remote Camera Control, Human Computer Interaction, Teleoperation, Usability Evaluation.

1 Introduction

Teleoperation has been regarded as an essential application strategy and widely applied in modern industry because of a variety of advantages. A device or machine is remotely operated by a person from a distance, which is able to effectively move human workers away from hazardous or difficult working environments, while potentially improving productivity and reducing costs. Regardless of whether the machine is directly manipulated by an operator, or granted full autonomy to execute its specific mission, at some level, human observation, supervision, and judgment remain critical elements of the entire teleoperation activity [7]. The direct perceptual link to the

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 319–331, 2009.

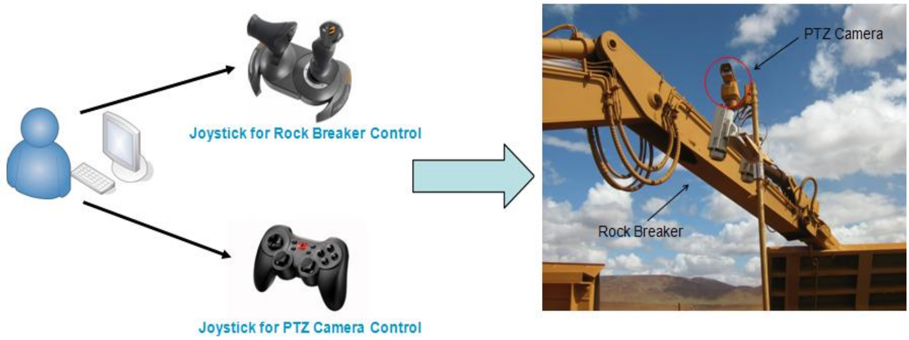


Fig. 1. Multi-task Situation for Rock Breaking in Remote Mining: for each device control, the operator has two joysticks, one hand does x, y and the other does z

remote environment often comes through a video feed supplied from one or more cameras as the foundation of situational awareness for the operator.

In practice, however, operators usually have to control multiple complex devices simultaneously, often more than they have hands, such as controlling a mechanical robot and a video camera at the same time. Figure 1 shows a typical teleoperation scenario for the rock breaking instance in mining. It is obvious that the operator is not capable of manipulating the two-joystick based 3D interfaces for the control of a rock breaker and to control a PTZ camera simultaneously as he can direct physical attention to only one task at a time. Under such circumstances, the operator needs to frequently switch hands between two different joysticks in order to accomplish the rock breaking task. This degrades the productivity of the entire process, increasing extra workload as well as the number of avoidable operational errors.

In this research, we focus on importing computer vision technology to undertake head tracking in interface design for teleoperation activities. The common remote control situation described above is modelled by using a physical game analogue: playing a table soccer game with two handles. This has the advantage of being more compelling for our student experimental subjects than a more abstract task. We use student experimental subjects as we have limited access to the operators. We then propose a novel design applying natural human head gestures for controlling a Pan-Tilt-Zoom camera as an effective approach to solve the camera control problem.

2 Head Tracking Technologies and Systems

Natural human head movements and gestures have served as a mode of interaction and communication throughout history. Responding to this common capability, much research has been done in trying to develop effective, robust and accurate head tracking technologies and systems to satisfy demand for building natural and interactive applications in the realm of human-computer interface design.

So far, various types of head tracking technologies have been developed. We can briefly classify these existing technologies according to the way head position is tracked into the following two main categories:

1. Sensor based head tracking.
2. Computer vision based head tracking.

The sensor based head tracking approach is fairly common. The typical configuration of this type of system comprises a set of sensors, which are required to be worn on a user's head (e.g. head-mounted tracker), and another hardware device for detecting the position of the sensor, receiving the transmitted head data. It can be either connected to a screen based display, or goggles (see Figure 2) for visual feedback and interaction.

There are a number of sensor based head trackers commercially available. TrackIR (Figure 3) is a typical head tracking device currently quite popular amongst gamers, especially in the simulation community [10]. This system consists of a small infrared camera placed on top of the monitor and a prepared baseball cap with three IR reflecting strips. The camera tracks the position of these reflective markers on the user's head, and reports head position with 6 degrees of freedom. Head orientation can then be used as input for many PC video games, for example, "fish tank VR", where a virtual world appears to be 3D as the view shifts depending on the angle of the user's current vision [14].

In recent years, much research effort has been expended on tracking and locating head pose, gestures and facial expressions from a video stream based on computer vision technologies. Compared with sensor based head tracking, this offers robust tracking quality, with more convenience and flexibility for the user as there is no need to wear any particular sensor device, and less cost for the hardware as usually only a normal webcam is needed.

In the computer vision area, head tracking generally starts with 3D face detection by defining corresponding facial features. For example, using facial geometry is a major strategy to estimate the face location as well as head motion [2]. In addition, color information is another powerful cue for locating the face [6] and other methods such as the use of depth information [8], classification of the brightness pattern inside an image window [12], etc. Figure 4 illustrates a commercialized real-time face tracking technology: FaceAPI, which provides a suite of image-processing modules created specifically for tracking and understanding faces and facial features with 6 degrees of freedom for head tracking [13].

Fig. 2. A Goggle Display for Head Tracking

Fig. 3. TrackIR System

Fig. 4. FaceAPI: Real-time Head Tracking with a Single Webcam

3 Head Tracking Applications in Human Computer Interaction

Head tracking is a key component in applications such as human computer interaction, person monitoring, driver monitoring, video conferencing, and object-based compression. Recently, one of the most popular ways of applying head tracking is to couple the virtual camera to a user's head position in order to achieve a more realistic and immersive experience of perspective in virtual reality or visual gaming. For instance, in [17], head tracking has been integrated into a first-person-shooter (FPS) game "Bullet Time" to control the user's view point. Another similar application with exaggerated head motions for game viewpoint control can also be found in [15].

In addition, there have also been attempts to develop head tracking based "hands-free pointing" interface for controlling the mouse cursor [16], by which a user can point his nose where he wishes to place the cursor on a monitor screen. "hMouse" is another head tracking driven camera mouse system [4], which provides alternative solutions for convenient device control with potential applications for people with disabilities and the elderly.

Other relevant applications are head tracking based user interfaces for navigation in virtual environments, remote control of devices [1] and head gestures (e.g. "Nodding" and "Shaking") based perceptual interface [3] [9].

4 Our Design of Head Gesture Based Remote Camera Control

The basic function of a PTZ camera is to Pan, Tilt and Zoom. With various functional combinations, the operator can obtain flexible control of its movement. With the integration of head tracking techniques with PTZ camera functions, we propose two sets of simple head gestures as interactive methods for remote control.

The first method (called "motion") operates according to natural human head motion. As shown in Figure 5, assuming initially that the user's head is directly facing the screen, when the user rotates the head to either left or right by a certain angle, the camera will pan in the corresponding direction. It will keep panning the view along that direction until the user moves their head back to the original position. Figure 6 shows similar interaction for the head tilting. When the user tilts their head up or

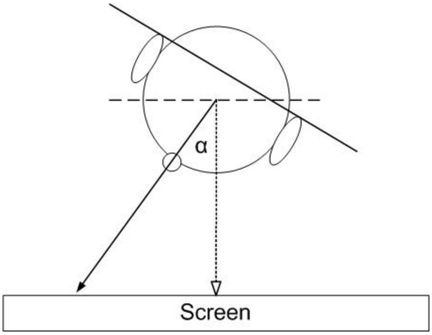


Fig. 5. Head Rotation

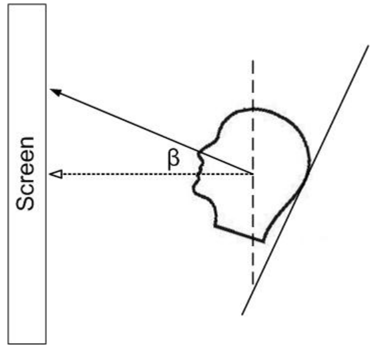


Fig. 6. Head Tilting

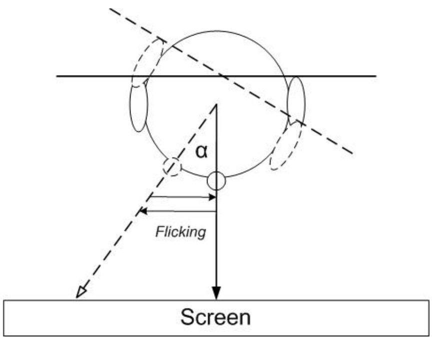


Fig. 7. Head Flicking for Panning

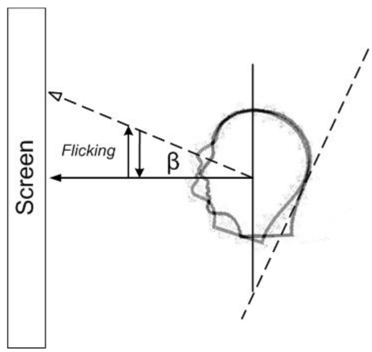


Fig. 8. Head Flicking for Tilting

down by a certain angle, the camera will correspondingly carry out the tilt function and not stop tilting until the head returns to the original position.

As for the zoom function, this specific control operates according to the distance between the user's head and the screen. For instance, if the user wants to have a more detailed view of the current video streams, he may naturally lean closer to the screen, effectively suggesting that the camera conduct a "zoom-in" function, and vice versa.

The other set of head gestures is based on human quick head movements, called "flicking". Head flicking based interactive control for camera functions is mostly like a switch. When a user quickly rotates his head to either the left or right direction then moves back to the original position, we consider this to be a head "flicking" along the corresponding orientation, which appropriately turns on the camera to start panning along this direction. When the user flicks to the opposite direction, it will switch the camera movement off and stop at the current position. Figure 7 and 8 are the relevant geometrical displays of head "flicking" for both pan right and tilt up actions respectively.

5 User Studies

A user evaluation experiment was conducted to assess how well these two head gesture based methods could perform the control of a remote camera in a model of a real-world teleoperation setting.

5. 1 Apparatus and Implementation

We integrated FaceAPI 3.03 [13] with a Logitech webcam [5] into our prototype system in VC++ that ran at 50Hz on a PC for real-time head tracking. The system used a Pelco ES30C PTZ camera [11] to perform the head gesture based control for our study. A keyboard based method was also implemented to simply control the PTZ camera by using the four arrow keys on the keyboard.

The display was a 19" monitor with a resolution of 1280 × 1024 pixels for showing the video stream from the camera to the user. A half size soccer table was placed under the monitor with several covers attached on one side to obscure the user's direct vision. Figure 9 and 10 show the experimental setup from front and back respectively.

5. 2 Participants

A total of 10 university students and staff (8 male, 2 female) participated in this evaluation, ranging from 21 to 48 years old with a mean of 29.6 years. All 10 were regular computer users with no previous experience in remote camera control. Four of them had some experience playing table soccer, and the rest had none. Most of the participants played computer games by using a keyboard occasionally (6 participants), one subject played quite often and the remaining 3 did not play games at all.

5. 3 Experimental Design and Procedure

The experiment was conducted using a 3 × 2 × 3 × 3 within-subject full factorial design. Factors were *control strategy* (Head Motion, Head Flicking or Keyboard), *table soccer experience* (Never or Occasionally), *computer game experience* (Never, Occasionally or Often), and *sequence of using three control methods* (S1: Motion → Flicking → Keyboard, S2: Keyboard → Motion → Flicking or S3: Flicking → Keyboard → Motion).

Please note in Figure 10, the experiment assistant was seated at the back. His role was to gently and consistently return the ball to the participant when it was out of reach of their soccer handles. Thus, participants were essentially playing a one-player game.

Particularly for the sequences, in order to avoid too many repeated trials which would affect the final results, a random selection of allocating participants into different experimental sequences was also carried out, which ended up with 5 subjects for the first sequence, 3 for the second and the remaining 2 for the last order.

Since the size of the entire play area was relatively small, we set the zooming level of the camera at a fixed value to only have a partial view of the field, leaving pan and tilt control to the participants. It effectively made the participants keep performing the control of the camera to find the ball throughout the whole experimental period, whenever the ball was out of the current area of vision.

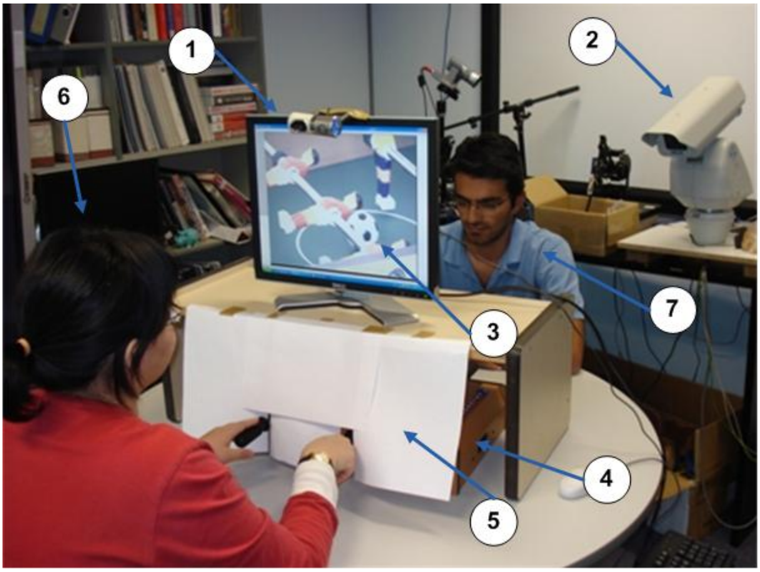


Fig. 9. Front View of the Test Setup, Webcam for Head Tracking (1), PTZ Camera (2), Video Stream from the PTZ Camera (3), Table Soccer (4), Covers for Obscuring Participant's Direct Vision (5), Experiment Participant (6), Experiment Assistant (7)

Fig. 10. Back View of the Test Setup

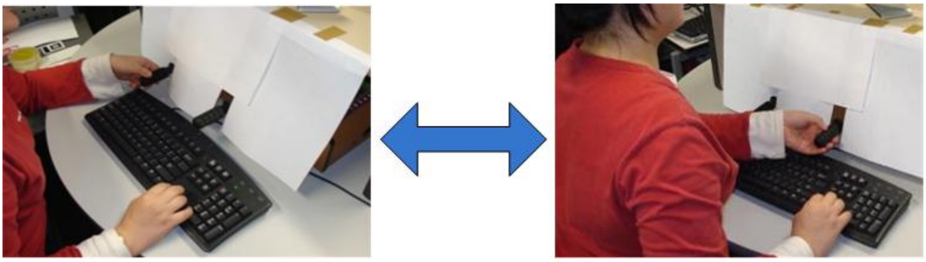


Fig. 11. One Hand Switching Between Handles, the Other Performing Camera Control by Keyboard

Participants were first given a short introduction (around 5 minutes) about the system, instructions on how to remotely control the PTZ camera by the two types of head gestures and the keyboard, and what kind of task they would be required to accomplish in the experiment. After that, subjects started the experiment with the randomly selected sequence of using those three control methods. No pre-training period was offered. For each method, participants had 5 minutes to play the table soccer game, and the number of kicks they made was recorded for the performance measure.

Once the table soccer game under all the three conditions had been finished, the participant was asked to complete a short questionnaire in which they compared their experiences with different control methods across several criteria for the subjective measures, including easiness, naturalness and time to get used to.

When conducting the keyboard based trial, there was no particular constraint for making the participant move both hands off from the two handles to the keyboard to adjust the view. As the control configuration of the keyboard was using only the four arrow keys, the participant could simply perform the camera control by using one hand pressing on the keyboard, leaving the other hand switching between two handles to kick (see Figure 11). In the future, we also intend to test if restricting the user in hand switching makes a difference.

6 Experimental Results

A repeated-measure ANOVA analysis was conducted on the performance measure to study the effects of all the factors, i.e. *control strategy, table soccer experience, computer game experience,* and *sequence of using three control methods*.

The overall average kicks were 24.53. The control method factor had a significant impact on the final performance, F (2, 22) = 5.6276, $p < 0.05$. Participants performed best by using the keyboard (M = 27.2, SD = 5.73), the mean kicks by using head motion control was fairly close to keyboard control (M = 25.9, SD = 8.29), but the head flicking method had much worse performance (M = 20.5, SD = 8.68). Figure 12 shows the mean kicks for each control method.

Whether the participants had table soccer play experience did not have any significant effect on how many kicks they made in the experiment, F (1, 22) = 0.0122, $p >$ 0.05. On the other hand, the factor of playing computer games using a keyboard turned out to have significant impact, F (2, 22) = 8.6814, $p < 0.01$. Participants who

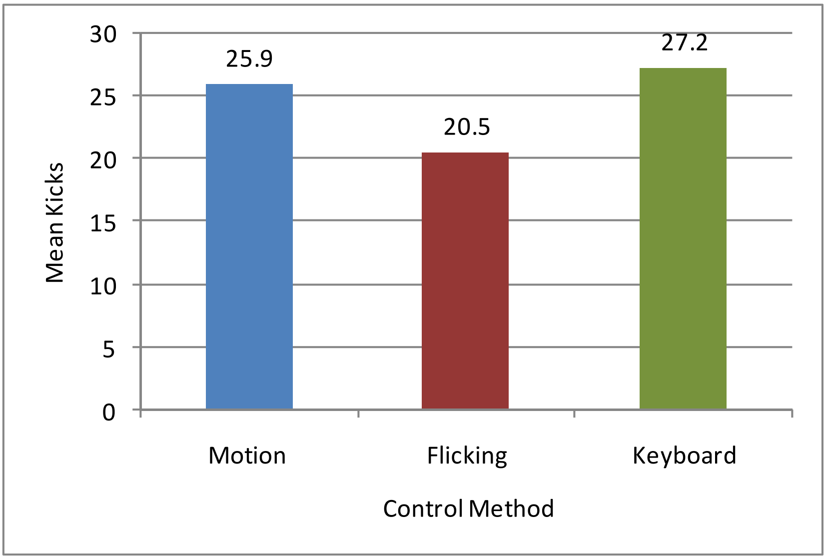


Fig. 12. Mean Kicks for Each Control Method

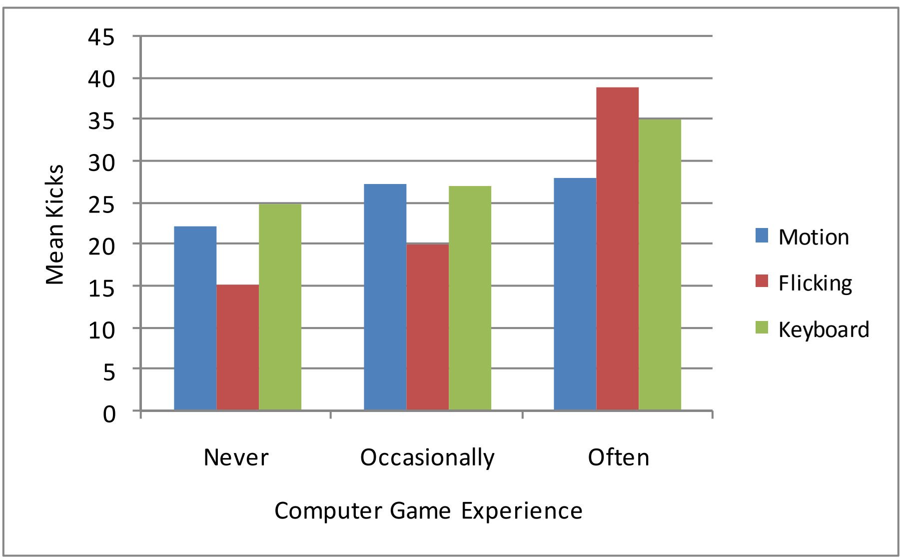


Fig. 13. Performance Comparison Based on Computer Game Experience

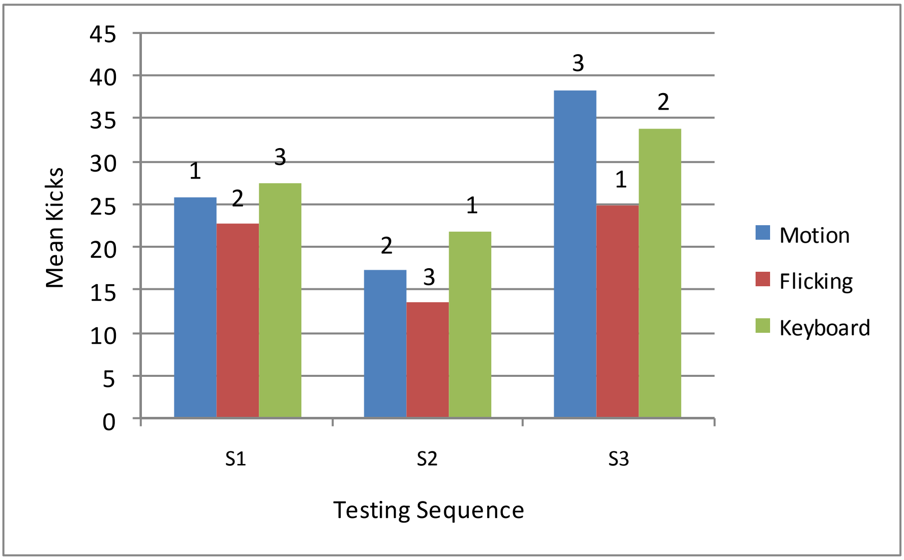


Fig. 14. Performance Comparison Based on Testing Sequence (S1: Motion → Flicking → Keyboard, S2: Keyboard → Motion → Flicking, S3: Flicking → Keyboard → Motion)

often played computer games using keyboards outperformed subjects with only occasional experience or no experience through all three different control conditions (see Figure 13).

In addition, the sequence of testing these three control strategies for each participant was highly significant for the performance, $F\ (2, 22) = 15.8212$, $p < 0.0001$. Participants following the last testing sequence on our list (Flicking → Keyboard → Motion) performed the best using all the control methods, compared with participants using the other two testing orders. Subjects starting with keyboard control had much worse performance in general (see Figure 14).

Table 1 illustrates the average scores participants rated for these three different control methods respectively according to their experience in the experiment. For the questions, we used a 4-point scale, rating from 1 (very difficult / very long) to 4 (very easy / very short).

Table 1. User Preference Results from the Subjective Measure

Average User Rated Point (out of 4)	Motion	Flicking	Keyboard
Q1: How easy/natural do you feel in the experiment?	*M = 2.6* *SD = 0.96*	*M = 1.9* *SD = 0.74*	*M = 3.2* *SD = 0.79*
Q2: How long did you feel to get used to the control method?	*M = 3.5* *SD = 0.53*	*M = 2.9* *SD = 0.99*	*M = 3.7* *SD = 0.48*

7 Discussion

Our objective results indicate that for this specific experimental setting, keyboards still performed the best by most of the subjects. We believe this is due to the fact that all the participants were quite familiar with using a keyboard, and initially there was no training time for them to get used to the two head tracking control methods. The reason for requiring the subjects to immediately start performing the experiment was to test how well users could pick up the head tracking based remote control. It is clear that our "head motion" based design provides quite comparable performance to the most conventional device (keyboard) even without any training.

As we mentioned in the previous section, another reason might be because users were actually switching only one hand between two handles, leaving the other hand for keyboard control in the experiment, which did not cost them much extra effort to trace the ball and make kicks.

Unsurprisingly, the "head flicking" strategy did not perform as well as the other head gesture based control. From our observations, when users were conducting trials they had to flick their heads quite frequently in order to find the ball. This is because the size of the viewing area was relatively small so that it required users to adjust the camera view quite often, which made the entire control be annoying and inefficient.

We have found similar results in the statistical analysis according to users' computer game experience using keyboards. Users with more gaming experience performed better not only in keyboard control but also in both head tracking controls. This is probably because those subjects already had more game based interaction experience, and in this particular game-like environment they may engage in the task more easily.

The results of testing sequence analysis indicate that this factor had a highly significant impact on the subjects' performance. A few interesting points have been discovered by the comparisons. Participants following the second testing order (Keyboard → Motion → Flicking) had a decreasing trend on the performance ($M_{\text{Keyboard}} = 22 > M_{\text{Motion}} = 17.33 > M_{\text{Flicking}} = 13.67$). In addition, the results of following the third sequence (Flicking → Keyboard → Motion) which used keyboard control between flicking and motion demonstrated another decreasing effect on these two head tracking controls ($M_{\text{Keyboard}} = 25$, $M_{\text{Flicking}} = 38.5 > M_{\text{Motion}} = 34$). On the other hand, the results of conducting motion or flicking control first in the sequences produced performance which was significantly improved over using keyboard control first (i.e. $M_{\text{S1-Motion}} = 26 > M_{\text{S2-Motion}} = 17.33$, and $M_{\text{S1-Flicking}} = 22.8 > M_{\text{S2-Flicking}} = 13.67$; while $M_{\text{S3-Motion}} = 38.5 > M_{\text{S2-Motion}} = 17.33$, and $M_{\text{S3-Flicking}} = 25 > M_{\text{S2-Flicking}} = 13.67$).

We suggest that as all the participants were good at using keyboards, they might be highly locked into this very familiar interface through the whole experimental period. This over-trained skill would affect the learning process for subjects to get used to operating the new interfaces introduced subsequently in the sequence. From our results, the use of keyboard control first actually depressed the performance of the two head gesture based methods.

The results of subjective measure are consistent with the performance measure. Keyboard control was ranked as the best, but users also suggested that the head motion control could be picked up very naturally without pre-training. Compared with these two methods, the head flicking control was the worst choice by users' consistent dislike.

8 Conclusion

In this paper, we considered the common problem of requiring an operator to control multiple devices simultaneously in current teleoperation, especially in remote mining. We presented our approaches of using two different sets of human head gestures to control a PTZ camera as potential solutions for this real-world situation.

The experiment we designed used a simple physical game analogue, modeling a multi-task environment for testing the users' performance through three different remote camera control strategies, including head motion control, head flicking control and keyboard control.

From the results, we demonstrate that the head motion based control is able to provide a comparable performance to using a keyboard even without the requirement of pre-training time, and the subjective measure of user's preference also indicates that the head motion is a comparable and effective method for this remote camera control case. Furthermore, we find that the sequence of conducting the three methods is the most significant factor. The use of keyboard control first depresses the success of using the other two head tracking methods.

If the results of our experiment are maintained or consistent in longer term training and use setting, it would suggest a seemingly paradoxical training regime of using the least familiar and worst control method for initial training to enhance subsequent performance. This warrants further investigation.

We believe our results map back to the mining teleoperation setting as follows. The results as to the two forms of head motion based control are likely to be directly applicable, so we expect flicking to be worst. The keyboard was familiar to our subjects, which most likely maps to joystick control in the mining setting, as the operators use joysticks a lot to control the rock breaker.

Our sequence results mapped to the mining setting would mean that operators presented with a teleoperation interface and initially presented with a joystick based control of the camera would have diminished performance with head motion based control. We could explain this by arguing that in the task encumbered mining setting, as long as the control of the camera is "good enough" minimal extra effort is expended on any later control strategy, as in real use the user knows they can resile to the "good enough" strategy. On the other hand, when initially presented with a novel interface, some effort is expended in learning to use it, which then has benefits for subsequent performance.

This explanation has two testable consequences when mapped back to our experimental setting. Firstly, that the same experiment with a less encumbered task would reduce or eliminate the effect of the keyboard in depressing head tracking based control. Secondly, that our physical model of the mining setting was sufficiently engaging to elicit such focus on the task, hence an equally encumbered but less engaging model would again reduce the effect of initial use of the keyboard.

Acknowledgements

The authors would like to express their appreciations to all the students and staff that participated in the experiment, and also thank Matt Adcock, Chris Gunn and Amir

Hadad for their great help on the implementation as well as the special assistance in the experiment.

Reference

1. Avizzano, C.A., Sorace, P., Checcacci, D., Bergamasco, M.: A Navigation Interface Based on Head Tracking by Accelerometers. In: 13th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2004), Kurashiki, Okayama, Japan, pp. 625–630 (2004)
2. Birchfield, S.: An Elliptical Head Tracker. In: 31st Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, vol. 2, pp. 1710–1714 (1997)
3. Davis, J.W., Vaks, S.: A Perceptual User Interface for Recognizing Head Gesture Acknowledgements. In: 2001 Workshop on Perceptive user interfaces (PUI), Orlando, Florida, USA, pp. 1–7 (2001)
4. Fu, Y., Huang, T.S.: hMouse: Head Tracking Driven Virtual Computer Mouse. In: 8th IEEE Workshop on Applications of Computer Vision (WACV 2007), p. 30 (2007)
5. Logitech, Inc. (2009), `http://www.logitech.com/`
6. Hsu, R.-L., Abdel-Mottaleb, M., Jain, A.K.: Face Detection in Color Images. IEEE Trans. Pattern Anal and Mach. Intel. 24(5), 696–706 (2002)
7. Hughes, S., Lewis, M.: Robotic Camera Control for Remote Exploration. In: ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2004), Vienna, Austria, pp. 511–517 (2004)
8. Malassiotis, S., Strintzis, M.G.: Real-time Head Tracking and 3D Pose Estimation from Range Data. In: 2003 International Conference on Image Processing (ICIP), Barcelona, Spain, vol. 2, pp. 895–862 (2003)
9. Morency, L.-P., Sidner, C., Lee, C., Darrel, T.: Contextual Recognition of Head Gestures. In: 7th International Conference on Multimodal Interfaces (ICMI 2005), Torento, Italy, pp. 18–24 (2005)
10. Natural Point, Inc.: TrackIR (2009), `http://www.naturalpoint.com`
11. Pelco, Inc. (2009), `http://www.pelco.com/`
12. Rowley, H.A., Baluja, S., Kanade, T.: Neural Network-based Face Detection. IEEE Trans. Pattern Anal and Mach. Intel. 20, 23–38 (1998)
13. Seeingmachines, Inc: FaceAPI (2009), `http://www.seeingmachines.com/faceAPI.html`
14. Smith, J.D., Nicholas Graham, T.C.: Use of Eye Movements for Video Game Control. In: 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE 2006), California, USA, p. 20 (2006)
15. Teather, R.J., Stuerzlinger, W.: Exaggerated Head Motions for Game Viewpoint Control. In: ACM International Academic Games Conference on the Future of Game Design and Technology (FuturePlay 2008), Toronto, Ontario, Canada, pp. 240–243 (2008)
16. Toyama, K.: Look, Ma – No Hands! Hands-Free Cursor Control with Real-Time 3D Face Tracking. In: Workshop on Perceptual User Interfaces (PUI 1998), San Fransisco, USA, pp. 49–54 (1998)
17. Wang, S., Xiong, X., Xu, Y., Wang, C., Zhang, W., Dai, X., Zhang, D.: Face Tracking as an Augmented Input in Video Games: Enhancing Presence, Role-playing and Control. In: ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2006), Montréal, Québec, Canada, pp. 1097–1106 (2006)

QualiTrack: Highspeed TUI Tracking for Tabletop Applications

Ramon Hofer, Thomas Nescher, and Andreas Kunz

ETH Zurich, inspire, Tannenstrasse 3,
8092 Zurich, Switzerland
{hofer,nescher}@inspire.ethz.ch, kunz@iwf.mavt.ethz.ch

Abstract. In this paper, we present a new technology to track multiple active Tangible User Interfaces (TUI) on a projection table. We use a commercial high speed infrared tracking camera with modified firmware. With a special tracking method, we reach update rates of up to 250 Hz with low latencies. At this tracking rate, we are able to track the position, state and the orientation of more than 10 active TUIs on the table. For this, we use specified bit codes which are transmitted by the devices. We developed dedicated hardware (SyncUnit) and software to keep the devices and the high speed camera synchronized. The system of camera, SyncUnit, and devices is fully hardware controlled and delivers event coded tracking data for further usage in interactive applications.

Keywords: Single Display Groupware, Tangible User Interfaces, CSCW, Tracking, high speed, low latency.

1 Motivation and Related Work

In applications such as brainstorming, design review and general idea finding tasks, interfaces are needed that satisfy the need for highly responsive and intuitive interactions. Sketches and fast drawing or writing gestures are used to exchange information between collaboration partners. Thus, it is essential that multiple strokes - being sketched very fast by different users on the same screen - are reliably tracked and visualized in order to provide a real and intuitive feeling to the users. In addition, hiding technology but still providing all advantages is a key issue in this special field of human computer interaction. In this paper, a technical approach is shown to support such tasks by providing a high speed multi device tracking system.

Many researchers have addressed tracking of multiple objects on a screen. One of the most popular methods for multi-device tracking on horizontal surfaces is optical tracking, where a camera is placed above or behind a projection screen representing the interactive surface. The TUIs are either identified by shape recognition as in the DigitalDesk setup [1], by marker detection as in the InfoTable [2], by occlusion detection as in [3], or by color detection [4]. Using marker or color tracking provides the possibility to unambiguously track several devices at once, since each device can contain a unique optical identification pattern (fiducials). When using infrared light (IR), the tracking is no longer visible to the user.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 332–335, 2009.

Simulated table top tennis [5] is one of the very responsive systems that provides update rates of 60 Hz and latencies of only 50 ms by using commercial ART cameras. In [6], an infrared tracking system was developed which is able to track position, orientation and state of up to 7 interactive devices on a table. The final update rate was 20 Hz. By using device specific bit codes, simultaneous identification of all active devices was possible, but the tracking rate was low.

Up to now, there is no system which provides a high update rate (>100Hz) with reliable identification of multiple devices and determination of their states as well.

2 System Components

The QualiTrack system is an optical system using a high speed IR camera and a back projection onto an opaque table. Each interactive device has an integrated IR receiver, at least one IR LED, and also a signal processor. The devices detect an IR synchronization signal with their receivers and answer with a unique device code (a binary code of n bits). This code is flashed back bit per bit (IR LED on/off) upon every synchronization signal. At every synchronization signal, a camera frame is recorded and processed. When as many frames as the bit code length are triggered and recorded, the recorded camera frames are combined and the code, the position, as well as the orientation of the devices are determined. A dedicated hardware device, the SyncUnit, was developed, which handles the synchronization of all components, i.e. IR flash, camera trigger and data transfer from the camera to the PC host system.

A commercial high speed motion tracking camera from Qualisys, the MCU1000 (up to 1000 Hz) is used. In order to achieve high resolution at a high update rate, the camera is used at 250 Hz, at which it still provides the full resolution of 658 x 496 (>250 Hz, the field of view is reduced). The camera contains an efficient hardware-based blob detection and directly delivers the coordinates of detected IR points.

A dedicated SyncUnit (controlled by an ARM7 board) ensures that the camera and the interaction devices run in sync all the time. This synchronization is independent of the PC host system and runs on its own. The SyncUnit generates the sync IR flash to trigger the devices and it also triggers the camera. After that, the preprocessed data from the camera is received. USB is used to transfer the data from the SyncUnit to the

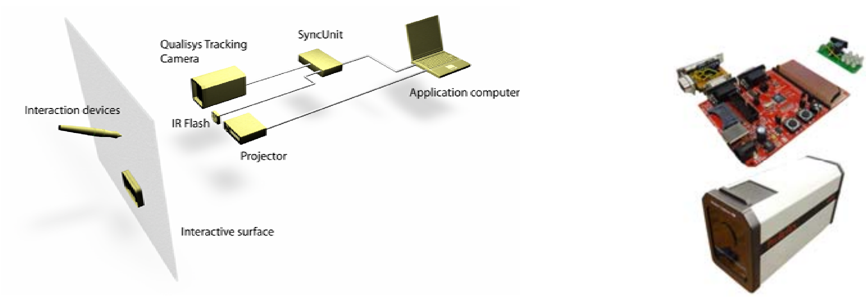


Fig. 1. Left: The QualiTrack setup. Right top: SyncUnit (ARM7 Board) and the IR Syncflash LED array. Right bottom: The Qualisys Camera ProReflex MCU1000 [7].

PC. A Java application processes the LED position data into TUI data. This includes the computation of TUI center points, their position, rotation, acceleration, rotational acceleration, and path prediction. The processed data is then sent to a client application via network using the TUIO [8] protocol.

The system distinguishes between two different kinds of devices: Devices with only one IR LED for drawing and pointing, and devices with 3 IR LEDs. The latter can also be tracked in orientation for other interaction possibilities. By pressing the devices onto the screen, a switch can be activated to change the state. An example for such a device is the Handle Tool. It serves as a navigation device. By pressing and rotating it the drawing area is zoomed. Shifting and pressing will drag the drawing.

3 System Characteristics

In order to unequivocally identify the TUIs, several post processing steps are required on the PC host system. Per frame, the SyncUnit transmits a set of LED position coordinates and the bit position number to the PC host. As every device has its unique n-bit code, the software on the PC host has to analyze n camera frames to identify the device. Hence, the real output frequency of the system of identified TUIs would drop to the camera frame rate divided by n. The smaller n is, the faster the system is, but the less TUIs can be used in parallel. To solve this problem, the devices' movements are predicted and each device's position is updated every time the camera sees it. This means that whenever a device enters the camera's field of view, the system has to wait for n frames to identify it (initialization phase). But for every subsequent frame, the device's position is updated whenever its bit code is 1 for a frame (LED switched on). Thus, the number of subsequent 0 in a device's bit code limits the update rate.

An 8-bit code was chosen for the system, in which only a single 0-bit is allowed between two 1-bits. Hence, the update rate only drops to half the camera frame rate. In order to further improve the tracking quality, the device's position is extrapolated using its previous velocity and acceleration. Thus, the device's position can be output at every camera frame using extrapolation. Although not all combinations of an 8-bit code can be used anymore, there is still a sufficient number of available IDs for all devices on the interactive surface of the table. Actually, this coding scheme would allow for having devices with low update rate as well (i.e. many subsequent 0).

There is another problem if a bitcode has many 1-bits: If many devices are used in parallel, they can also be active in the same camera frame. Thus, more potential collisions (extrapolated LEDs lighting up at the same location than another device's LED) are likely in particular if devices are close to each other. By employing position extrapolation, the area in which two devices could collide can be narrowed. Additionally, the bitcode should be chosen so that few devices have a 1-bit at the same bit position in order to minimize collisions. Apart from the 8-bit code, 2 more arbitrarily usable bits are appended to the code, which allow each device to signal its status to the PC host. This is used for instance to check whether a button is pressed. Thus, a status update is only possible after every 10th frame. The device codes were set with respect to the aspects mentioned before: update rate and collision prevention (see Fig. 2). For proofing the system's responsiveness, a pong game was realized.

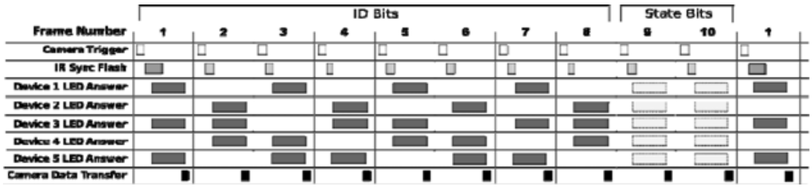


Fig. 2. Bit codes of the first 5 devices and timing of all components (1 Frame = 4 ms)

4 Conclusion and Future Work

A high speed tracking system for the detection of tangible user interfaces on a back projection tabletop surface is presented. It provides update rates of 250 Hz for fast interactions of up to 1m/s with reliable object identification and state recognition. The system forwards tracking data using the TUIO protocol to any TUIO compatible client. The amount of interactive devices is not restricted, but compromises in update rate and collision avoidance have to be made. We are currently working on creativity applications that need high spatial and temporal accuracy and low latencies, such as writing and sketching applications. It is planned to evaluate if such high update rates are necessary at all for the envisioned interactions. The goal is to set up a taxonomy that shows which low update rates are no longer tolerable for such applications as well as thresholds, above which high update rates are unnecessary.

References

1. Wellner, P.: Interacting with Paper on the DigitalDesk. Communications of the ACM 36(7) (1993)
2. Rekimoto, J., Saitoh, M.: Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. In: Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, David Lawrence Convention Center Pittsburgh, Pennsylvania, USA, ACM Press, New York (1999)
3. Smart Technologies, DViT (Digital Vision Touch) Technology, `http://www2.smarttech.com/st/de-DE/Products/SMART+Boards/Overlays/`
4. Nishimoto, K., Amano, K., Usuki, M.: pHotOluck: A Home-use Table-ware to Vitalize Communication at Mealtimes by Projecting Photos onto Dishes. In: TableTop, Adelaide, Australia. ACM Press, Adelaide (2006)
5. Rusdorf, S., Brunnett, G.: Real Time Tracking of High Speed Movements in the Context of a Table Tennis Application. In: VRST 2005, ACM Press, Monterey (2005)
6. Ganser, C., Steinemann, A., Kunz, A.: InfrActables: Multi-User Tracking System for Interactive Surfaces. In: IEEE VR 2006, Alexandria/Virginia, USA (2006)
7. Qualisys Motion Tracking Systems, `http://www.qualisys.com/`
8. Kaltenbrunner, M.: TUIO - A Protocol for Table Based Tangible User Interfaces. In: GW 2005, Vannes, France (2005)

Augmenting Surface Interaction through Context-Sensitive Mobile Devices

Alex Olwal

School of Computer Science and Communication
KTH (Royal Institute of Technology), Sweden
alx@csc.kth.se

Abstract. We discuss the benefits of using a mobile device to expand and improve the interactions on a large touch-sensitive surface. The mobile device's denser arrangement of pixels and touch-sensor elements, and its rich set of mechanical on-board input controls, can be leveraged for increased expressiveness, visual feedback and more precise direct-manipulation. We also show how these devices can support unique input from multiple simultaneous users in collaborative scenarios. Handheld mobile devices and large interactive surfaces can be mutually beneficial in numerous ways, while their complementary nature allows them to preserve the behavior of the original user interface.

Keywords: Touch, mobile, interaction techniques, interactive surface.

1 Introduction

Touch-sensitive surfaces enable dynamic user interfaces that avoid the physical constraints of traditional input devices, and are often viewed as intuitive and easy to use. The user experience in such systems is, however, often affected by calibration issues, parallax, limited resolution, and occlusion by the user's finger. The rigid surfaces also lack the tactile feedback provided through texture, actuation, physical properties and mechanics of dedicated input controls. The interaction tends to instead rely on single- or multi-touch gestures, on-screen widgets or custom interface elements.

Several projects have investigated the use of handheld or moveable displays to address these limitations, to extend and complement the capabilities of interactive surfaces. The M-Pad system [7] augments interaction on electronic whiteboards with handheld touch-screen devices, while the Pebbles project [5] focuses on office and home appliances, and collaborative scenarios for shared displays. Spatially aware handhelds can act as tangible lenses that provide detailed information about underlying media [3, 4], increase resolution on large digital surfaces [2, 8], and improve the precision of the user's interaction [6].

This work explores how these concepts can be combined to complement a large touch-sensitive surface with a commercially available mobile touch-screen device, to leverage its denser arrangement of pixels and touch-sensor elements for precise direct-manipulation, and its rich set of mechanical on-board input controls for increased expressiveness and feedback, as illustrated in Figure 1.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 336–339, 2009.

2 Focused Mobile Interaction in the Context of a Large Display

The content in a mobile display may be linked to the user's manipulations on an interactive surface, to enable synchronized presentations of relevant data. Such setups can leverage the respective advantages of the two devices, in which the handheld's better brightness, contrast, viewing angles and pixel density enable a high-fidelity focus, in the context of the larger surface [1]. A surface with device-tracking capabilities, such as our extended LightSense framework [6], can thus allow the mobile display to enhance the visual quality of the regions it is moved over. It can also be used as a magic lens to display alternative representations of the underlying data, such as a satellite view that is overlaid to complement a terrain map.

An alternative mode allows the user to control the view shown on the mobile display, held by the non-dominant hand, through direct pointing to the region of interest on the large surface, by the dominant hand. This mode does not depend on device tracking and enables more flexibility in posture for ergonomic viewing and interaction distances. See Figure 1.

Fig. 1. a–b) Multiple mobile devices are wirelessly connected to the surface and provide individual visualizations (road maps and satellite images) of the region indicated by the user's touch. c) The higher input resolution of the mobile device's touch screen allows the user to create precise annotations, while also being able to work in a comfortable pose.

3 Tactile Controls and Precise Touch-Screen Input

Device movement or direct touch with the finger can also be used to indicate the *point of interest for interaction*, allowing the mobile device to extend the input capabilities of the surface with its tactile controls and high-resolution touch screen.

The zoom action is an example of a popular operation on interactive surfaces, and is often implemented as a "two-finger stretch" or "pinch" gesture. It does, however, rely on the surface's quality and ability to sense multiple touch points, which, for example, excludes all single-touch displays. We have, however, previously discussed how such dependencies and limitations can be avoided through complementary techniques based on interaction through a mobile touch screen or by using the controls of a tracked mobile device [6]. This work introduces an additional technique based on a jog wheel (a rotary encoder that is a side-mounted equivalent of the mouse wheel, which is conveniently located under the user's thumb on many mobile devices), to support rapid changes in zoom level about the indicated point of interest on the large display. The wheel can be rolled downwards/upwards for zooming in/out and enables fine adjustments and large changes in the same mode, due to its mechanical properties. The wheel on the Sony Ericsson P1i device used in this work moves in 24-degree

Fig. 2. a–b) The lack of tactile buttons on the surface can be addressed by using the physical keyboard of the mobile device. Here, the user enters "Stadium" on a point of interest. c) Simultaneous input from multiple uniquely identified devices allows collaborative scenarios. A blue arrow has been drawn on the left device, while a red arrow and an outline of the stadium has been added using the right device.

steps, where passive tactile feedback is provided through small "snaps" that are felt during slow rolling actions. The continuous wheel also allows very fast rolling for large zooming magnitudes, but there is, in contrast to many other approaches, little risk of losing context, as the finger is always pointing to the target location. The zooming action is also easily reversible by rolling the wheel in the opposite direction. Devices that lack jog wheels can use the up/down keys, or one of the previously described touch-screen techniques [6] for zooming.

Tactile controls are perhaps even more important for text input. A physical keyboard's form factor and need for regular maintenance often make it impractical on interactive surfaces, which instead tend to favor on-screen keyboards. The lack of tactile feedback, however, demands high visual attention, and affects typing speed. Figure 2a–b shows how a mobile device's keyboard can be used as an alternative means of text input in such systems. It shares the mechanical and tactile benefits of ordinary computer keyboards, while emphasizing portability and compactness.

Similarly to previous work [4, 5, 6, 7, 8] of interaction through a mobile device's touch-screen, we provide precise annotation that is propagated to the large display, taking advantage of the higher pixel density and touch-sensor resolution of the mobile device. (See Figures 1c and 2c.)

3.1 Input from Multiple Uniquely Identified Users

One of the main benefits of multi-touch surfaces is the ability to support simultaneous interaction for multiple users. It is, however, not straightforward to identify an individual's touch input without the use of custom hardware and tethering. A user's mobile device can, on the other hand, provide implicit identification to allow unique input from multiple wirelessly connected users [5]. It can also surpass physical, computational or sensing constraints that affect how many touch-points single- or multi-touch surfaces can detect and track.

Our implementation supports collaborative annotations and text input in individual colors from multiple mobile devices, as shown in Figure 2. A user can touch a location on the large surface to indicate the starting point for text input, and can continue typing, even if another user moves the input focus to another part of the screen. Freehand annotations from different touch-screen devices are mapped to the current rectangular focus region on the surface and visualized with unique colors.

3.2 Implementation

The system uses our extended LightSense framework [6], where Java ME mobile clients communicate over Bluetooth with a Java server on the PC. The main application interfaces the touch-overlay, communicates with the Java Server over UDP and renders the graphics on the large surface using C++/OpenGL. Sony Ericsson provided P1i devices (touch screen and QWERTY-keyboard), while Nokia provided a 5800 XpressMusic (touch screen) and E71 (QWERTY-keyboard) device.

4 Conclusions and Future Work

This paper discusses the benefits of complementing an interactive surface with mobile devices, where bidirectional communication synchronizes graphics and user input. Such symbiotic scenarios, where the input and output mechanisms are shared over multiple connected devices, make it possible to leverage the respective advantages of mobile devices and large interactive surfaces. We provide examples of how the handheld device's higher pixel density, input resolution and rich set of onboard tactile controls, can expand the expressiveness of the large digital surfaces that provide the context for interaction. The possible input can be expanded further, by leveraging other built-in sensors in mobile devices, which could make typically exotic technology readily available on digital surfaces. Preliminary versions of the described techniques were used by trained, non-technical staff to demonstrate zooming and annotation on interactive surfaces to a large number of visitors during a five-day trade show. The staff's enthusiasm and interest in the techniques, combined with how fast they learned to use them, was encouraging. We plan to further study our techniques in formal user evaluations, with regard to usability, preference and performance.

References

1. Baudisch, P., Good, N., Bellotti, V., Schraedley, P.: Keeping things in context: A comparative evaluation of focus plus context screens, overviews, and zooming. In: Proc. CHI 2002, pp. 259–266 (2002)
2. Benko, H., Ishak, E.W., Feiner, S.: Collaborative mixed reality visualization of an archaeological excavation. In: Proc. ISMAR 2004, pp. 132–140 (2004)
3. Fitzmaurice, G.W.: Situated information spaces and spatially aware palmtop computers. Commun. ACM 36(7), 39–49 (1993)
4. Mackay, W.E., Pothier, G., Letondal, C., Bøegh, K., Sørensen, H.E.: The missing link: augmenting biology laboratory notebooks. In: Proc. UIST 2002, pp. 41–50 (2002)
5. Myers, B.A., Nichols, J., Wobbrock, J.O., Miller, R.C.: Taking Handheld Devices to the Next Level. IEEE Computer 37(12), 36–43 (2004)
6. Olwal, A., Feiner, S.: Spatially aware handhelds for high-precision tangible interaction with large displays. In: Proc. TEI 2009, pp. 181–188 (2009)
7. Rekimoto, J.: A multiple device approach for supporting whiteboard-based interactions. In: Proc. CHI 1998, pp. 344–351 (1998)
8. Sanneblad, J., Holmquist, L.E.: Ubiquitous graphics: combining hand-held and wall-size displays to interact with large images. In: Proc. AVI 2006, pp. 373–377 (2006)

Designing Novel Image Search Interfaces by Understanding Unique Characteristics and Usage

Paul André[1], Edward Cutrell[2], Desney S. Tan[2], and Greg Smith[2]

[1] ECS, University of Southampton, UK
pa2@ecs.soton.ac.uk
[2] Microsoft Research, Redmond, WA, USA
{cutrell,desney,gregsmi}@microsoft.com

Abstract. In most major search engines, the interface for image search is the same as traditional Web search: a keyword query followed by a paginated, ranked list of results. Although many image search innovations have appeared in both the literature and on the Web, few have seen widespread use in practice. In this work, we explore the differences between image and general Web search to better support users' needs. First, we describe some unique characteristics of image search derived through informal interviews with researchers, designers, and managers responsible for building and deploying a major Web search engine. Then, we present results from a large scale analysis of image and Web search logs showing the differences in user behaviour. Grounded in these observations, we present design recommendations for an image search engine supportive of the unique experience of image search. We iterate on a number of designs, and describe a functional prototype that we built.

Keywords: image search, log analysis, design.

1 Introduction

Most major search engines treat general Web search and image search very similarly. A searcher types in one or more keywords into a text box and results appear in a paginated view ordered by ranking algorithms that are largely based on textual features. While the experience for image and Web search are similar, we believe that there are important differences that are not well-recognized or exploited by search engine providers. The goals of searchers using image search are likely quite different from those using general Web search. In addition, images occupy a high-dimensional semantic and visual space that could be used to improve retrieval and the presentation of/interaction with search results.

Researchers have introduced many innovations in retrieval for image search (e.g., textual and visual features [2], distance measures [20], clustering [21], query-by-image, etc.), although few of these techniques are used in practice in major image search engines. In contrast to the excellent work in retrieval techniques, there has been little research in understanding image search behaviour, particularly in comparison to Web search. We believe we can contribute more immediate value by recasting existing technologies deployed for image search within a new interface and interaction model designed for the unique aspects of the image search experience.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 340–353, 2009.

In this paper we explore ways to design more compelling image search experiences by sharing two main contributions. First, we conducted informal interviews with researchers, designers, and managers working to develop and deploy a major commercial search engine. From these interviews, we derived four main characteristics that make image search unique from its Web search counterpart. Following this, we also report on a large-scale analysis of query logs to characterise some of the differences between image and Web search in the user activity of a major commercial search engine. Second, based on these observations, we generate design recommendations for interfaces tuned for the needs of image search. To exemplify these recommendations, we describe a functional prototype interface for image search that we designed to support the unique features of image search.

In the following sections, we review the literature and some existing image search engines. We then describe the methodology and results of our interviews and query log analysis. We detail how these results form the basis for the design of a new image search interface. We conclude with recommendations for evaluation and future work.

2 Related Work

2.1 Image Search Engines

The major Web search engines (e.g. Google, Yahoo, Baidu, Microsoft Live, Ask) all provide a similar experience for image search: keyword-based query resulting in a grid of image thumbnails. Various query refinements are available, including image size, aspect ratio, colour and different kinds of content (e.g., illustrations, or images containing a face).

Besides the major search engines, there are a wide range of specialised, original or experimental image search engines on the Web (we counted at least twenty-one and imagine there are many others). Though space precludes a full listing, three of the most interesting examples are: captivating colour search [13], in which images are returned based purely on one or two chosen colours; Viewzi [19], which presents results in a number of 'views' – flipping the pages of a cookbook when searching for recipes, a combined photo and tag cloud, an Amazon book view; and Getty Catalyst search [8], which leverages rich ontological and tag metadata to provide easy refinement of searches.

While there are a number of innovative image search engines, they tend to either be purely aesthetic-driven or cater to niche communities, neither of which increases functionality for the average image searcher. Our goal in this work is to explore a design space that allows appealing aesthetic treatment while still supporting greater efficiency in the task-centric behaviour seen in the larger search engines.

2.2 Improving Retrieval and Presentation of Images

The majority of academic work on image search has focused on back-end technologies. Such improvements include similarity space projection [16] in which images and queries are projected into a third space such that relevant images are kept close to the corresponding query; a framework to reduce the semantic gap in content-based image retrieval [20]; and automatic image annotation through a domain-specific ontology

[7]. Though all are promising, these techniques can face serious challenges to Web scale implementation, and as far as we can tell, none have seen commercial use.

A number of projects utilise new user interfaces to exploit work on the back-end. In many cases, these comprise different visual representations of various clustering techniques [2,3,5,21]. Other innovations in feature extraction have led to different interfaces. Liu et al. [15] display a one page overview of results, with thumbnails cropped to regions of interest based on low-level features, and similar images displayed near each other. In Visual Islands [25], raw features are mapped to a 2D space, and the images are snapped to a grid maintaining spatial layout of the 2D space. EnjoyPhoto [26] presents a subset of 'high-quality' images scraped from forums with rich metadata, and the user interface utilises fish-eye and force transfer implementations for in-place browsing.

CueFlik [6] is an interactive machine-learning interface to create rules for re-ranking images based on image characteristics extracted from representative examples the searcher provides. IntentSearch [4] categorises a query image into one of several predefined intention categories, with a specific similarity measure used inside each category to combine image features for re-ranking. At the end of 2008, Microsoft Live Image Search released a feature allowing a user to search for 'similar images to this'. However, the user interface remained essentially unchanged.

To our knowledge, few papers have focused purely on supporting the image search user experience through novel user interfaces. One exception to this is work by Porta [18], which describes several techniques for presenting all images within a collection in a short time to support users who do not know precisely what to search for.

In summary, although there are many innovative image retrieval technologies in the literature, the major search engines have been slow to adopt them. In order to provide a Web-scalable prototype that performs well for most scenarios of image search (i.e. the long tail), we chose not to focus on complex retrieval algorithms, but instead on presenting current search technologies within a new user interface and interaction setting that is specifically designed for the unique qualities of image search. However, to better understand those unique qualities, it is necessary to understand how users actually use current technologies to search for images.

2.3 Image Search Usage

Previous work on understanding image search *strategies* has largely focused on classifying the kinds of queries posed by professionals. Analysis of image requests of journalists [17] found that over half were for a named person or object, and that browsing was an essential strategy, though it was not generally well supported. Similarly, a study of a newspaper image archive [22] indicated that around 40% of requests were for a named person, with a significant number of requests for general objects or themes. In a study of queries to a commercial database by professional users [14], proper nouns accounted for only 9% of queries, whereas nouns accounted for around 50% of searches, and again browsing and exploration were seen to be under-supported.

The most relevant work in understanding image search *behaviour* was undertaken in 2001 by Goodrum and Spink [10] using the Excite search engine. They found that image queries contained on average 3.74 words (though at least one of these was the image request term, e.g. 'picture' or 'image'). They also reported a high percentage of unique terms, and that modification of queries was common (accounting for 60% of queries). An analysis of image search patterns [9] found that the most commonly occurring patterns were rapid browsing of thumbnails and individual images. Hung [12] analysed search activities of journalism students and found that most activities in all tasks were related to browsing and enlarging images. Finally, a recent query log analysis [26] showed that more than 20% of image search queries are related to location, nature, and daily life categories.

The above work underlines the importance of flexible query modification and rapid browsing in the practice of image search. Like general Web search, image search appears to be used by a wide variety of people in support of many different tasks. However, much of the detailed query analysis was done on search logs that were collected at the dawn of the Web. Current commercial Web search engines now have dedicated image search verticals, and the number and variety of people using these services is dramatically larger than it was in 2000. As a result, we believe it would be helpful to explore how Web search and image search differ with current interfaces.

3 Understanding Characteristics and Usage of Image Search

3.1 Informal Interviews to Characterise Unique Goals and Intent

To better understand the fundamental characteristics of image search, we spoke with 8 researchers, designers, and managers associated with Live Image Search. From these discussions, we derived 4 characteristics of image search that may be relevant for interface design.

1) The search results themselves may fulfil a user's needs. Image search, like general Web search, can be both exploratory (information about art at the Louvre) or goal-directed (looking for a specific fact or image). In some cases, the 'answer' to the query can be found in the results page for either image or Web search and a user need never click on any result. One might find the answer to the query, "When was the Louvre built?" in the text of result snippets in Web search. Similarly, the query "What does the pyramid at the Louvre look like?" could be found just by examining a page of thumbnails returned by an image search.
2) Image search is often more exploratory than Web search. Searchers may be looking for an image with a particular visual 'style', with a predefined 'type' in mind, or with certain characteristics, and they may not be able to express those requirements until the desired image is found. This type of exploratory search within the rich visual feature space of images is not well supported today.
3) Image search is often used purely for entertainment to play or explore in visual space with no end goal. We see examples of this in Flickr's 'Interesting Photos' page, or the playful image search sites that allow search by something abstract such as colour [13]. The inherently visual nature of image search lends itself particularly well to the artful presentation of results.
4) Image search encourages tangents. Related to the previous point, the visual nature of image search makes it very easy to become sidetracked when something else of interest catches their eye, even if the initial query is task-focused. And if a searcher is browsing for entertainment, this kind of tangent is actually *desired*.

3.2 Search Log Analysis to Understand User Behaviour

Building on findings in the interviews, we conducted large scale analysis of search logs from the same search engine. Results presented in this paper are derived from a one day sample (October 10th 2008), allowing a broad view of over 55 million queries. The logs contain anonymised information about the query, the number of result pages seen, the number and location of clicks, and time spent looking at results. To remove variability caused by geographic and linguistic variation in search behaviour, we filtered to only queries generated in the English speaking United States ISO locale.

We examined search activity at two different levels of granularity: the query level (i.e., an individual query with a number of result page views), and the session level (i.e., multiple queries within a single session). We demarcated sessions by 30 minutes of no activity (as used in [24]). For both image and Web search, we explored similar questions at each level of granularity.

Table 1 shows the query level statistics comparing image and Web search. The statistics for a single query incorporate all interactions within that query (e.g., clicking a result, clicking to see the next page), and are reset when a new query is issued. First, we see that the average number of result clicks per query is higher for image search than Web search. Second, the average depth (number of consecutive results pages viewed for a single query) for image search is more than twice that of Web search[1]. If we filter to look only at the queries that had result clicks, thereby excluding all queries in which the user just looked at results on the results page without clicking, then we see that 30% of all queries for image search contained a result click, significantly lower than the 43% for Web search. However, when a query did contain a click, image queries had more clicks than Web queries.

Overall, the dwell time (i.e., the amount of time a user spends viewing a page of search results) was significantly higher for image than Web queries.

Table 1. Comparison of search log statistics at the query level, for image and Web searches

Query level	**Image**	**Web**
Avg. depth	3.47	1.25
Median depth	2	1
Avg. clicks	0.73	0.63
Median clicks	0	0
% with > 0 clicks	30%	43%
Avg. clicks	2.5	1.5
Median clicks	1	1
Avg. dwell time (s)	85	52

[1] It should be noted that in Live Image Search, pagination is implemented through a continuous scrolling mechanism, so total depth may be somewhat inflated over search engines with a strict click-based pagination. However, the general trend of increased average depth confirms intuitions in discussions and the difference is quite dramatic.

Table 2. Comparison of search log statistics at the session level, for image and Web searches

Session level	**Image**	**Web**
Avg. queries/session	2.45	1.59
Median queries/session	1	1
Avg. clicks per session	1.87	0.96
Median clicks/session	0	0
% with > 0 clicks	35%	42%
Avg. clicks	5.4	2.3
Median clicks	3	1
Avg. dwell time (s)	312	109

Table 2 details the session-level statistics for the same measures shown in Table 1. A session consists of a set of queries issued by a user in near temporal proximity. The session ends when there is a period of inactivity of at least 30 minutes between queries for a given (anonymous) ID. At this level we are able to gain a broader view of user behaviour, focusing on the number of queries issued within a session, and which of those queries contained clicks on search results.

At the session level, we see that, on average, more queries are issued and more results are clicked in an image search than a Web search session. Second, even though a higher proportion of image search sessions resulted in no search result clicks, the average number of clicks for image search was significantly higher than for Web search among sessions with a click. Finally, echoing findings at the query level, the average dwell time for a session was significantly longer for image than Web search.

There are several reasons that multiple queries may appear in a session, including different kinds of query refinements and consecutive unrelated searches. In order to understand this behaviour, we manually examined query text from around 1000 sessions chosen randomly from both image and Web search logs that contained sessions of two or more queries. We term these 'query trails'. We categorised them into 3 groups, those containing *unrelated*, *tangentially* related, or obviously *related* queries. Though there is some overlap between these categories and the classification is subjective, this scheme provides a reasonable estimate of the progression of search terms within a session. Table 3 displays the breakdown of classification, with examples.

Table 3. Classification and examples of search log queries

Query trail classification	**Image**	**Web**	**Example**
Unrelated	20%	35%	reebok → family guy; fitness → deal or no deal
Tangential	10%	5%	horned lizard → horned lizard desert → desert color
Related	70%	60%	spiders → hawaiian spiders → hawaiian sugarcane spiders

The majority of query sessions for both image and Web search comprise a trail of *related* queries, although there are somewhat more for image than for Web search. In contrast, Web search has a larger proportion of trails with *unrelated* queries. This suggests that when searchers turn to image search, they tend to explore a single line of questioning or topicality, iterating on related queries. This is less true for Web search, where searchers are more likely to engage in unrelated queries within a session.

3.3 Summary and Analysis of Results

The findings from our interviews and query log analysis suggest several interesting implications for the design of interfaces for Web-based image search. In both query- and session-level search statistics, image searchers view more pages of search results (average depth), they spend more time looking at those pages of search results, and they click on more results than Web searchers. This implies that either query relevance is not as important for image as for Web search or that query relevance is simply much worse for image search than Web search. We hypothesize that the increased click average is mainly due to two factors. The first is that there is often no definitive answer to an image query – the sought after image could be one of many, and it is a subjective result. The second is that we believe image thumbnails offer visually appealing objects that searchers may click on solely for the aesthetic value.

Compared to Web searches, image searches contain a greater number of queries that have no result clicks. However, when image searches do contain clicks, they contain significantly more. We hypothesize that the reason there are more queries with no clicks is that the answer (or lack of) is contained in the results page itself (i.e. low relevance is obvious and no result clicks are necessary, or a question such as 'what does a passion fruit look like' is answered immediately), or the user is merely browsing for enjoyment. The greater number of tangential query progressions implies that there is a diverse range of results returned that inspire a new search and/or image searchers are more conducive to going 'off-task' than Web searchers. The large number of related query trails shows that refinement and iteration are important in both image and Web search. However, we believe that a facility for such refinement is particularly important in image search where a goal (a mental picture of the desired image) may be far more difficult to specify with traditional keyword queries.

In the following section we consider how the findings from our log analysis, as well as from interviews and related work, can be used for design recommendations.

4 Design Suggestions

Based on our findings, we derive design suggestions that range from pure aesthetics to specific functionality in order to better support new image search experiences.

4.1 Supporting Exploration

Exploratory interaction paradigm. Our analysis suggests that image searches are more exploratory than Web search: more search pages viewed, 80% of query trails directly or tangentially related, and looking for a non-definitive answer within a multi-featured visual and semantic space. The existing page-based grid view of images, though clean and simple, does not visually or conceptually convey the user's path through this space. A literal representation of such a space would likely be overwhelming (e.g., lists of features mapped onto a multi-dimensional representation). However, any simple metaphors or concepts to change the interaction and interface in a way that supports the exploration of associated terms, or image-based queries, would be an advantage. Because image searchers tend to view many more results (more queries per session, more page views per query, and more clicks per session), support for rapid browsing would also be desirable. A highly interactive system that immediately displays more results is very important.

Entry point. Entry points for the major image search engines are virtually indistinguishable from Web search: a query box on a clean, blank page. Such a practice reinforces a task-centric approach to image search; a searcher is offered a blank slate that they must fill with a keyword expression. Exploratory search is better served by allowing searchers to frame a search in various ways. One way to do this is to prepopulate the opening page with sample images, just as some "experimental" image search sites do. Such images may be popular pictures, things taken from current news, or based on former searches (e.g., the new tab shown by the Google Chrome Web browser [11]). A searcher can easily launch a new query, but these opening images also give a starting point for non-directed browsing. We believe this would provide a more compelling visual experience, an opportunity to learn about current news stories or popular searched for topics, and increase the chance of serendipitous findings [1].

4.2 Fun and Aesthetics

In contrast to Web search which is dominated by text, image search is overtly visual. We believe that the plain grid view of current search engines fails to adequately exploit this fact. Because the search results themselves are visual artefacts, there are a variety of visual layout techniques that, while not necessarily enhancing functionality *per se*, can provide a beautiful aesthetic and a more compelling search experience.

As noted in section 3.1, a common use of image search is pure image *browsing*, whether for general entertainment or for specific information needs. As such, aesthetics are likely very important compared to other search engines that are used for more utilitarian tasks.

4.3 Query Refinement

The image search logs show a significant amount of query reformulation towards a specific goal (around 70% of all queries). Such reformulation supports our belief that an image searcher generally has a good idea of the type or style of the desired result prior to issuing their query. Though this may also be true of Web search, in image search, it is often much more difficult to express this textually. For example, an image searcher may be looking for the 'mood' or framing of an image, for the dominant colour or perhaps a particular pose or position of a subject.

Some requirements can be expressed using existing refinements, such as aspect ratio, or black and white. Others require more complex features of the image (e.g., edge histograms, colour histograms, face detection, etc.). Some existing image search sites allow query by image, or a 'more like this' query. Because of the mental model that searchers have of images, and the extensive reformulation as seen in the logs, such features should be considered essential refinement options.

Refinement based on semantics should also be considered. In this context, semantics of an image could be acquired from the labels it has in the image database, the query terms that led to clicks on that image, or user-generated tags. Since tags can be at different granularities or conceptual levels, they allow for another type of refinement: broadening, narrowing, or even 'side-stepping the current query. They may also be used alongside visual features in a 'more like this' query.

Query and Refinement History. The easy refinement proposed above addresses the concern that current image search engines do not allow users to fully *explore* the rich visual and semantic space that images inhabit. However, such simple refinement and exploration can quickly lead to a complex history that is hard to retrace. A searcher may well ask, "How did I get here and how do I find my way back to that good content I saw a few minutes ago?" Therefore we need to design a history that details the current path as well as a means to navigate that history. Within a session, this would represent explicit queries (i.e. text typed by a user) and all the refinements that may have been added (e.g. metadata filtering, more like this image). Between sessions this could act in a similar way to Web browser history and allow revisitation of previous queries.

4.4 View and Save Images

Since about twice as many image search results are clicked in a session than Web search results, it is worth considering how one might support easily viewing many images (as discussed above in exploratory search), and saving them for later viewing. Though we have spent time considering the fun and aesthetics involved in image search, there is still a core productivity-oriented component. For such tasks, where the end goal is important, it may be useful and efficient to allow saving of particular images from a result set, allowing the user to later revisit the saved set. Such a feature could be integrated with the within- and between-session histories.

5 Design Concepts and Implementation

Based on the suggestions derived from related work, discussions and log analysis, we briefly present some novel concepts in four design storyboards, before discussing a final design prototype we implemented.

Fig. 1 shows four early storyboards. Initially, we wanted a radically different interface that facilitated the rapid browsing of thumbnails. In Fig. 1-A we focused on a spherical metaphor. The area in the centre is a 'black hole' from which images could emerge and spread in a circle, eventually shrinking and falling off at an edge, or the reverse. Also in this figure we see how an image could be used to seed a new query, exploring similar visual and semantic content.

Focusing on the importance of query modification, our next storyboards (Fig. 1-B and C) explored how a tag cloud and context menu could enable exploration and refinement, and an opportunity to learn about the search in a wider context. One can also see some of our investigations into the effect of subtle changes such as background colour or rotation of images on user experience – the simple act of slight random rotation of images evoked positive response in informal testing.

The storyboard in Fig. 1-D further explores the notion that the query and subsequent modifications are central to the search experience. The idea is that with the right visual feedback for selection and refinement, optimising a query to find the appropriate image is a simple task.

5.1 Final Design

The design we chose to prototype incorporates many of the features distilled from related work, interviews, and from log behaviour. Of course, this design is not intended to be 'complete,' but an interpretation or instantiation of the design suggestions in section 4. We highlight some of the limitations and future work in section 6. Our design emphasizes the importance placed upon query modification seen in previous designs and highlights *exploration* of visual space, which unlike most of the previous

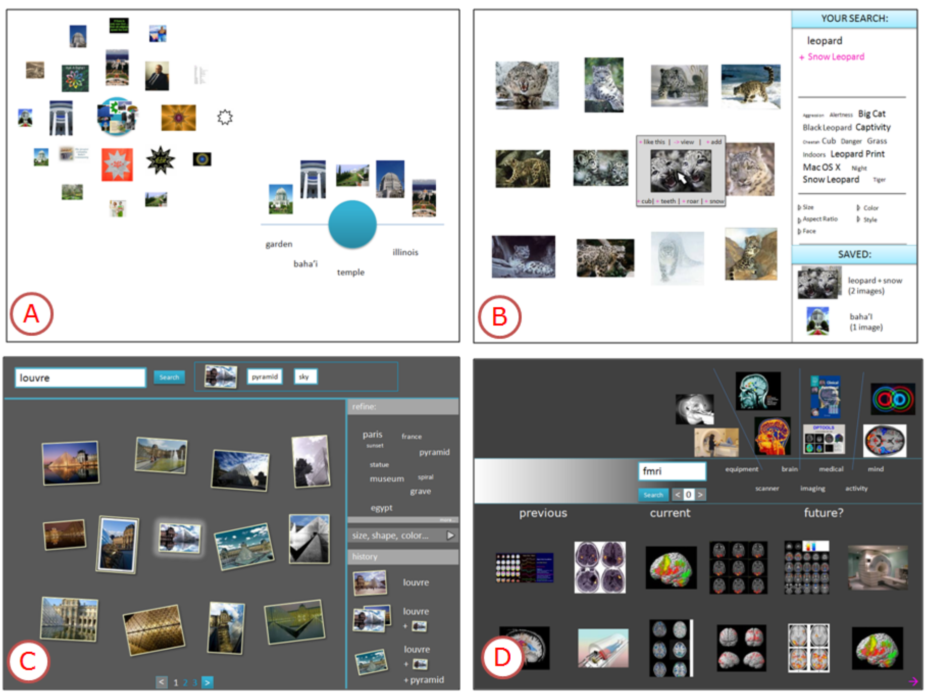


Fig. 1. Four image search engine designs, based on design recommendations from interviews and log analysis

designs is also alluded to in the presentation of results. Nicknamed 'Tendril', it uses a metaphor of tendrils (or ivy vines) for an organic display of images. The user seeds a search, and explores the resulting path (tendril) through the space. The overall space that a tendril could 'grow' into is infinite, and so we can imagine a true exploration of that space. Alterations or modifications to the query result in branching off the initial path. Figure 2 shows the interface and selected details from the working prototype.

Layout and Launch Page. The right hand side of the screen is used to display the search box and current query, history, and refinement possibilities. In the main area, a selection of images on the front page is used to highlight the exploratory (and not pure keyword) search experience, as well as to facilitate serendipitous browsing (Fig. 2-B).

Result Display. To emphasise the exploratory nature of the search engine, and to offer a distinctive and compelling experience, the result display takes the form of a 'tendril' extending across the screen, leading the searcher through the imagined vast result space (Fig. 2-A). Conceptually this is grounded in exploration of unseen semantic and visual space, something that in later iterations we hope to convey more obviously with the camera panning out to see the 'world view' before zooming in again on a particular part.

Context Menu and Query Modification. Hovering over an image presents a context menu with familiar actions such as 'view', and also enables query modification so prevalent in related work and in our log analysis. Query modification is a form of pivoting, and could be used to find an image 'more like this', or to use tags or other metadata to broaden/narrow/sidestep the current search. We have chosen to expose the ability to search for 'more like this' (since it was available through Live Image Search XML), and we have specifically designed the user interface to promote and take advantage of such a feature. The user can select to search for more like this from the context menu, or (soon) drag an image to the search box. Such a feature is especially useful in those cases where users may have an idea of what they are looking for but are unable to express it in words. Semantic exploration in the form of keywords (either taken from an image database, or from Flickr or a similar service) is also desirable. While examples such as Getty and Flickr have this metadata, current search indexes do not, or do not expose it in an efficient manner. In the future, we would like to include it in the manner described in Fig. 1-B/C.

Search Box. Along with the initial query, this area acts as a control area for query modifications (Fig 2-E), where refinements or 'similar images' added to the query are shown. This is also the region into which related images are dragged for refinement.

Branching. As modifications are made to the query, the previous path through the search space is no longer valid, and so a new path is 'branched' off. An example of this is seen in Fig. 2-D, where a new query has been issued based on an image.

History. With the ability to easily explore and refine, a complex path could soon be built. The history allows a user to see the path to their current search results, and to 'step back' if, for instance, the latest modification was not to their liking. A miniature version of the tendril on that page (each tendril is unique) provides a representation that could be used to aid navigation back to a previously seen page (Fig. 2-C).

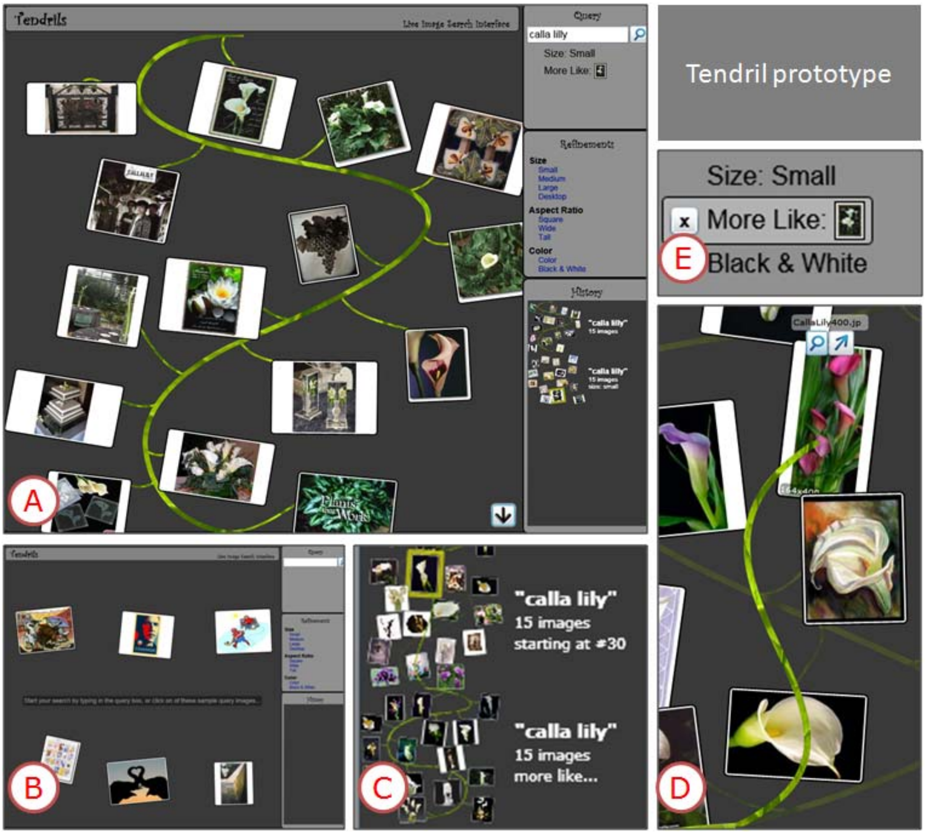


Fig. 2. The Tendril design concept. (A): a full page with query and results. (B): the launch page with a selection of random images. (C): enlarged view of history. (D): example of branching after querying for 'more like this'. (E): enlarged view of query area.

5.2 Implementation Details

The Tendril prototype is implemented as a Silverlight 2 application running in the browser and an associated Windows Communication Foundation Web service, both hosted on an ASP.NET Website. Queries and refinement clauses entered into the interface are translated into query options and submitted asynchronously back to the Web service on the Tendrils Website. To utilise existing functionality (notably the 'search for similar images' option) in Live Image Search, the service builds the appropriate query string in Live Image Search syntax and submits the query to search.live.com/images, asking for a response in XML format. The XML response is returned to the Silverlight application and parsed for the individual search results. The application then submits additional asynchronous requests to the Tendrils Web service for a thumbnail for each result, and inserts the thumbnails into graphical templates in the Silverlight interface as they are returned. It is worth noting that designing this implementation as a novel front-end that sits on a fully functional search engine allows us to tap into the vast and growing index of images, as well as the functionality offered by that engine.

6 Limitations and Conclusions

In this paper we have considered how understanding the unique characteristics and usage of image search can contribute to the design of a focused and compelling image search experience. We offer two main contributions. First, we offer a qualitative and quantitative analysis of differences in behaviour between image and Web search. Second, we use those insights to ground recommendations for the design of image search experiences. Additionally, we present five designs grounded in these recommendations including a full working prototype.

While Web log analysis provides a broad description of what people do, we do not have the context of their experience. In the future, we would like to observe and interview a variety of individual image searchers to better understand their needs and behaviour in specific contexts of use.

We continue to refine our prototype, in terms of both performance and aesthetics. In the future we would like to take one of our other design concepts to prototype, and do a formal evaluation of use against each other, and against commercial search engines, to determine exactly if and how our design suggestions improve the image search experience and goals therein. These concepts are only a few examples in a world of potential new designs in image search. We hope that the analysis and design recommendations in this paper will inspire others to consider improving holistic experience as well as retrieval in their own designs for image search on the Web.

Acknowledgments. We would like to thank Dan Liebling for data gathering and analysis assistance, Richard Qian, Steve Beck, Steven Drucker, Gonzalo Ramos, Gang Hua, and the many others we talked to for various helpful and informative discussions.

References

1. André, P., Teevan, J., Dumais, S.T.: From X-Rays to Silly Putty via Uranus: Serendipity and its Role in Web Search. In: CHI 2009 (2009)
2. Cai, D., He, X., Li, Z., Ma, W., Wen, J.: Hierarchical clustering of WWW image search results using visual, textual & link information. In: MM 2004, pp. 952–959 (2004)
3. Chen, Z., Wenyin, L., Hu, C., Li, M., Zhang, H.: iFind: a Web image search engine. In: SIGIR 2001, p. 450 (2001)
4. Cui, J., Wen, F., Tang, X.: Real time google and live image search re-ranking. In: MM 2008, pp. 729–732 (2008)
5. Ding, H., Liu, J., Lu, H.: Hierarchical clustering-based navigation of image search results. In: MM 2008, pp. 741–744 (2008)
6. Fogarty, J., Tan, D., Kapoor, A., Winder, S.: CueFlik: interactive concept learning in image search. In: CHI 2008, pp. 29–38 (2008)
7. Gao, Y., Luo, H., Fan, J.: Searching and browsing large scale image database using keywords and ontology. In: MM 2006, pp. 811–812 (2006)

8. Getty Catalyst Search, `http://www.gettyimages.com/Catalyst/Default.aspx`
9. Goodrum, A.A., Bejune, M.M., Siochi, A.C.: A State Transition Analysis of Image Search Patterns on the Web. In: Bakker, E.M., Lew, M., Huang, T.S., Sebe, N., Zhou, X.S. (eds.) CIVR 2003. LNCS, vol. 2728, pp. 281–290. Springer, Heidelberg (2003)
10. Goodrum, A., Spink, A.: Image searching on the Excite Web search engine. Inf. Process. Manage. 37(2), 295–311 (2001)
11. Google Chrome Web Browser, `http://www.google.com/chrome/`
12. Hung, T.-Y.: Search Moves and Tactics for Image Retrieval in the Field of Journalism: A Pilot Study. J. of Educational Media & Library Sciences 42(3), 329–346 (2005)
13. Idée Colour Search, `http://labs.ideeinc.com/multicolr/`
14. Jörgensen, C., Jörgensen, P.: Image querying by image professionals. Journal of the American Society of Information Science & Technology 56(12), 1346–1359 (2005)
15. Liu, H., Xie, X., Tang, X., Li, Z., Ma, W.: Effective browsing of Web image search results. In: MIR 2004, pp. 84–90 (2004)
16. Liu, Y., Qin, T., Liu, T., Zhang, L., Ma, W.: Similarity space projection for Web image search and annotation. In: MIR 2005, pp. 49–56 (2005)
17. Markkula, M., Sormunen, E.: End-User Searching Challenges Indexing Practices in the Digital Newspaper Photo archive. Information Retrieval 1(4), 259–285 (2000)
18. Porta, M.: Browsing large collections of images through unconventional visualization techniques. In: AVI 2006, pp. 440–444 (2006)
19. Viewzi Search Engine, `http://viewzi.com`
20. Wang, C., Zhang, L., Zhang, H.: Learning to reduce the semantic gap in Web image retrieval and annotation. In: SIGIR 2008, pp. 355–362 (2008)
21. Wang, S., Jing, F., He, J., Du, Q., Zhang, L.: IGroup: presenting Web image search results in semantic clusters. In: CHI 2007, pp. 587–596 (2007)
22. Westman, S., Oittinen, P.: Image Retrieval by End-Users and Intermediaries in a Journalistic Work Context. In: IIiX 2006, pp. 171–187 (2006)
23. Westman, S., Lustila, A., Oittinen, P.: Search strategies in multimodal image retrieval. In: IIiX 2008, vol. 348 (2008)
24. White, R.W., Drucker, S.M.: Investigating behavioral variability in Web search. In: WWW 2007, pp. 21–30 (2007)
25. Zavesky, E., Chang, S., Yang, C.: Visual islands: intuitive browsing of visual search results. In: CIVR 2008, pp. 617–626 (2008)
26. Zhang, L., Chen, L., Jing, F., Deng, K., Ma, W.: EnjoyPhoto: a vertical image search engine for enjoying high-quality photos. In: MM 2006, pp. 367–376 (2006)

Crossmedia Systems Constructed around Human Activities: A Field Study and Implications for Design

Katarina Segerståhl

University of Oulu, Department of Information Processing Science
P.O. Box 3000, FIN-90014 Oulu, Finland
katarina.segerstahl@oulu.fi

Abstract. Many interactive systems today span across a range of interoperable IT artifacts, forming crossmedia systems. They aim at providing pervasive and synergistic support for human activities. This paper reports a three-month-long qualitative field study exploring the use of a crossmedia fitness system to support physical training. The main concern is how the system – through the configuration of its components – supports the primary activity. Users' primary motivation, elaborateness of their activities, internalization or externalization of their actions and their perceived threshold toward using distinct IT artifacts determined the utilization of the system and each of its components. Compositional aspects of the system, such as its hierarchical structure, distribution of functionality and functional modularity influenced its ability to support different ways of training. The article contributes by shedding light on aspects that influence the synergistic use of IT artifacts and by proposing implications for designing crossmedia systems.

Keywords: Pervasive computing, crossmedia systems, interaction design, human activity, case study.

1 Introduction

The purpose of this study is to understand (and ultimately to help design for) a complex interaction – such as where multiple devices and applications are used together to carry out a set of interrelated tasks. The focus is on *crossmedia systems*, i.e., combinations of interoperable media (devices and applications) that are systemically constructed around a distinct space or activity.

Recent emphasis in mobile computing on 'smart' and one-size-fits-all design has led to the "chaos of infinite combinatorial possibilities" [10]. Yet, there exists a manageable range of situated types or recognizable spaces of human activities, e.g., learning, collaborating, shopping, and sporting [10]. Design that is targeted on a few appropriate themes is more likely to be appreciated within those spaces [10]. Alongside the development of all-power-in-one devices, there appears to be growing interest in developing crossmedia systems that are purposefully structured around a specific space. Examples include: a system for realtors to manage open houses and sales [1]; laboratory information management systems [13]; the Nike + iPod, and the Polar training systems for fitness and sports; as well as systems for managing personal

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 354–367, 2009.

information across mobile and non-mobile use contexts (e.g., smartphones that connect with calendars and other applications on laptops and desktop computers).

What is common for all of these systems is that they aim at providing pervasive task support through an optimal configuration of multiple intertwined IT artifacts that each address distinct contextual settings of the primary activity. Devices and applications in such configurations are often retrospectively bridged. This may be the case when existing technologies are combined to form mash-ups, applications are built on legacy platforms, or when the components of a system have been designed in different organizations or teams. Technical interconnectivity is often achieved, but actual interoperability (in terms of affordance for combinatorial use) has not always been supported all the way. Also, the situated and subjective nature of human activity is challenging to design for. However, crossmedia systems hold potential for supporting a range of human activities that could be better utilized by focusing on their structure and composition.

This study investigates the use of a system that is designed to support physical training. As is characteristic of crossmedia, the system is designed with distinct presuppositions about how the training activity is structured. I will investigate, how the system *composition*, i.e., the roles and functionality of its components and the ways in which they are intended to work together, actually corresponds with users' real-world training practices. Based on my observations, I will discuss implications for the design of crossmedia systems in terms of their structure and composition.

In the next section I will briefly introduce the background for the study; then I will describe the field study as well as its findings and analysis. In the final sections I will discuss the implications of the study focusing on the design of crossmedia systems.

2 Background

In 2000 Steven Feiner [5] introduced the concept of hybrid user interfaces in which multiple heterogeneous displays and interaction devices are used in "synergistic combination to benefit from the advantages of each" [5]. One example of this vision in practice is the iRoom which was created in an exploratory project at Stanford University. The iRoom is an interactive workspace where multiple technologies are mapped to a single defined physical location and configured to provide combinatorial support for meeting activities [7]. Today, there are a number of information systems that expand across diverse applications and devices. Yet, despite technical advances and increased interoperability, there are still issues that remain to be addressed regarding the everyday interactions with these systems. This is especially the case when crossmedia systems expand across both stationary and mobile use contexts.

Oulasvirta's and Sumari's [14] research outlines several strategies that users develop and adopt for putting together sets of media to support their tasks. Users employ these behavioral strategies in order to get the best out of the information appliances that they carry around with them, at the same time trying to fit them into their environments [14]. Users constantly find new ways of coping with and using information technology by carrying out the practice of recombinant computing, i.e., serendipitously interweaving diverse media to enrich information and to gain pervasive task support [12, 14]. Feiner predicted that such use of hybrid user interfaces is difficult to predict and structure, and may become so diversified that special 'environment management' will be needed [5].

However, the systems could as well be designed focusing on pre-identified situated spaces. This would provide grounds for incorporating environment management into the systems' composition. Applying the concept of crossmedia to the design of information systems may help to support the different aspects of human activities and form a conceptual basis for designing systems that may better afford the synergistic use of their components.

2.1 Crossmedia Systems and Synergistic Specificity

The term *crossmedia* has emerged in the context of modern communications research converging with the fields of pervasive computing and human-computer interaction (HCI) [20]. According to Filgueiras et al. [6], crossmedia can be defined as "the collaborative support of multiple media to delivering a single story or theme, in which the storyline directs the receiver from one medium to the next, according to each medium's strength to the dialogue." Crossmedia may not only deliver a story or theme, but may also be used for supporting a specific goal or objective through functionality that is distributed across media [3]. According to Boumans [3] the characteristics of crossmedia include that, more than one medium is involved in supporting a message/story/goal, and that the delivery or support of the common message/story/goal purposefully spans across the different media.

What distinguishes crossmedia systems from, for example cross-platform systems is that instead of enabling an application to run on multiple platforms or for content to be accessible through multiple media, it is spread across them for optimization of modalities and interaction resources. Crossmedia systems reach their full intended potential when their components are used in combination, or synergistically. Schilling [16] has termed *synergistic specificity* as "the degree to which a system achieves greater functionality by its components being specific to one another" within a particular configuration. Systems high in synergistic specificity may be able to provide functionality that more modular systems cannot.

The functionality of some systems relies on optimizing the components' ability to work with each other. In such systems, detaching the components or using them in isolation would result to a loss in performance or even paralyze the whole system [18]. An example of high synergistic specificity can be found in the Apple iPod products where the mobile devices are dependent on the desktop application as the source of power and content. There are also crossmedia systems that are more flexible due to a higher degree of modularity, i.e., they can be used in different combinations and parts of them can be operated independently. A more modular system achieves synergy when used in such a way that the benefits of its components are merged.

Crossmedia systems are designed by establishing interoperable combinations of media around a distinct theme (space or activity) and developed with the intent of enabling and supporting the synergistic use of their components. In the next section I will report the findings of a field study that explores the use of a crossmedia system in real-world settings.

3 Field Study: Training with the Polar Fitness System

The focus of our investigation is in how a crossmedia system, through the configuration of its components, supports people in training to maintain and improve their physical fitness. The study was carried out in cooperation with Polar Electro, a world leading manufacturer of heart rate monitoring and sports technology [9]. The Polar fitness system, comprising a wearable heart rate monitor and an interoperable web service was used as an example of a crossmedia system.

3.1 The Polar Fitness System

The Polar fitness system is targeted to active, fitness-oriented, but ordinary users thus excluding, e.g., professional athletes. It comprises a wearable heart rate monitor, the Polar FT60, and a web service, www.polarpersonaltrainer.com. Along with the heart rate monitor, accessories such as a GPS receiver, a heart rate monitoring strap, and a USB dock for transferring data to the web are included.

The wrist unit provides a fitness program with weekly targets that help in managing exercising throughout the week. It also provides information and instruction during exercise and helps the user to achieve desired heart rate zones and other exercise related goals. After each exercise the wrist unit provides feedback and a weekly summary with suggestions for the upcoming week. The web service includes a training calendar and tools for creating long-term training programs and detailed exercise plans, and information and instructions for heart-rate-based exercise. It also provides progress charts, graphs and summaries for analytic and long-term follow up, a place to document exercising and a long-term storage for exercise data. Figure 1 illustrates how the system and its components are intended to support the different aspects of the training activity.

Fig. 1. The Polar fitness system spans across the training activity via its components by providing support for planning, exercise and follow-up

3.2 Method and Approach

A qualitative multiple case field study was conducted with 12 participants over a period of three months. The aim was to investigate and understand the aspects in users' training practices that influenced the use of the fitness system and each of its components. Through this understanding it was possible to analyze how the system structure corresponded with the actual training activity.

The study was announced at two fitness centers and participants could apply by filling in an application. A total of 30 applications were received out of which 12 applicants were selected to take part in the study. The method of purposeful intensity sampling [15] was used for selecting the participants. The sample was constructed to match the criteria of potential real-world users that had previously been identified by Polar. Within these criteria, a sample that expressed diversity in terms of age, gender, and professional background as well as activity and training background was formed.

Four group interviews were carried out throughout the study, each followed by a three-week-long self-documentation period when the participants used the target technology in their daily lives. Participants filled out cross-sectional questionnaires at each group meeting. During the self-documentation periods, structured diaries were filled in by participants on a daily basis. During the study each participant was observed and interviewed. Observations were one and a half hours each and they were conducted applying the principles of contextual inquiry [2]. The inquiries were carried out either at a gym when the participant was performing a regular workout or by joining the participant's run or jog. Research instruments such as diaries and questionnaires were developed upon analysis throughout the study. See Fig. 2 for an overview of the data collection process. All inquiries were recorded and documented with notes. Interviews were recorded and professionally transcribed and diary and questionnaire input were extracted to tables for analysis. The grounded theory [17] approach was applied in systematic coding of the data and in discovering emergent categories.

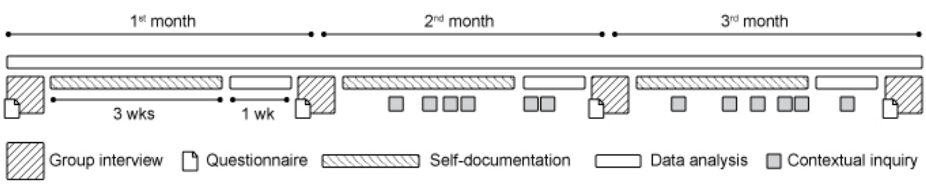


Fig. 2. The data collection process

3.3 Participants

Participants were screened according to the pre-identified characteristics of potential real-world users. They were between ages 23 and 40. Seven of them were female and five were male. All participants were active in sports and reported training approximately three to five times a week with a history of regular exercise.

The participants' exercising activities included jogging, cycling, and indoor exercise such as gym workouts or taking group exercise classes. Participants came from various professional backgrounds. Within this group of subjects, it was found that basic skills regarding everyday IT (e.g. PC/laptop, mobile phone, digital camera) use were well established. Experience using internet for daily or utilitarian purposes (browsing, e-mail, internet banking) was also generally common. However experience of combinatory use of multiple devices was not common. The participants and their utilization of the target technology will be investigated in more detail in the following sections.

4 Findings

I will begin by describing the participants and the reasons they had for training. Then I will describe the ways in which they carried out their training related activities. After this I will tackle the use of the fitness system and its components. Finally, I will discuss how the use of target technology and the ways of training fluctuated over the course of the study.

4.1 Objectives of Training

As the participants' primary motivation for training and the ways in which they carried out their planning, exercise and follow-up activities were analyzed, three categories of motivational types were identified that corresponded with styles of training. These are *improvement* (n=4), *self-monitoring* (n=3) and *maintenance* (n=5) oriented individuals.

The improvement oriented valued objectives such as performance, fitness, health and enjoyment generally higher than appearance and weight. They also more often set long-term goals for their training, such as an event or a specific improvement in their performance. As one of them told about her goal:

P#09: *"I'm training for my first marathon. To train and improve technique, I'm attending a running school. I also train with my personal trainer once every week. I plan my other exercises so that they support my running practice. "*

The self-monitoring subjects considered all objectives, performance, fitness, health and enjoyment, as well as appearance and weight as equally important. However, instead of looking to achieve a specific improvement or goal, they were more curious about how their training generally influenced their personal state regarding each of the objectives. Self-monitoring subjects would enjoy collecting and analyzing data about themselves and their actions even if there was no specific goal or event that they would do it for.

The maintenance-oriented subjects were very much aware of the importance of training and they felt the objectives as the end toward which they should train regularly. As opposed to the improvement and self-monitoring oriented, they would not, however, take on an analytic approach to their training. They had more often difficulties in maintaining a regular exercise regimen and in staying motivated.

The differences identified in the analysis of training objectives were reflected in subjects' individual training styles and became meaningful in the ways in which they would utilize the target technology. The following section will explore the different styles of training that were identified.

4.2 Styles of Training

The improvement oriented subjects carried out their planning, exercise and follow-up activities in relatively elaborate ways. They would systematically plan their training several weeks or months in advance. Their exercises were often goal-oriented, including detailed interval-exercises for improving performance as well as lower intensity ones to train endurance. They analyzed their performance and progress with calendars, diaries and notebooks usually in relation to their long-term goal.

The self-monitoring oriented subjects were also prone to elaborateness through their curiosity and eagerness to constantly learn more about themselves and their training. Most of them would plan ahead one week at a time (as opposed to training serendipitously or in longer cycles). They would set weekly goals often with respect to an abstract idea of what should be accomplished in terms of diversity and intensity of exercise. When exercising, they often conformed to moderately structured exercises (e.g., warm-up/work/cool-down). The self-monitoring subjects carried out analysis of exercise amounts and intensities with respect to their weekly targets. They would often revisit their exercise data for motivation.

The maintenance oriented subjects, instead of planning or analyzing, would more often exercise on an ad hoc basis. They would not usually plan ahead, but fit exercise wherever they could. Sometimes they would mark down their exercises into calendars and, for motivational purposes, look back at what has been accomplished.

The identified styles of training were not constant. Some fluctuation was observed over the course of the study among all participants. The training activity was fairly sensitive to participants' various life situations, e.g., periodic changes in their work schedule, holidays, and situations in which they needed to prioritize time for family and friends. This became apparent in the participants' diaries:

P#12: *"A couple of busy weeks ahead, going out of town for three days for business and got a couple of exams coming up. So for the next two weeks I will train whenever I can find time. Then I'll hopefully get back into my normal rhythm."*

4.3 Use of Target Technology

Over the first three weeks of the study the subjects were very enthusiastic about the target technology, namely at this point the heart rate monitor. Their threshold to appropriate it and learn its use was relatively low. After three weeks had passed, participants were explicitly encouraged to explore the web service. This was necessary in order to collect data on the potential uses of the different components. As the users finally started to explore the web service and interoperate the two media, the dynamics behind the process begun to show. Fig. 3 illustrates some of the individual differences in the use of the system that were identified.

The maintenance oriented would use the heart rate monitor for motivation and guidance, but did not see much use in the web service. Four of them had little experience of using the web or any other technologies for anything but mandatory purposes (such as banking) and were reluctant to invest time into learning the use of the web service. They also did not believe that it would provide additional value.

One of them tried it out and was at first very interested, but after a week found that it just wasn't useful enough for her as she only trained a week at a time and for that the heart rate monitor was adequate.

A similar case occurred among the self-monitoring type, where a participant reported being interested in the web service when first exploring it, but abandoned it later because it was too advanced or complicated. She continued to use her traditional calendar and diary for follow up.

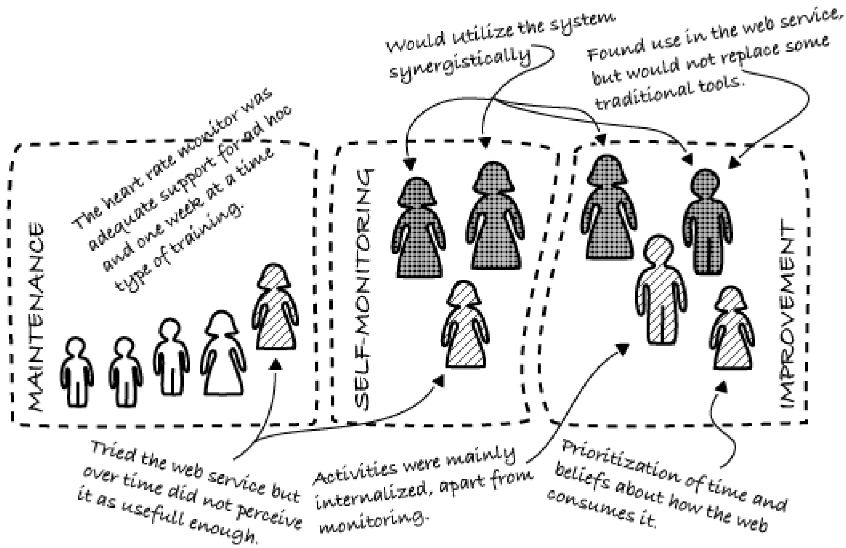


Fig. 3. Summary of individual differences in how participants used the system. Individuals' size indicates the relative elaborateness of their activities and their position indicates their proneness towards utilizing the web service as an addition to the heart rate monitor.

P#05: *"I uploaded my exercises online earlier and it was pretty interesting. But there's no way I can find time for really using it [the web service]. The wrist unit is great, because it's always with me when I exercise and it still stores all my weekly exercises. Right now I just don't need anything more detailed for my purposes."*

The self-monitoring and the improvement oriented participants that carried out their exercise at a higher level of elaborateness were generally more interested in the web service. They would carry out their activities at a level where actions more often needed to be externalized with tools, such as notebooks, graphs, and training diaries. For example, one of the participants was interested in running a marathon and felt that she needed all the support to accomplish her training as efficiently as possible. For this purpose, she perceived the web service as useful.

P#09: *"For the past two years I've kept a diary where I've been writing down almost everything about my training. It's so much more convenient to be able to transfer all that information automatically to the computer and it's also easier to analyze it there."*

Half of the improvement oriented users would find the web service useful for them, especially in that it provided long-term training programs. They would use these to refresh their exercise routines and to make them more efficient.

There were two participants among the improvement oriented who were less prone towards the web service. One of them had most of his training internalized, i.e., he could memorize the weights he was working out with and had his entire exercise plan 'in his head'. He also had a clear perception of his longer term goals and was able to follow them as he trained. He had years of experience from team sports, which probably influenced his ability to carry out systematic training with few external tools. He used the heart rate monitor mainly to regulate intensities during exercise and to collect data. The other participant was in a life situation where she had very little time to begin with and perceived the web as being a "time-thief." She was very reluctant to use the web service for her training and felt that it would compromise the little time that she had with family and for actual exercising. She managed her personal goals and improvement with the heart rate monitor and used its fitness-trend feature and weekly summaries to make judgments about her progress.

5 Analysis and Discussion

The findings of the study were organized and analyzed using some concepts from activity theory. Activity theory [8,11] is one of the "contextualist" or "post-cognitivist" theories used in current HCI as a framework for understanding and analyzing the role and use of modern technology in mediating human activity.

Activity theory is constructed around five key principles, which include 1) *intentionality*, proposing that an activity is always targeted at an object, i.e., the motivation of the 'doing'; 2) *hierarchical structure of activity*, suggesting that depending on the level of analysis, an activity can be broken down into temporally and hierarchically related components; sub-activities, actions and operations; 3) *internalization/ externalization*, meaning that actions may be internalized, i.e., become unconscious or automatic like riding a bicycle, or as actions become more complex, they may be externalized with tools, e.g., using a spreadsheet to carry out complex calculations; 4) *tool mediation*, emphasizing that both physical as well as conceptual tools or artifacts are a fundamental part of any human activity and play a distinct role in shaping it; and finally 5) *historicity/development*, stressing that past events contribute to tools' current use and shape. In this study I focus on the mediated relationship on an individual level. As the underlying motivation is in advancing the design of crossmedia systems, I also put special emphasis on the utilization of the target technology across the different cases of training. Using constructs from activity theory as tools, I will first summarize the dynamics of utilization that were observed in this study. Then I will analyze, how real-world practice influenced the use of technology and then, based on these analyses, I will draw implications for how crossmedia systems could better respond to human activity.

5.1 Dynamics of Utilization

The participants' primary motivation together with the level of elaborateness that they trained on, were the primary determinants for their proneness towards adopting and using multiple media. The need for external tools was also determined by the level of internalization or externalization of the users' actions. Individual, experiential and social issues as well as the users' life situations also influenced the participants' utilization threshold (or proneness) towards the target technology and its components. (See Fig. 4)

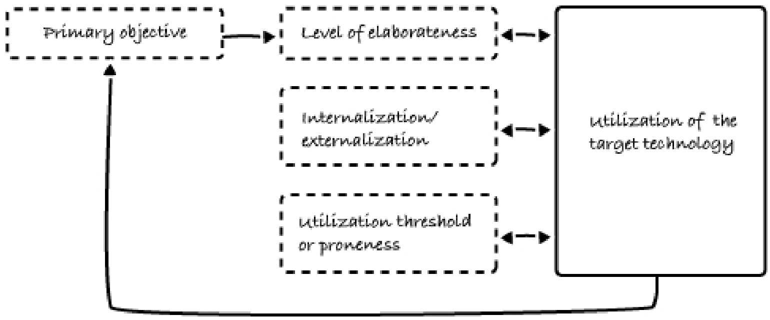


Fig. 4. The dynamics of utilization

As it was used, the fitness system also changed the ways in which subjects trained and sometimes even their primary objectives. For example, one of the participants discovered how he could use heart rate information to regulate his recovery times between weight lifting sets when he was working out at the gym. Another participant would train regularly to maintain his health at the beginning of the study, but set new, more improvement oriented goals for himself as the tools, i.e., the expert training programs, became available to him.

By breaking down the dynamics of utilization and understanding the role of the target technology and its components in each case, it is possible to identify distinct ways in which crossmedia systems could better respond to human activities.

5.2 Activities in Practice

In addition to identifying the dynamics of utilization, it is important to understand how they are actuated and shaped in practice. When participants' training was studied in practice, it was found that there were a number of other personal activities, such as work and socializing, that penetrated into the area of the supposed ideal training activity and sometimes distorted it. People's activities are not such isolated ideals, and therefore these ideals cannot function as the sole basis of system design.

The underlying ideal of a distinct situated type of the training activity, fundamentally characterized the composition of the fitness system. One example of this is the assumption that there is such a thing as systematic training that would require more advanced tools (such as the web service) for support. It was assumed that users who would train in this way would find the functionalities as they needed them. However, this need only became apparent when the participants were encouraged to explore the web service – the ideal of the activity did not exist in practice as strong enough to trigger proactive exploration for additional support. Apparently the conceptual representation of a human activity behind the system is not enough – it needs to actively support the synergistic use of its components and support users in finding the functionalities that they may benefit from.

The idea of a distinct situated type or activity is to some degree a functional one when designing crossmedia systems. It is the first step in defining the main task domains that a system should span as well as its basic requirements. However, compositional aspects that allow for crossmedia systems to respond to the fluctuating nature of human activity also need to be taken into account.

5.3 Implications for Composition of Crossmedia Systems

Based on my understanding of how a crossmedia system – through the configuration of its components – may support the different ways of carrying out an activity as well as respond to its individual fluctuations, I will now draw implications for design.

Hierarchical Structure. Each of the devices and applications within a crossmedia system has a distinct role, i.e., each addresses a specific domain of the activity as their primary. For example, the role of the heart rate monitor was to support the exercise context, whereas the role of the web service was to support the planning and analyzing contexts. The roles of the components should not only be considered with respect to the primary activity but also with respect to each other. Users did not always conform to the ideal structure of the training activity. For example, depending on their life situation, they would sometimes not plan their exercising at all and leave follow-up for later. Exercise, however, was the core activity that could be carried out independently of the other, supporting activities. For exercising, users relied on a single component, the heart rate monitor, as their primary and incorporated the secondary component, the web service, when they needed or had time for it. Thus, the heart rate monitor was the primary component in the system as the source of content (exercise data) and core functionality. Incorporating a hierarchy of components into the system's configuration may increase the system's ability to respond to different user needs and fluctuation of practices.

Distribution of Functionality. The very idea of crossmedia systems is to optimize interaction resources for the different contextual settings that may occur throughout a human activity. When functionalities are appropriately distributed or specialized across the different media within the system they will each contribute to the synergy with their individual strengths. The complexity of a system's individual components as well as redundancy of functionality and content may be reduced by distributing functionality according to the system's hierarchical structure. This means that more advanced functionality, that is likely to benefit a portion of users, could be assigned to an additional media, not necessarily the primary one. Users, who need extended support for their activities, as was observed, may more often have a lower threshold for adopting multiple media. However, with some this threshold is still very high. Users who need the more advanced functionalities, need to be better supported in extending their use practices across more of the systems' components for holistic support. In contrast, users who do not need all the possible functionalities would not be burdened with their added complexity on a single medium.

Functional Modularity. The roles that are assigned to each component help to identify their primary strengths and functions in the system, however, these roles and strengths should not be too strictly defined. As was found, people may carry out an activity in many different ways due to their overall objectives, experience and life situation. Human activities are fluctuating. They may vary by individual and also over time within one's context of life. In order for a crossmedia system to be able to support activities throughout this fluctuation, a certain degree of functional modularity needs to be maintained. Functional modularity means that regardless of the intended synergistic structure of the system, its components or at least its primary or master component should allow for it to be operated to some degree also independently. For example, in addition to supporting exercise, the heart rate monitor provides limited support for planning and follow-up as well. For more advanced support, users may use the web service, but when they do not have time for it or their threshold to using it is too high, they will be adequately supported by the heart rate monitor as well. The degree of functional modularity is determined by the mutual dependencies among the components and can be applied when the components' hierarchical position allows for it.

Crossmedia systems may provide optimal means for supporting many levels of a human activity by providing layered support through their various components. Users who carry out their activities at a lower level of elaborateness, do well using only a part of the system. Whereas more advanced levels of an activity may be supported by providing more extensive functionalities through additional media. A high level of synergistic specificity, i.e., making it impossible to use one component without the use others, may intimidate users that only need a certain level of support and perceive the threshold to use the whole system as too overwhelming. However, having the opportunity to extend the use of tools as the nature of one's activity evolves over time, towards more elaborate or advanced, is one of the benefits of these kinds of systems. The roles of the system components need to be identified with respect to the primary activity, as well as with respect to each other. The hierarchical structure of the system is determined by the roles of, and dependencies among its components. Identifying this hierarchical structure provides a framework by which distribution of functionality may be applied. Functional modularity helps to maintain a moderate degree of flexibility that enables the system to be used on several different levels of an activity and by different users.

Even though some of the implications presented above may be applicable outside the scope of this particular study, they are still based on an analysis of a specific system and activity domain. Further research is needed in order to form a broader understanding of how crossmedia systems could be structured in various application environments, such as learning, and shopping. It would also be interesting to investigate, whether there are distinct compositional patterns that support some situated types better than others. Despite being designed to enable synergistic use practices, crossmedia also needs to *actively* support them and for that, the development of interaction techniques that lower users' thresholds to adopt synergistic use practices is needed. An interesting direction for research and development would be systems or IT environments that combine context aware and ubiquitous computing with the concept of crossmedia to support everyday life through the activities that constitute it.

6 Conclusion

This paper reports a three-month-long qualitative field study with 12 participants using a crossmedia fitness system, comprising a heart rate monitor and a web service, to support their training. The aim was to investigate how the composition of the crossmedia system influenced the ways in which it and its components would be utilized. It was found that the quality of users' primary motivation to exercise, elaborateness of their activities, internalization or externalization of their actions as well as their perceived threshold to using distinct IT artifacts determined the utilization of the system and each of its components. Structural and compositional aspects of the system influenced its ability to support different ways of training.

Implications are focused on the design of crossmedia systems in terms of their structure and composition. A system structure that corresponds with the target human activity and incorporates a hierarchical structure of roles among its components may be able to provide a clear system image and appropriately layered support for the primary activity. However, a certain degree of modularity needs to be maintained in order to support the different levels of elaborateness that an activity may be carried out on. Further research is needed to form a broader understanding of how crossmedia may support other situated types or human activities as well as to identify different compositional patterns that may be applied to crossmedia.

Acknowledgements. I wish to thank the RichWeb project, National Technology Agency of Finland (TEKES), Graduate School of Software Systems and Engineering (SoSE), Oulun Yliopiston Tukisäätiö, and Polar Electro Oy for funding parts of this research. I am grateful to Marja Harjumaa for her contribution to the study. I also wish to thank Kari Kuutti and Tanja Kotro for commenting the manuscript and the Stanford HCI Group, Harri Oinas-Kukkonen, Esa Juusola, Tanja Kalliojärvi, Sirpa Hope, and Petteri Siekkinen for their support and Svante Kärkkäinen and SATS Oulu fitness centers for cooperation.

References

1. Aneja, S., Lim, Y.-K.: Designing for Totality of Mobile and Non-Mobile Interaction: A Case Study. In: Ext. Abstracts CHI 2007, pp. 1685–1690 (2007)
2. Beyer, H., Holzblatt, K.: Contextual Design: Defining Customer-Centered Systems. Morgan Kaufmann Publishers/ Academic Press, San Francisco (1998)
3. Boumans, J.: Cross Media: E-content report 8. Published in a series of E-Content Reports by ACTeN (2004), `http://www.acten.net`
4. Edwards, K.W., Newman, M.W., Sedivy, J.Z.: Building the Ubiquitous Computing User Experience. In: Ext. Abstracts CHI 2001, pp. 501–502. ACM Press, New York (2001)
5. Feiner, S.K.: Environment Management for Hybrid User Interfaces. IEEE Personal Communications, 50–53 (October 2000)
6. Filgueiras, L., Correa, D., Oliveira Neto, J., Facis, R.: X-gov planning: how to apply cross media to government services. In: ICDS 2008, pp. 140–145. IEEE, Los Alamitos (2008)
7. Johanson, B., Fox, A., Winograd, T.: Experiences with Ubiquitous Computing Rooms. IEEE Pervasive Computing, 67–74 (2002)
8. Kuutti, K.: Activity Theory as a Potential Framework for Human-Computer Interaction Research. In: Nardi, B. (ed.) Context and Consciousness. Activity Theory and Human-Computer Interaction. The MIT Press, Cambridge (1996)
9. Laukkanen, R., Virtanen, P.: Heart Rate Monitors: State of the Art. Journal of Sports Sciences 16(3) (suppl. 1), 3–7 (1998)
10. McCullough, M.: Digital Ground. Architecture, Pervasive Computing, and Environmental Knowing. The MIT Press, Cambridge (2005)
11. Nardi, B.: Context and Consciousness. Activity Theory and Human-Computer Interaction. The MIT Press, Cambridge (1996)

12. Newman, M., Sedivy, J., Neuwirth, C., Edwards, W.K., Hong, J., Izadi, S., Marcelo, K., Smith, T.: Designing for serendipity: supporting end-user configuration of ubiquitous computing environments. In: Proc. of the 4th conference on Designing Interactive Systems, London, pp. 147–156. ACM, New York (2002)
13. Nilsson, M.: Advancements and Trends in Medical Case-Based Reasoning: An Overview of Systems and System Development. In: Proc. International FLAIRS Conference, Miami, USA, pp. 178–183 (2004)
14. Oulasvirta, A., Sumari, L.: Mobile Kits and Laptop Trays: Managing Multiple Devices in Mobile Information Work. In: Proc. CHI 2007, pp. 1127–1136. ACM Press, New York (2007)
15. Patton, M.Q.: Qualitative evaluation and research methods, 2nd edn. Sage Publications, Newbury Park (1990)
16. Schilling, M.A.: Towards a general modular systems theory and its application to interfirm product modularity. Academy of Management Review 25, 312–334; Reprinted in Garud, R., Langlois, D., Kumaraswamy, A. (eds) Managing in the Modular Age: Architectures, Networks and Organizations. Blackwell Publishers, Oxford (2000)
17. Seaman, J.: Adopting a Grounded Theory Approach to Cultural-Historical Research: Conflicting Methodologies or Complementary Methods? International Journal of Qualitative Methods 7(1) (2008)
18. Simon, H.: The Architecture of Complexity. Proceedings of the American Philosophical Society 106, 467–782 (1962)
19. Steen-Johnsen, K.: Globalized Fitness in the Norwegian Context: The Perfect Meets the Popular. International Review for the Sociology of Sport 42(3), 343–362 (2007)
20. Wiberg, C., Jegers, K., Bodén, J.: Cross Media Interaction Design. In: Presented at the workshop HCI and New Media Arts: Methodology and Evaluation at CHI 2007 (2007)

Query Suggestion for On-Device Troubleshooting

Frédéric Roulland, Stefania Castellani, Ye Deng, Antonietta Grasso,
and Jacki O'Neill

Xerox Research Centre Europe
6 Chemin de Maupertuis 38240 Meylan, France
firstname.lastname@xrce.xerox.com

Abstract. This paper describes a novel query suggestion tool we have designed and implemented to help users of office printing devices better formulate their queries, while searching a troubleshooting knowledge base provided as a service on the device itself. The paper traces the main motivations of the design of the query suggestion tool and outlines its technical details with an emphasis on its combination of features in relation to prior work.

1 Introduction and Motivation

This paper describes a novel query suggestion tool that we have designed to help users more easily formulate their queries when searching solutions while troubleshooting a printing device. The tool is part of an on-device support that lets users search a troubleshooting knowledge base (TKB). The problem that we are considering is one of search by non-experts in the specific technical domain of office device troubleshooting. To investigate troubleshooting practice, we carried out an ethnographic study of the call centre [1]. An immediate finding from the ethnography was that customers' telephones, for communicating with experts, and PC's, for online help, were rarely located by the printer/copier. It is an extra burden for customers to have to leave the ailing device to find out how to fix it. This inspired the design of a system to access the TKB on the device. The ethnography also revealed the work that customers and experts put in to co-construct suitable problem descriptions, suggesting that users are likely to have difficulties specifying their problems without expert help. There is a major terminology issue: customers do not always know the technical names of parts or how to best describe the problem they experience such that they can retrieve troubleshooting instructions from the TKB [9]. Initial designs to improve the online service included features designed to help customers better explore query results sets [10]. However, users still had problems formulating their initial query. Moreover, turning to designing on-device access to the TKB we also had to take into account a number of characteristics of its situated use. Although larger than normal copier/printer screens, the interface is relatively small. Sizing and spacing need to be set to 'finger-size'. In addition, a soft keyboard was required for entering queries. Together these conditions provided the inspiration for the query suggestion tool.

In this paper we present the details of the working of the tool with an emphasis on its combination of features in relation to prior work.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 368–371, 2009.

2 Related Work

The problem of how to support people while they search for information and they are not sure exactly what it is they are looking for has been studied both from the human (e.g. [2]) and the technology side. On the human side, it is important to consider the search domain. When considering searches of the web the quantity of content is such that almost any search will return results and the problem is one of appropriately filtering the results. In comparison, in domain specific databases, there may be much fewer and less diverse content and the problem of finding the right terminology becomes more central [3]. On the technology side, the predominant approach has been the one of Information Retrieval, based on the model of matching user's requests to the documents' content by using keywords. This model led to a first generation of systems used to access both the unstructured content of the Web and domain specific databases. However, a problem with these systems is that the users tend to be very generic in the formulation of their queries which often only contain two or three keywords [4]. One approach to address these issues is relevance feedback [5], which incorporates user feedback on the relevance of the retrieved documents and uses it to provide more precise answers. However, outside of the laboratory, this mechanism has drawbacks due to the lack of immediate incentives for the users to "instruct" the system to perform more precisely. Another approach is to complete query formulation through *query expansion* (e.g. [6]). However these systems, in an attempt of saving the user from additional effort, work as a black-box and remain too generic to really help the user in carving appropriate queries [7]. To go toward systems where the user is more active, recently there has been the development of *query suggestion* techniques. For example in [8] query suggestion has been tested on a domain specific search engine proving improved precision and recall of the results. Additionally, query suggestion, since it asks the user to pick just one of the suggested terms, seems to be more appropriate than query expansion when inter-document variation is small but the user has no prior knowledge of the terms used in the documents. This is the situation for example in the TKB in which editors enforce a controlled vocabulary for technical terms. Interestingly, several web search engines, e.g. Google, have introduced query suggestion mechanisms, taking the popularity of various searches into account in the ranking. However, little work has been done in investigating how to tune these mechanisms to domain specific contexts as our, which has led us to the design of a novel mechanism mixing popularity and domain vocabulary based suggestions.

3 Query Suggestion

The design of our query suggestion tool has been guided by the aim of improving the quality of the searches and reducing the burden of typing. Our approach can be summarized as follows: a collection of expressions candidate for suggestion is created from the TKB, so that only queries that would retrieve at least one result from the TKB are taken into account. They are then ranked according to the frequency of their occurrence in past queries and the relevance of the suggestions in the TKB, so that top ranked suggestions are both likely to correspond to a valid user problem description and to provide good search results. The collection of query candidates for suggestion is created from the TKB contents, indexed using Natural Language Processing techniques to identify noun phrases, and the queries examined as containing a noun phrase used in the system. We decided to use the frequencies of queries in the ranking as a way to favour popular suggestions. Indeed, an analysis on a sub-set of queries of the TKB logs showed us that users can be observed to have some favourite ways of expressing their common problems. To compute the frequency of a suggestion we compare the set of lemmatized meaningful words of the expressions selected from the TKB with the set of lemmatized words computed for each past query stored in the logs. When the set of words of the expression is a subset of the set of words of the old query we increment its frequency. In addition, the content and structure of the TKB is also used to rank suggestions. Expressions that, if used as a query, would generate results scored highly by the search engine are favoured. For computing the score component of a suggestion a search using the suggestion is performed and the list of results for this search is retrieved. We simply use the highest score of the results returned by the search. This parameter balances the previous one and helps differentiate suggestions that cannot be effectively ranked from the frequencies in the logs.

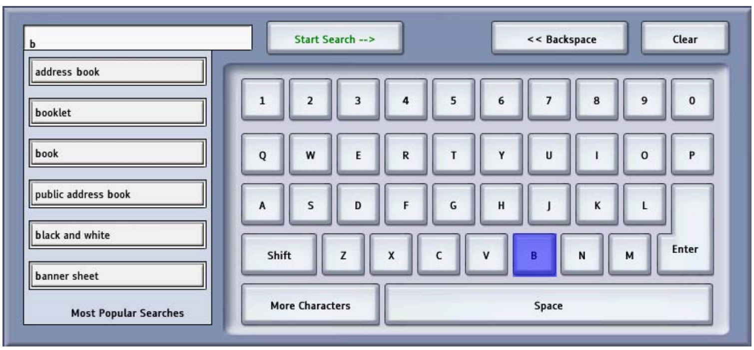


Fig. 1. The touch screen query interface

The query suggestion tool is part of a troubleshooting system accessible on the device and assist users when they want to search the TKB. The query interface is shown in Figure 1. It consists of a simplified keyboard and, on the left side of the screen, a text box that displays the query typed by the user below which is the list of the “best” six suggestions provided by the tool. At any moment the user can select one of them from the list and the user input will be replaced with the selected suggestion. As in Google Suggest[1], when typing queries, the sequence of characters already typed in is used to filter the ranked collection of suggestions. The filtering we perform is however different from Web query suggestion systems, as it needs to point people toward expanded technical expressions they may not know or let them understand what they are using may be appropriate only in part. To achieve this goal, in our proposal, valid suggestions are ones that contain words starting with what is being typed or that contain the words already typed by the user. For example, if the user is typing “li”, suggestions could be expressions with one word starting with “li” like “light” but also

[1] http://www.google.com/support/websearch/bin/answer.py?answer=106230&hl=en

"black lines". In order to implement the filtering, the query is tokenized into words which are then lemmatized. Then, a search is performed over the list of suggestions, returning every expression containing the set of words of the query. When the query contains several words and no suggestion is available for the whole expression, we propose suggestions for the longest, latest part of it. For example, if the user has entered "noise car" (for "noise in the cartridge"), and the system does not find any suggestions for the whole input then it looks for suggestions for "car" returning "cartridge", "toner cartridge", etc.

4 Conclusion and Future Work

We designed the query suggestion tool to help users formulate their queries by making the contents of the TKB available as users type queries and by reducing the burden of typing on the soft keyboard. Some preliminary tests of the tool, performed in order to both understand how quickly accurate query suggestions would appear on the interface and investigate the tool's usability, produced promising results and we are using them to refine the tool for which we will perform more extensive user tests.

References

1. Castellani, S., Grasso, A., O'Neill, J., Roulland, F.: Designing Technology as an Embedded Resource for Troubleshooting. Journal CSCW special issue on CSCW, Technology and Diagnostic Work (2009)
2. Crabtree, A., O'Neill, J., Tolmie, P., Colombino, T., Castellani, S., Grasso, A.: The practical indispensability of articulation work to immediate and remote help-giving. In: 20th CSCW, pp. 219–228. ACM, New York (2006)
3. McCray, A., Tse, T., Understanding, T.: search failures in consumer health information systems. In: American Medical Informatics Symposium, pp. 430–434 (2003)
4. Silverstein, C., Marais, H., Henzinger, M., Moricz, M.: Analysis of a Very large Web Search Engine Query Log. In: ACM SIGIR Forum, vol. 33(1), pp. 6–12 (1999)
5. Salton, G., Buckley, C.: Improving retrieval performance by relevance feedback. Journal of the American Society for Information Science 41(4), 288–297 (1990)
6. Qiu, Y., Frei, H.P.: Concept Based Query Expansion. In: 16th SIGIR, pp. 160–169. ACM, New York (1993)
7. Muramats, J., Pratt, W.: Transparent Queries: Investigating Users' Mental Models of Search Engines. In: 24th SIGIR, pp. 217–224. ACM Press, New York (2001)
8. Feuer, A., Savev, S., Aslam, J.A.: Evaluation of phrasal query suggestions. In: Sixteenth ACM CIKM, pp. 841–848. ACM, New York (2007)
9. O'Neill, J., Grasso, A., Castellani, S., Tolmie, P.: Using real-life troubleshooting interactions to inform self-assistance design. In: Costabile, M.F., Paternó, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 377–390. Springer, Heidelberg (2005)
10. Roulland, F., Kaplan, A., Castellani, S., Roux, C., Grasso, A., Pettersson, K., O'Neill, J.: Query Reformulation and Refinement Using NLP-Based Sentence Clustering. In: Amati, G., Carpineto, C., Romano, G. (eds.) ECiR 2007. LNCS, vol. 4425, pp. 210–221. Springer, Heidelberg (2007)

Acquisition of Animated and Pop-Up Targets

Guillaume Faure[1,2], Olivier Chapuis[1,2], and Michel Beaudouin-Lafon[1,2]

[1] LRI -- Univ. Paris-Sud & CNRS, Orsay, France
[2] INRIA, Orsay, France

Abstract. Pop-up targets, such as the items of popup menus, and animated targets, such as the moving windows in Mac OS X Exposé, are common in current desktop environments. This paper describes an initial study of pointing on pop-up and animated targets. Since we are interested in expert performance, we study the situation where the user has previous knowledge of the (final) position of the target. We investigate the effect of the DELAY factor, i.e. the delay before the target pops up (for pop-up targets) or the duration of the animation (for animated targets). We find little difference between the two techniques in terms of pointing performance (time and error), however a kinematic analysis reveals differences in the nature of the pointing movement. We also find that movement time increases with DELAY, but the degradation is smaller when the target is farther away than when it is closer. Indeed, larger distances require a longer movement time therefore the target reaches its destination while the participant is still moving the pointer, providing more opportunity to correct the movement than with short distances. Finally we take into account these results to propose an extension to Fitts' Law that better predicts movement time for these tasks.

Keywords: Pop-up targets, Animated targets, Movement analysis, Fitts'.

1 Introduction

Pointing is a fundamental action in graphical user interfaces (GUIs) that has been the subject of much research to both understand pointing actions and improve pointing techniques. Most of this research has focused on static targets that are displayed before pointing starts and stay still during the pointing action. GUIs however have always featured *pop-up targets* that appear after the pointing movement starts, e.g., pop-up menus and dialog boxes. More recently, GUIs have also started to feature *animated targets* that move while the pointing gesture is being performed, e.g., the windows in Mac OS X Exposé. While we know that animation enhances user interaction with the system [13, 24] by providing a continuous feedback that increases the user's sense of direct and indirect interaction [26], its effect on pointing has not, to the best of our knowledge, been studied. In particular, beyond the empirical evidence that users take advantage of their knowledge of the behaviour of targets to anticipate their apparition or final position, we are not aware of any systematic study of this phenomenon.

This paper presents what we believe is the first controlled study designed to better understand the acquisition of animated and pop-up targets. Our main factor is DELAY, the delay before the target pops up (for pop-up targets) or the duration of the

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 372–385, 2009.

animation (for animated targets). We have focused on short values of DELAY — between 0 and 500*ms* — which are typical of GUIs, and we have operationalized the common situation where users can anticipate the final position of the target by ensuring that they have prior knowledge of its location.

The experiment compares pointing performance (in time and error) for static, animated and pop-up targets under various values of DELAY. We found a strong effect of DELAY: movement times increase with the delay, but the degradation is smaller for longer distances to target than for shorter ones. We also found little performance difference between pop-up and animated targets, but an interesting qualitative difference when looking at the kinematic profiles. In particular, this analysis shows evidence of how users anticipate the final position of the target.

Finally we show that Fitts' Law does not accurately predict movement times for pop-up and animated targets. Using the insights gained by the analysis of the experimental results, we propose an extension to Fitts' Law that takes into account the unique characteristics of these tasks to better predict movement time.

The rest of the paper is organized as follows. After some background and related work, the next section provides a set of examples and motivates the present study. The subsequent sections describe the experiment, present the performance results (movement time and errors), analyse the kinematic aspects of typical movements (velocity vs. distance and number of sub-movements) and propose an extension to Fitts' Law. The paper then concludes with a discussion and directions for future work.

2 Background and Related Work

The acquisition of an on-screen target by means of a pointing device (*pointing* for short) is one of the basic tasks in Graphical User Interfaces and one that has been extensively studied in Human-Computer Interaction (HCI). Previous research in this area includes novel techniques to facilitate target acquisition, comparison of input devices performance, and models to predict movement time and better understand target acquisition tasks of various types. The major theoretical tool for studying pointing is Fitts' Law [11], which predicts the movement time to acquire a target of width W at a distance D. The most widely used form of Fitts' Law in HCI [19] is:

$$MT = a + b\log_2\left(\frac{D}{W}+1\right)$$

where a and b are empirically determined constants and where $\log_2(D/W+1)$ is called the *index of difficulty* (*ID*) of the task. While a number of variations of this law have been proposed (see [23] for a review), the Psychology literature has not reached a consensus yet on an explanation of this Law.

The most popular explanation of Fitts' Law in the HCI community is Meyer et al.'s sub-movements model [20]: first, a fast sub-movement towards the target (the ballistic phase of the movement) is performed, then if this movement does not hit the target, another (probably smaller) sub-movement towards the target is performed, and this process continues until the target is reached. An important characteristic of the sub-movements model that may explain the logarithmic form of Fitts' Law is that the distance vs. velocity graph shows a bell shape with a clear velocity peak.

The originality of our study is that the target is not always visible on the screen during the acquisition movement. Most previous work in HCI has investigated the acquisition of an initially invisible target using *navigation* tasks, e.g., scrolling and pan-and-zoom, rather than a pure pointing task. A notable exception is Cao et al. [7] where the targets are initially invisible and are revealed by moving a display window attached to the cursor (coupled cursor). The pointing task in that case involves two phases: first, reveal the target, then acquire it. In our study the targets are revealed automatically (without user action) and users know in advance where the targets are so they can anticipate their final position. We therefore expect a more integrated motion than in the above study, with dynamic adaptation to the target behaviour.

Some studies have shown that the shortest delay needed for sensory information to affect hand movement is approximately 100 *ms* [17]. For example Flash and Henis [12] have observed movement adaption between 100 and 200 *ms* after an experimentally controlled modification of the cursor trajectory. Closer to HCI, Zhai et al. [28] have shown that users can take advantage of *unpredictable* target expansion, suggesting a dynamic adaptation of the user's pointing movement.

Menus are the main example of pop-up targets and have been widely studied in HCI. With linear menus, it is assumed that the menu pops up immediately and studies have focused on factors such as the number of items and the depth of the hierarchy of menus rather that the pop-up delay (see, e.g., [9]). With marking menus [18] and their variants, there is a distinction between novice mode, where the user *waits* for the menu to pop up, and expert mode, where the menu does not appear at all. Here, the absence of motion during the delay is used to activate novice mode (see also [14]).

Regarding animated targets, previous studies in Psychology, e.g., [15], and in HCI [21] have addressed the acquisition of moving targets. These studies however have considered *capturing tasks*, i.e. acquiring the target while it is moving whereas in our case the target can be acquired only when it has reached its destination. Other examples that involve moving targets include facilitation techniques that bring the targets (with a possible animation) close to the cursor after the movement has started, e.g., Drag-and-pop [4] and Vacuum [5]. These techniques however make it difficult for the user to anticipate the final position of the target.

Memorization of target positions is an important point of our experimental design and has been the subject of previous work, e.g., Hornof and Kieras [16]. However, we only consider the memorization of a single target in short term memory. This phenomenon has been studied in the domain of Human Vision (see, e.g., [2]) where it has been shown that humans can recover spatial positions via saccades (rapid eye movements) from a few memorized gaze positions.

Surprisingly, we have not found in the literature any previous study of pointing on memorized "invisible" targets. The reverse paradigm however, pointing with an invisible cursor, has been widely studied (see, e.g., [6]) to better understand whether arm movements follow a position model or an amplitude model. It is interesting to note that our study seems to support the amplitude control model.

3 Motivation and Examples

In the rest of this article we call *Anim* the condition where the subject is pointing on an animated target and *Popup* the condition where the subject is pointing on a pop-up target. When the duration of the animation is null (DELAY = 0), the *Anim* condition is equivalent to the *Popup* condition: the target appears immediately when movement starts. This is different however from the acquisition of a persistent target, when the target is visible at its destination before the movement starts. This latter condition is the one usually considered in pointing experiments and will act as a control. We call it the *Static* condition.

One motivation for the design of the experiment was the following scenario. The user wants to drag-and-drop a hidden icon from the desk[1] to an area of the working window, e.g., to attach a file to an email. The user typically must trigger a command to make the desk visible, grab the icon and then make the original working window visible again to drop the icon at the desired destination. With Mac OS X, the user may use Exposé to make the desk visible: all windows are moved outside the screen with an animation and then animated back to their original position when the user starts the drag operation. This corresponds to the *Anim* condition because the target of the drop operation is an animated window. With Microsoft Windows and in most modern X Window environments the user may use the "show desktop" feature to immediately remove the windows and make the desk visible, start dragging and then show the windows again. This corresponds to the *Popup* condition with, ideally, a very short delay. A variant of this technique is *desk pop* [, 2007.

10]: the desk is moved to the foreground with a semi-transparent background, providing access to the icons while keeping the windows accessible. This case corresponds to the *Static* condition.

Dialog boxes are another examples of animated and pop-up targets with a potential *a priori* knowledge of the position of the targets by the user. Such windows often pop up at the same position and contain only a few buttons, e.g., "Ok" and "Cancel". Users often know the action to perform before the dialog box pops up, and we have observed [8] that they tend to anticipate the display of the dialog box by moving their mouse pointer toward the target button before it appears. A typical example is the "how to download" dialog box of the Firefox browser, which shows up when clicking a link to non-HTML content. Due to network delays, the box pops up with an unpredictable lag and yet users anticipate its appearance. Mac OS X features an example of animated target: some dialog boxes "slide" out of their parent window's titlebar and here too, users anticipate the location of the button they want to click.

Finally, an interesting category of pop-up targets is given by web navigation. Users often revisit the same web pages [22] and end up knowing their layout. They also often follow the same navigation paths [25]. For example, a user loads a page, clicks on the login button of a pre-filled form that loads a new page, and finally clicks on a link to navigate to the desired page. The delays involved in displaying the various pages and the targets they contain depend not only on the speed of the network connection but also on the browser's rendering algorithm and the structure of the page. Understanding the effect of delays on pointing may thus have some implications on browser and web page design and, more generally, GUIs.

4 Experiment

We conducted an experiment to study pop-up and animated targets by reducing the problem to a one-dimensional Fitts-like pointing task. Fitts' Law is inherently a 1D

[1] The background area of the screen containing icons.

model and considering 2D pointing would have involved additional factors [1] that we did not want to include in this first study.

4.1 Apparatus

The experiment was conducted on a 2.33 Ghz Core2-Duo Macbook Pro with an ATI Radeon X1600 graphics card connected to a 24 inch, 1920x1200 pixels LCD monitor. The experiment was implemented in Java with *SwingStates* [3]. We used a 1000 dpi LogiTECH RX 250 optical mouse with the default Mac OS X acceleration. The animations were linear translations computed at a rate of 60 steps by second.

4.2 Participants

12 unpaid adult volunteers (10 male and 2 female) participated in the experiment. All were right-handed and aged from 23 to 35 years old (average 27.6, median 26). All were experienced computer users familiar with mouse pointing.

4.3 Task and Procedure

A trial is a 1-D target acquisition task decomposed as follows. First, a green target is displayed on the screen (Figure 1.a) and the participant clicks on it when ready. The green target then disappears and the main target (in grey) appears for one second (Figure 1.b). This is the *memorization* phase where the participant is informed of the final position of the main target. After the main target disappears, a red target is displayed (Figure 1.c). In the *Static* condition the main target stays on screen while in the other two conditions the main target disappears. In all three conditions, the participant then has to click on the red target, which starts the recording of the movement time. In the *Static* condition, the main target remains static for the rest of the trial. In the *Anim* condition, the main target is smoothly animated to its final position in DELAY *ms*. In the *Popup* condition, the main target appears after DELAY *ms*. The trial ends when the participant successfully acquires the main target (Figure 1.d) after is has appeared at its final position.

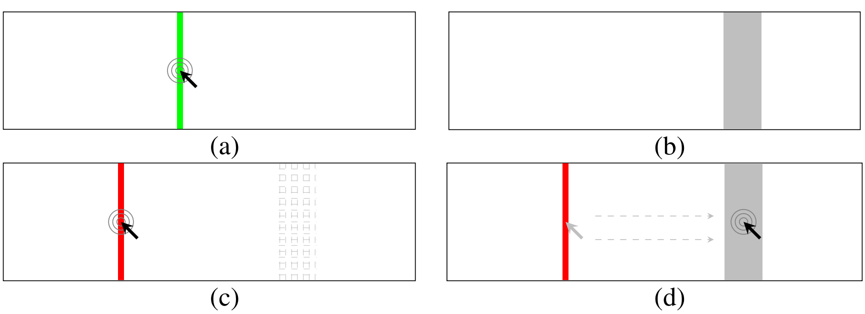


Fig. 1. (a) Start of trial: click green target; (b) memorization phase: show main target for one second; (c) target on screen (*Static*) or hidden (other conditions); click red target to start pointing; (d) end of trial: click main target (after animation or pop up according to condition)

The participants' movement time is recorded from the time the mouse button is released on the red target until the successful button press on the main target. If the participant clicks outside the target or before it reaches its final position, the trial is considered an error but continues until a successful click. Participants are instructed to perform the task as fast and as accurately as possible. To prevent participants from using the cursor to memorize the position of the target the mouse pointer is hidden throughout the memorization phase. Finally, in the *Anim* condition, the main target moves from the closest border of the screen to its final position in the opposite direction to the acquisition movement.

4.4 Design

The experiment is a partial 2×5×3×3 within-participant design with the following factors:

- 2 "techniques"conditions (TECH): *Anim* and *Popup*;
- 5 "delays" conditions (DELAY): *Static* and 0, 200, 350 and 500 milliseconds (*ms*) representing the duration of the animation or the delay for the pop-up;
- 3 widths (W) for the main target: 16, 32 and 64 pixels;
- 3 acquisition distances (D) to the main target: 256, 512 and 768 pixels.

The design is only partial (not fully factorial) because crossing TECH and DELAY leads to only 2+2×3=8 conditions since the DELAY conditions *Static* and 0 are the same for both TECH conditions: an animation of 0 *ms* makes the target pop up immediately. *Static* is not really a "delay" condition but it is a convenient way to consider it as both a control and an extreme condition (note that this transforms DELAY into a nominal factor, but we will later transform it into a continuous one).

We use 350*ms* as the median animation time because this duration is commonly used in graphical user interface animations, e.g., Mac OS X Exposé. We ran pilot studies in order to determine the range of reasonable animation and pop-up times, resulting in 200*ms* for the lower bound and 500*ms* for the upper bound.

We divide a run of the experiment into two parts. The first is made of two blocks, one with DELAY=*Static* and the other with DELAY=0. The second part of the experiment corresponds to the crossing of TECH (*Anim* and *Popup*) with the DELAY conditions >0, i.e., 2×3=6 blocks. We block by the TECH condition and use a 3×3 latin square to cross the 2 possible orders of TECH with the 3 non-zero DELAY conditions. This gives 6 counter-balanced orders that we cross with the first two blocks. We divide the 12 participants into two groups of 6. Both groups use these 6 orders, but the first group starts with DELAY=*Static* while the second group starts with DELAY=0.

Each of the 8 blocks described above has 7 replications of the 3×3=9 combinations of D by W, presented in a random order. The first replication is considered a warm up. A pause is offered to participants at the beginning of each replication.

To summarize, the total number of logged trials in the experiment is 8 blocks × 9 width-distance combinations × 6 replications × 12 participants = 5184 trials. We logged 72 trials for each full condition, 6 for each participant. The experiment lasted from 42 to 54 minutes (average 48, median 47).

5 Results

In this analysis, movement time MT is measured until a *successful* button press in the target (as opposed to the first button press). This has the advantage of accounting for penalties caused by errors. We duplicate data for *Static* and DELAY = 0 in order to simulate these conditions for both *Popup* and *Anim*. This allows us to perform a standard full-factorial repeated measures analysis of variance MT ~ TECH × DELAY × D × W × Random(Participant). Outliers[2], which represent 0.35% of the trials, are removed from all analyses[3].

Table 1. ANOVA for MT ~ TECH × DELAY × D × W × Random(Participant)

Factors	DF	DFDen	F	P
TECH	1	22	4.66	0.0539
DELAY	4	44	97.74	< 0.0001
D	2	22	68.63	< 0.0001
W	2	22	481.94	< 0.0001
TECH × DELAY	4	44	1.43	0.2396
TECH × D	2	22	0.19	0.8303
TECH × W	2	22	1.05	0.3653
DELAY × D	8	88	8.33	< 0.0001
DELAY × W	8	88	1.90	0.0686
D × W	4	44	2.67	0.0444
TECH × DELAY × D	8	88	0.19	0.9920
TECH × DELAY × W	8	88	0.35	0.9436
TECH × D × W	4	44	0.51	0.7274
DELAY × D × W	16	176	1.11	0.3469
TECH × DELAY × D × W	16	176	1.05	0.4071

Table 1 shows the results of the repeated analysis of variance for movement time. As expected D and W have a (strong) significant effect on movement time: the harder the task, the longer it takes to complete (Figure 2). TECH fails to reach significance for movement time and we find no significant interaction with the other factors. The difference in mean between *Anim* and *Popup* is only 17 *ms* in favor of *Popup*. This difference (see Figure 3) is less than 2% of the mean movement time[4].

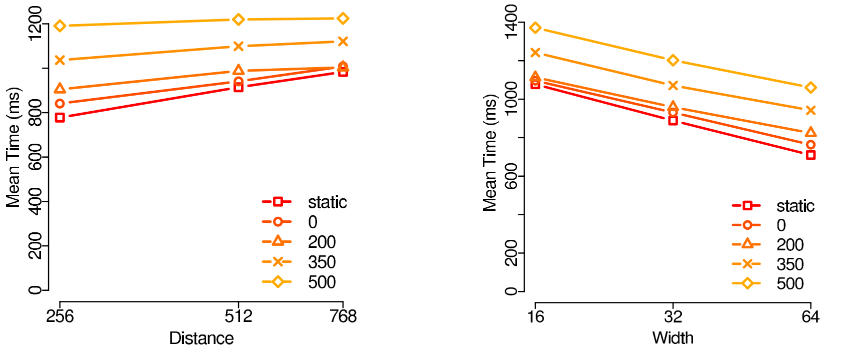
We observe a (strong) significant effect of DELAY on movement time (Figure 3). A Tukey post-hoc test ($\alpha = 0.05$) shows no significant difference in mean between *Static* and DELAY = 0 nor between DELAY = 0 and DELAY = 200. However, it shows a significant difference in mean between *Static* and DELAY = 200 (74 *ms* in favor of *Static*, a speed-up of 7.7%). It also shows that DELAY = 200 is significantly faster than DELAY = 350 (120 *ms*, 11.1% speed-up) and that DELAY = 350 is significantly faster than DELAY = 500 (126 *ms*, 10.4% speed-up).

We also observe a significant interaction effect between DELAY and D (Figure 2.a), suggesting that the effect of distance on movement time is less strong as

[2] Defined as a movement time 2 standard deviations a away from the mean movement time (for each PARTICIPANT, TECH, DELAY, D and W).

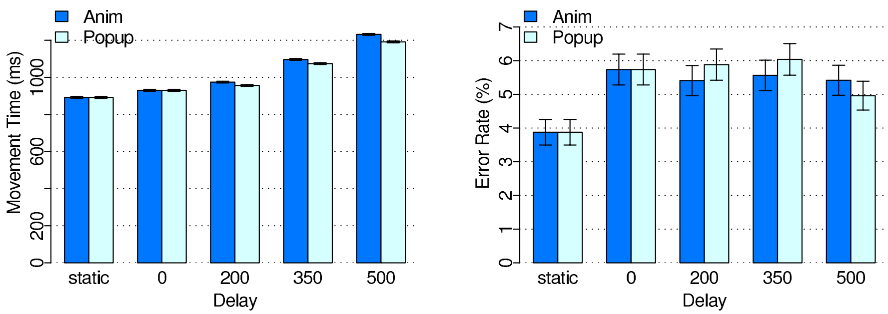
[3] Including the outliers yields results that are very similar to those described here.

[4] Note that the data duplication for *Static* and DELAY = 0 does not influence these statistical results since they cancel each other out.



(a) (b)

Fig. 2. (a) Movement time as a function of distance for each DELAY condition. (b) Movement time as a function of width for each DELAY condition.



(a) (b)

Fig. 3. Movement time (a) and Error rate (b) for each DELAY and TECH

DELAY increases from *Static* to DELAY = 0 to DELAY = 500. Indeed, a post-hoc Tukey test (α = 0.05) shows that there is a significant difference (in mean) between each distance for *Static* and DELAY = 0, a significant difference between distances 256 and 512 but not between distances 512 and 768 for DELAY = 200, a significant difference between distances 256 and 768 but not between distances 256 and 512 and distances 512 and 768 for DELAY = 350, and finally no significant difference at all for DELAY = 500.

Figure 2.a also suggests that the difference in movement time between *Static*, DELAY = 0 and DELAY = 200 decreases as distance increases. Indeed, a post-hoc Tukey test shows a significant difference in mean between *Static* and DELAY = 200 for D = 256 but no such difference for D = 768.

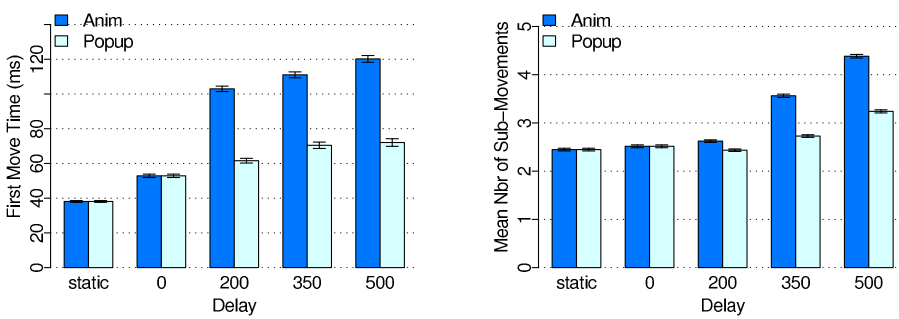
The above phenomenon can be explained by the fact that participants start their acquisition movement as soon as possible, i.e., before the end of the animation or before the target pops up, and that this ballistic movement is more precise when the target pops up or stops its animation earlier. For example, with DELAY = 500 and D = 256, participants can typically move the pointer close to the target before it pops up (or finishes its animation), but then have to wait before performing the final adjustment. On the other hand, with DELAY = 200 and D = 768, the target pops up (or finishes its animation) before the end of the ballistic movement and participants can adjust their movement as they go.

Regarding errors, we measured an error rate of 5.25%, a typical value for a pointing experiment. Figure 3.b shows the error rate as a function of DELAY for each TECH. Logistic two by two Pearson tests[5] show a significant difference ($\chi^2 = 4.880$, p=0.0272 for DELAY = 0) between *Static* (an error rate of 3.87%) and the other DELAY conditions (error rates between 5% and 6%.) There is no significant difference for errors among the remaining DELAY conditions. Other Pearson tests did not reveal any significant effect on errors for the other factors TECH, D and W.

6 Kinematic Analysis

A more detailed analysis of the pointing movements shows that participants effectively start their movement before the target pops up (when TECH = *Popup* and DELAY > 0) or before the target ends its animation (when TECH = *Anim* and DELAY > 0). Figure 4.a shows the time taken by the participants between the mouse button release (when they click on the red starting target) and the first mouse movement. These are clearly shorter than DELAY when DELAY > 0. Moreover, an ANOVA shows no significant difference between the DELAY conditions for TECH = *Popup*, nor between the non-zero DELAY conditions for TECH = *Anim*. However, as shown in Figure 4.a, the first mouse move times for TECH = *Anim* and DELAY > 0 are significantly longer than those for the other conditions (~40 *ms*).

An analysis of the kinematic record of the movements shows that Popup and Anim have slightly different profiles (see Figure 5 for an example, similar shapes arise with the other distances). The distance/speed curves for Popup are similar to those for the Static and DELAY = 0 conditions: they all are bell-shaped with a velocity peak at about 50% of the distance. The curves for *Anim* however are qualitatively different: they are right-skewed with a lower acceleration and a velocity peak at about 60% of the distance to target (these differences are statistically significant). Moreover, in the



(a) (b)

Fig. 4. (a) Time before first mouse move as a function of DELAY and TECH. (b) Mean number of sub-movements as a function of DELAY and TECH.

[5] Tests adapted to “discrete” measures.

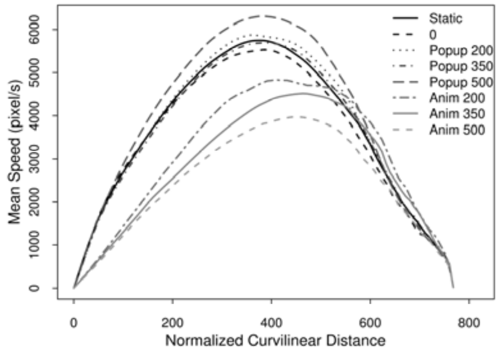


Fig. 5. Mean speed as a function of the (normalized curvilinear) distance for D = 768 for each DELAY × TECH condition (grey curves with a lower velocity peak are the *Anim* curves)

Anim condition we observe slower movements and in particular a lower velocity peak as DELAY increases. In the *Popup* condition we do not observe any such pattern.

A possible explanation of these phenomena is that in the *Anim* condition the participants try to follow the moving target or are distracted by it, leading to a slower movement. Conversely, in the *Popup* condition, participants move directly to the memorized position of the target, leading to a movement close to the classical target acquisition profile (note that in Figure 5.a the movement is even faster for DELAY = 500, but this is not the case for the other distances).

To confirm the above interpretation and to better understand the end of the movement we computed the number of sub-movements for all non-error trials. To do so, we count how many times the mouse does not move during at least 50 *ms* (to account for a null velocity) after the velocity peak and before the final mouse press. We add one to this number to account for the final movement, and interpret the result as a good estimate of the number of sub-movements. Figure 4.b shows the mean value of this number as a function of TECH and DELAY. An analysis of variance shows a significant effect of TECH and DELAY and an interaction between these two factors[6]. There are more sub-movements for *Anim* than for *Popup* and the difference (about one sub-movement) is significant for DELAY > 200.

In summary, we found that while *Popup* and *Anim* have similar performance, the pointing motion for *Popup* is closer to static pointing than to *Anim*. While *Popup* seems to follow a simple amplitude control model (jump to the anticipated position of the target and then adjust), *Anim* features a more complex movement.

7 Extending Fitts' Law

In order to better understand the effect of DELAY we used Fitts' Law to analyze movement time. Table 2 shows the Fitts' Law parameters and the adjusted r^2 for all data by DELAY. We averaged movement times across participants, repetitions and

[6] D and W also have significant effects – more sub-movements for smaller W and more sub-movements for D = 512 and 768 than for 256 – but there is no significant interaction effect with TECH and DELAY.

TECH since we have shown that this factor did not have a significant effect on MT. In other words, we take the mean of all trials for each DELAY, D and W conditions. As we expected, the fit for the complete dataset is not good because of a strong effect of DELAY (Figure 6.a). The fit for each DELAY condition is good for the *Static* condition (traditional Fitts' pointing) but it degrades as DELAY increases because of the unusual effect of distance on movement time. The fit for each DELAY condition can be improved by using Welford's model [27] $MT = a + b.\log_2(D) + b.\log_2(W)$ (r^2 = 0.95), but this model fails to properly fit the complete dataset (r^2 = 0.545).

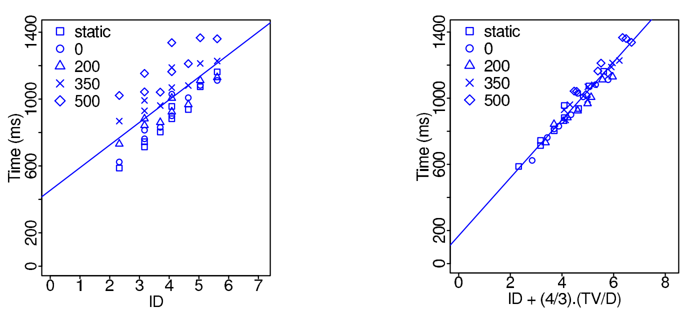
In order to compute regression models that include DELAY, we create a continuous factor TV as follows: We map *Static* to 0 *ms*; We map DELAY = 0 to 100 *ms* since this is the delay needed for sensory information to affect the physical movement [17]; Finally we map the other values of DELAY to themselves.

We can now consider the simple model $MT = a + b.ID + c.TV$, which dramatically improves the fit: $MT = 307 + 135.ID + 0.64.TV$, adjusted r^2 = 0.8921. Using Welford's model instead: $MT = a + b.\log_2(D) + c.\log_2(W) + d.TV$ improves the fit: $MT = 977 + 74.\log_2(D) - 158.\log_2(W) + 0.64.TV$, adjusted r^2 = 0.9478. This can be further improved to r^2 = 0.9665 by adding yet another term, $\log_2(D) * TV$, to account for the D by DELAY interaction observed in the previous section.

However good the fit, none of these models are particularly intuitive. They simply assume a linear effect of DELAY and an unusual interaction of DELAY with Distance. The problem with this approach however is that the resulting model has four free variables (*a*, *b*, *c*, *d*) while we have only three main factors (D, W, DELAY). If we are to be consistent with Fitts' Law, we should have at most three variables.

Table 2. Fitts' Law parameters *MT=a+b.ID*

DELAY	a	b	adj. r^2	# pts.
ALL	454	135	0.4998	45
Static	180	174	0.9656	9
0	299	153	0.9074	9
200	462	120	0.8973	9
350	599	116	0.7974	9
500	731	115	0.6654	9



(a) (b)

Fig. 6. Regression with Fitts' Law (a) and with the modified *ID* including a $\frac{TV}{D}$ term (b)

We have seen in the previous section that movement time increases with TV but that this depends on distance: as distance increases, the degradation on movement time for a given value of TV decreases. Thus, the combined effect of TV and D may be captured by the ratio TV / D . We therefore consider the following model:

$$MT = a + b.ID + c\frac{TV}{D}$$

The regression with this model yields $MT = 171 + 174.ID + 237.TV / D$ and a good adjusted r^2 = 09414. Taking advantage of the fact that $237/174 \approx 4/3$, we use the new index of difficulty $ID + 4/3.TV / D$ to plot the resulting data in Figure 6.b.

Of course, the validity of this model should be further tested and may be altered by considering larger sets of distances, target widths and DELAY conditions. Note that an important property of our experimental design is that the values for DELAY are shorter than the expected movement times in the *Static* condition. If TV is larger than the pointing time in the static condition, a different model is likely to apply since the user would have to move the pointer towards the memorized position of the target, wait until the target pops up (after *TV ms*), and finally acquire the target of width *W* at a distance *err*(*D*) where *err* is a function modeling the distance error when pointing to a memorized invisible target at a distance *D*. A candidate model for this situation could be $MT = TV + a + b.\log_2(err(D)/W + 1)$, but this remains to be tested.

8 Discussion and Future Directions

The work presented in this paper is the first to study the acquisition of pop-up and animated targets. We did not find significant differences in performance between the acquisition of a static target and the acquisition of a “memorized” target that pops up immediately. However, we found a significant difference regarding errors in favour of static pointing and we also found that static pointing is significantly faster than pointing on animated and pop-up targets with a delay longer than 200 *ms*. This suggests that delays and animations can indeed impair performance, and that techniques that keep the context (and the target) visible should be preferred.

We did not find significant performance differences between the acquisition of targets that pop up immediately and animated or pop-up targets with a delay of 200 *ms*. However, we observed large differences when delays increase to 350 and 500 *ms*, the magnitude of the difference being comparable to the increase in delay. This suggests that animations should be kept close to a duration of 200 *ms* whenever the acquisition of a target inside the animated object is desirable. Also, GUIs should do their best to keep the delay for popping up a potential target under 200 *ms*. Note however that other design factors may come into play in the real world that are more important than pointing performance, such as the ability to perceive causality with animations.

Our study suggests that users are able to adapt their pointing movement dynamically. It is interesting to note that this adaptation depends on the feedback given to the user: animation affects the nature of the movement, slowing it down and leading to more sub-movements than pop-up. Since we did not observe a significant difference in performance between the acquisition of an animated target and a pop-up target (under the same delay) it is possible that users are able to take advantage of the target animation during the acquisition movement to better predict the final position of the target. We intend to pursue this hypothesis in future work.

Another area for future work is to extend the scope of the study. In this initial study, we conducted a "pure" experiment idealizing the real world. Now that we have a better understanding of the basic factors involved, we can start to examine more realistic situations, such as 2D pointing with real windows and icons. We also want to consider more complex tasks that involve pop-up/animated targets such as those described in section 3. Finally we need to study the effects of other factors such as the animation type, e.g., slow-in/slow-out, and the degree of position memorization. For pop-up targets, we also plan to investigate larger delays and the extreme case of memorized invisible targets in order to refine our movement time model.

Acknowledgements

We wish to thank the member of the InSitu lab and in particular Stéphane Huot, as well as our experiment participants and the anonymous reviewers for their feedback.

References

1. Accot, J., Zhai, S.: Refining Fitts' law models for bivariate pointing. In: Proc. Human Factors in Computing Systems (CHI 2003), pp. 193–200. ACM Press, New York (2003)
2. Aivar, M., Hayhoe, M., Chizk, C., Mruczek, R.: Spatial memory and saccadic targeting in a natural task. Journal of Vision 5(3:3), 177–193 (2005)
3. Appert, C., Beaudouin-Lafon, M.: SwingStates: Adding state machines to Java and the Swing toolkit. Software: Practice and Experience 38(11), 1149–1182 (2008)
4. Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., Zierlinger, A.: Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch- and pen-operated systems. In: Proc. INTERACT 2003, pp. 57–64. IOS Press, IFIP (2003)
5. Bezerianos, A., Balakrishnan, R.: The vacuum: facilitating the manipulation of distant objects. In: Proc. Human Factors in Computing Systems (CHI 2005), pp. 361–370. ACM Press, New York (2005)
6. Bock, O., Eckmiller, R.: Goal-directed arm movements in absence of visual guidance: evidence for amplitude rather than position control. Exp. Brain Res. 62(3), 451–458 (1986)
7. Cao, X., Li, J.J., Balakrishnan, R.: Peephole pointing: Modeling acquisition of dynamically revealed targets. In: Proc. Human Factors in Computing Systems (CHI 2008), pp. 1699–1709. ACM Press, New York (2008)
8. Chapuis, O.: Gestion des fenêtres: enregistrement et visualisation de l'interaction. In: Proc. Journées Francophones d'Interaction Homme-Machine (IHM 2005), pp. 255–258. ACM, New York (2005)
9. Cockburn, A., Gutwin, C., Greenberg, S.: A predictive model of menu performance. In: Proc. Human Factors in Computing Systems (CHI 2007), pp. 627–636. ACM Press, New York (2007)

10. Faure, G., Chapuis, O., Roussel, N.: Power tools for copying and moving: Useful stuff for your desktop. In: Proc. Human Factors in Computing Systems (CHI 2009), pp. 1675–1678. ACM Press, New York (2009)
11. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47, 381–391 (1954)
12. Flash, T., Henis, E.: Arm trajectory modifications during reaching towards visual targets. J. Cognitive Neuroscience 3(3), 220–230 (1991)
13. Gonzalez, C.: Does animation in user interfaces improve decision making? In: Proc. Human Factors in Computing Systems (CHI 1996), pp. 27–34. ACM Press, New York (1996)
14. Hinckley, K., Guimbretiere, F., Baudisch, P., Sarin, R., Agrawala, M., Cutrell, E.: The springboard: multiple modes in one spring-loaded control. In: Proc. Human Factors in Computing Systems (CHI 2006), pp. 181–190. ACM Press, New York (2006)
15. Hoffmann, E.R.: Capture of moving targets: a modification of Fitts' law. Ergonomics 34, 211–220 (1991)
16. Hornof, A.J., Kieras, D.E.: Cognitive modeling demonstrates how people use anticipated location knowledge of menu items. In: Proc. Human Factors in Computing Systems (CHI 1999), pp. 410–417. ACM Press, New York (1999)
17. Jeannerod, M.: The neural and behavioural organization of goal directed movements. Clarendon Press, Oxford (1988)
18. Kurtenbach, G., Buxton, W.: Issues in combining marking and direct manipulation techniques. In: Proc. User Interface Software and Technology (UIST 1991), pp. 137–144. ACM Press, New York (1991)
19. MacKenzie, I.S.: Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction 7, 91–139 (1992)
20. Meyer, D., Smith, J., Kornblum, S., Abrams, R., Wright, C.: Optimality in human motor performance: Ideal control of rapid aimed movements. Psych. Review 95, 340–370 (1988)
21. Mould, D., Gutwin, C.: The effects of feedback on targeting with multiple moving targets. In: Proc. Graphics Interface (GI 2004), Canadian Hum.-Comp. Comm. Soc., pp. 25–32 (2004)
22. Obendorf, H., Weinreich, H., Herder, E., Mayer, M.: Web page revisitation revisited: Implications of a long-term click-stream study of browser usage. In: Proc. Human Factors in Computing Systems (CHI 2007), pp. 597–606. ACM, New York (2007)
23. Plamondon, R., Alimi, A.: Speed/accuracy trade-offs in target-directed movements. Behavioral and Brain Sciences 20(2), 279–349 (1997)
24. Schlienger, C., Conversy, S., Chatty, S., Anquetil, M., Mertz, C.: Improving users' comprehension of changes with animation and sound: An empirical assessment. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4662, pp. 207–220. Springer, Heidelberg (2007)
25. Tabard, A., Mackay, W., Roussel, N., Letondal, C.: Pagelinker: integrating contextual bookmarks within a browser. In: Proc. Human Factors in Computing Systems (CHI 2007), pp. 337–346. ACM Press, New York (2007)
26. Thomas, B.H., Calder, P.: Animating direct manipulation interfaces. In: Proc. User Interface Software and Technology (UIST 1995), pp. 3–12. ACM Press, New York (1995)
27. Welford, A.T., Norris, A.H., Shock, N.W.: Speed and accuracy of movement and their changes with age. Acta. Psychologica 30, 3–15 (1969)
28. Zhai, S., Conversy, S., Beaudouin-Lafon, M., Guiard, Y.: Human on-line response to target expansion. In: Proc. Human Factors in Computing Systems (CHI 2003), pp. 177–184. ACM, New York (2003)

An Optical Pen Tracking System as Alternative Pointing Device

Ingmar Seeliger, Ulrich Schwanecke, and Peter Barth

University of Applied Sciences Wiesbaden,
Unter den Eichen 5, 65195 Wiesbaden

Abstract. A webcam together with a pen can replace a mouse as pointing device for many common user interaction tasks. We have implemented an image-processing component integrated in a tool that acts as mouse alternative. The image-processing component tracks the head of a pen based on shape and colour information retrieved in a quick, integrated initial pen-calibration phase using Hough transform triggered by a motion detection cycle. The tracked 2D position of the pen-head seen by the webcam is used to smoothly position the mouse cursor. Combined with auto-clicking we can replace mouse-based user interaction. The system tolerates changing lighting conditions, does not need time-consuming camera calibration and works with off-the-shelf webcams. First user experiences show that this technology can partially replace mouse interaction for Repetitive Strain Injury (RSI) patients as well as completely replace mouse interaction within dedicated environments such as presentation booths or simple games.

Keywords: Image processing, human computer interaction, mouse alternative, Repetitive Strain Injury (RSI).

1 Introduction

Repetitive Strain Injury (RSI) [18,5] is probably the most common occupational disease inflicted by human computer interaction [9]. Damages and injuries are mostly located in the musculoskeletal system of hand, wrist, arm, elbow, shoulder and neck resulting from thousands of repetitive motions mainly with keyboard and mouse. While ergonomically formed keyboards, trackballs, vertical mice, etc. are at least partially suitable for preventing RSI they cannot annihilate the repetitive nature of the typical work on the computer.

A promising approach is variation, which means switching among input devices while completing the daily work. While typing can often be replaced with speech recognition there are only few viable and cost-effective alternatives available that replace the mouse or mouse like devices as pointing device. We propose a simple to setup and intuitive to use alternative pointing device requiring only a classical coloured pen and a webcam as additional hardware. By moving the pen in front of the camera the user positions the mouse cursor. We concentrate on image processing capabilities that avoid extensive and complicated camera calibration while still providing a decent recognition in most environments.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 386–399, 2009.

This paper is structured as follows. First, we differentiate from related work in section 2. In section 3 we provide the overall architecture and design of the pen tracking system based on the environmental assumptions and technical requirements. The initial pen calibration phase extracting colour and shape information is detailed in section 4, while actual pen tracking is handled in section 5. Using an alternative pen tracking phase based on particle filters is first motivated and then realized in section 6. We evaluate the suitability of the pen tracking system as mouse alternative in section 7 with a small group of diverse test persons and conclude in section 8.

2 Camera Based Tracking

The main task in order to replace a classical pointing device lies in identifying and tracking a position of a body part such as hand, finger, head, or eye [3]. Disregarding special hardware solutions most research concentrates on processing a succession of camera images to identify and track a position.

An example for hand tracking is the Visual Touchpad [13]. The input is extracted from two cameras. The user's hands move parallel to a horizontal black surface. This multitouch-interface enables for example drag and drop, selection, zooming and operating a virtual keyboard. Finger tracking on large displays is done e.g. by Visual Touchscreen [4]. Here, the cursor is controlled by pointing movements with the finger. The system uses one camera and identifies the fingertip and the central point between the eyes, so that a line through fingertip and central point defines the destination on the display. Another finger tracking system is FingerMouse [6], a stereo camera system for wearable computing. The depth computed from the images of the two cameras helps to identify fingers as the nearest objects and then finger movement can be used for controlling a cursor. For eye tracking typical setups need two cameras to reconstruct the pupil in 3D space [11] while there are also approaches with one camera [7] allowing free head motion. An advantage of both systems is the absence of a complicated camera calibration. Beside camera based tracking there are several systems that require special equipment in order to track positions, among them are multimodal interfaces [12], the wiimote [15], and step-user interfaces [14].

Our aim is to control the cursor by tracking the position of a coloured pen with a simple webcam. In contrast to finger tracking using a pen has two advantages. Firstly, tracking is much more reliable because pens have a regular shape and interference with facial skin colour is avoided. Secondly, user interaction is intuitive because explicitly grabbing an input device reflects grabbing a mouse.

3 Requirements and Design

The pen tracking system ought to be a mouse alternative for daily routine work, like operating office, email, and browser applications. Based on this premise we first define the requirements and identify potential factors influencing tracking the pen position and then describe the overall architecture of our pen tracking system.

3.1 Requirements

On the one hand, the system's resolution has to be high enough to accurately position the cursor on the control elements such as menu items and links. The replacement of the mouse has to be comprehensive and on the window system level as software will not be rewritten to take advantage of our solution.

We do not require any special surrounding but allow usage by different persons at any place and time. Thus, we require an independence of background, brightness and colour of clothes. Furthermore, users of all ages typically tremble while holding a pen up in the air. The pen tracking system has to flexibly compensate the trembling according to the individual behaviour of different users.

The acceptance of such a solution depends heavily on an easy setup. Therefore, we must refrain from a complicated camera calibration phase. The system should be cost-effective and only require a simple webcam. Therefore, we need independence of the camera resolution and must provide a working environment starting with a resolution of 320x240 pixels.

3.2 Camera and Lighting

Web *cameras* differ in many parameters influencing the tracking results. The resolution determines the maximum precision of the pointing device and the focal distance defines the optimal distance of pen and camera. Images are distorted because of the bending of optical lenses and the sensitivity of the CCD-chip (Charge Coupled Device), which is an image sensor that converts an optical image to an electric signal, determines brightness and noise in the images. Figure 1 shows some of these effects by means of three different webcams with the same position and lighting.

However, the biggest influence on the images is caused by different *lighting conditions*. We may operate in an environment with artificial or natural light and have direct and diffuse light that influence the image. Direct light can lead to blasting, causing loss of the important colour information. In addition to the light sources in the room, the computer monitor directly lights the user and often changes its colour depending on the background. Figure 2 shows the influences of different lightings.

The effect of different *positions* of the camera with respect to the main light source is significant, as shown in figure 3. Some of the positions make it nearly impossible to

Fig. 1. Different cameras with the same position and lighting, (1) Unibrain Fire-i Digital Camera (Firewire), (2) Phillips FunCam (USB), (3) LogiTech QuickCam Zoom (USB)

Fig. 2. Different lighting on the same object (1) Natural light in the afternoon, (2) Natural light in the evening (3) Artificial light from the ceiling, bright monitor light

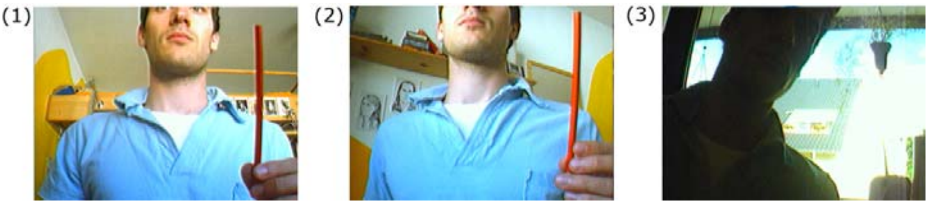


Fig. 3. Different positions of the camera regarding the main light, (1) Camera points in the opposite direction, (2) Lateral position to the light, (3) Camera points to the main light

detect a pen in the image. Ideally, the camera points in the opposite direction of the light source such as the room window. The user lighting is frontal and the illumination of the room is constant. As long as the light source is not from the back, the lighting is acceptable.

3.3 Design

The overall system design comprises of two parts, image processing providing pen positions and integration of the tracking system into a GUI toolkit. Image processing is further refined. One part covers the initial calibrating phase, where parameters are extracted that allows identifying the pen in an image. The other part covers using the previously extracted parameters to provide a continuous stream of valid pen positions. Both parts merit a separate section, section 4 and section 5 respectively.

For the integration of the pen tracking system into a GUI toolkit, we decided to provide a platform dependent (Windows) stand alone application, that positions the mouse cursor and emits click events through platform specific APIs. Mouse drivers typically move the cursor farther on faster mouse movement. Note that we cannot mimic such a behaviour as there is no analog to "lift the mouse and set to a new position" with a pen. Therefore, we assign absolute screen positions to the positions of the pen in the camera image. To simulate clicking, we implement a simple "auto-clicker", that simulates a single left click whenever the cursor comes to a halt for a while after a movement. Other clicks and scrolling is initiated when special control keys are pressed on the keyboard simultaneously to the automated click.

In the following, we focus on the image processing features to extract valid positions and disregard implementation details regarding the window level integration of the pen tracking system.

4 Pen Calibration

In order to track the position of the pen, we first need to extract enough characteristic information to identify the pen in an image. Therefore, we extract in an initial calibration phase, where the user moves the pen once horizontal from on edge of the image to the other, four main characteristics – oblongness, movement, linear edges and monochromaticity – that serve later to identify the pen.

By *oblongness* we mean a measurement to compare different objects when more than one object is segmented as possible pen. In such cases it can serve as decision support because the pen is a very longish object.

The *movement* of the pen can be detected with help of difference-image-methods. Note, that differences with regard to a reference background are not sufficient, as the user typically moves, and if only a little, behind the pen. Thus, we compute difference images I_t based on three consecutive images G_{t-1}, G_t, G_{t+1} with the sensitivity parameter s as follows.

$$I_t(x,y) = \begin{cases} 1 & \mathit{if} \left|G_{t+1}(x,y) - G_t(x,y)\right| > s \wedge \left|G_t(x,y) - G_{t-1}(x,y)\right| > s \\ 0 & \mathit{otherwise} \end{cases} \tag{1}$$

Examples are shown in figure 4. Note that this method alone is not sufficient to reliable identify pen movement, but is used for a first approximation.

Linear edges are reliably found with Hough transform [8] after creating an edge image with a proper filter such as Laplace, Sobel, or Canny. Again, we might find not only the edges of the pen but also other edges in the image. Furthermore, the edges are not necessarily straight lines due to the bending of the camera's optical lens.

The *monochromaticity* and therefore the colour of the pen is a very important and practical information for segmentation and is the most significant characteristic feature used later in the tracking phase. The pen should be chosen in a way that it differentiates strongly from the other colours in the background. As the background and lighting differs, different pens might be better suited in different circumstances. Thus, the pen can be freely chosen.

Note, that the free choice of the pen and the different lightings make a calibration step before every use mandatory. In this calibration step the pen must be identified in the image and its main characteristic feature, the colour, adequately identified.

Fig. 4. Analysis of movement with difference-images, (1) With background as reference-image, (2) With three following images, movement of the body, (3) With three following images, only movement of the arm

4.1 The Calibration Process

The result of the calibration is a colour class with linear borders for the actual pen, which serves as basis for later tracking. We have chosen sense-orientated HSV colour space among the colour spaces YCrCb and YUV most common in analog video, the device-dependent RGB and the chromatic RG colour space. HSV was the most compact representation of some experimentally created pen colour classes tolerating the most error pixels. As calibration has to be done frequently, we require only a simple calibration movement. The pen is moved vertically from right to left through the image within two seconds. During this movement, the colour class is calculated.

The whole process is depicted in figure 5. First, to reduce noise, images are preprocessed with a Gaussian filter. Then, actual movement is detected with a difference- image-method based on three images. The result is the rough shape of the pen plus disturbance areas such as the moving hand or head. This is followed by a line search based on a Hough transform on the edges of the difference-images. Not only the lines are calculated, but also the line segments of the detected edges. Edges with a high deviation from the vertical are removed because of the vertical pen movement during calibration. From the remaining edge lines, in most cases just the two pen edges remain, the longest edge is interpreted as pen. The median x-value of this edge is used as actual horizontal x position of the pen. We assume to have started the calibration, if the detected edge is in the right third of the image, i.e. the pen movement has just started. If the movement does not reach the left side the calibration is cancelled and restarted. In case of a successful calibration we compute the colour class based on the difference-images and the Gaussian filtered images. To this end, we use a fast region labelling with an union find algorithm [2] that removes all but the biggest regions. Only the pen and – maybe connected with it – some parts of the hand remain. These images are now used for masking the Gaussian filtered images, so that the pen is successfully segmented during movement. Finally, the parts of the hand operating the pen, which consist of all pixels below the line segments, are removed.

Figure 6 (1) shows a collage of the segmented pens during the movement. This is the basis for the creation of the colour class. First, the image is converted in HSV colour space and second, a linear colour class represented by a rectangular bounding-box is calculated which contains 90% of all non-black pixels of the image. The remaining 10% are treated as distorted noise-pixel, which result from discoloration caused by reflections especially at the border of the pen as shown in figure 6 (1). The ranges of the bounding-box in the dimensions of saturation and brightness depend on the brightness and lighting in the room. The range of hue is a static interval because of the monochromaticity of the pen, only translated depending on its actual colour. An example colour class is shown in figure 6 (2).

We expand the colour class to account for the discontinuous and fast changing saturation and brightness values of the image of the pen. The expansion is limited by the first erroneous background pixel falling into the colour class. This is based on the Gaussian filtered images and the fitting inverted masks.

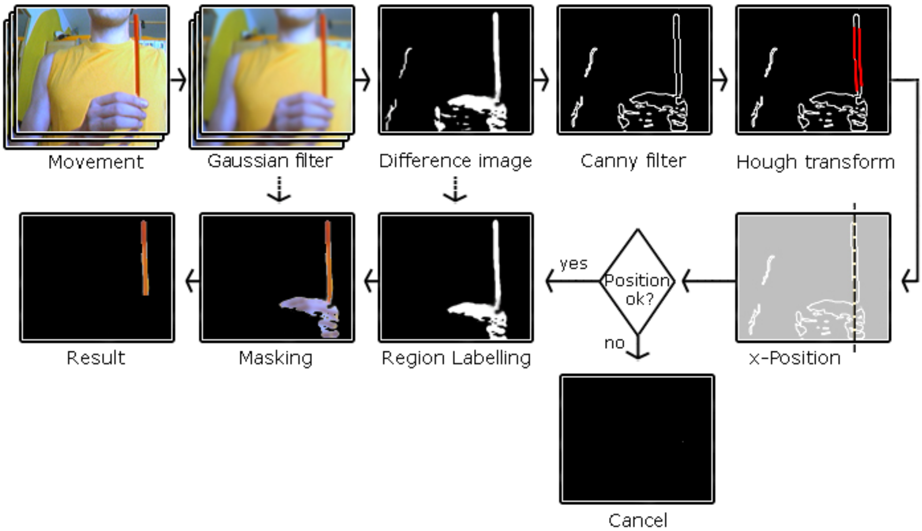


Fig. 5. Process of segmentation for pen calibration

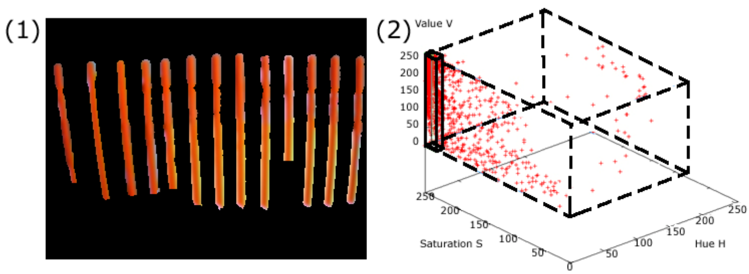


Fig. 6. Calibration results, (1) Segmented pen during movement, (2) Colour classes of the segmented pen. Dashed box: Class with all pixel. Solid box: Class without 10% noise-pixel

5 Pen Tracking

The main task while tracking the position of the pen is the *segmentation of the pen* in the stream of images from the webcam. In order to achieve fluid interaction this step needs to be fast.

The segmentation process consists of several steps depicted in figure 7. The first step is the reduction of noise with a Gauss-Filter and conversion into HSV colour space. Thereafter, the colour is segmented with the help of a fast colour class test [2]. The result is a binary image where the areas with the approximate colour of the pen are white and all others black followed by region labelling.

Next, we determine which of these regions the searched pen is by assigning probabilities of a hit to these regions and select the one with the highest probability. Note that the region with the biggest surface area is not necessarily the pen. Bigger parts of the pen are commonly not found because of light reflections which reduce the surface area of the segmentation. Therefore, we use the area only as secondary characteristic to eliminate single error pixel and small areas. The oblongness of the pen serves as primary characteristic. It can be calculated with the help of moment-based shape features [10]. Especially the eccentricity

$$\varepsilon = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{(\mu_{20} - \mu_{02})^2} \qquad (2)$$

where μ_{20}, μ_{02} and μ_{11} are the translation-invariant centric moments of second order is a robust measure of the roundness of an object. Thereby ε varies between 0 for a perfect circle and 1 for a straight line.

Finally, we use the shape to identify the pen being the most longish object.

Note that the pen is frequently put down while typing. In this case other longish segmented objects may be chosen as pen by mistake. For this reason, it is necessary to factor the actual movements of the pen into the recognition so that static background objects are not erroneously identified as pen. To this end, the difference-image is calculated and compared with the segmented pen. When a movement is discovered in this area, the found pen is valid and the coordinates of the pen are used for position the cursor. In any other case we conjecture to have either a static background object and assume that the pen is outside of the camera area, or we have the pen but it did not move sufficiently far or fast to change the cursor position. In case the distance of the actual coordinates to the last ones is small, the segmented area is in fact the pen.

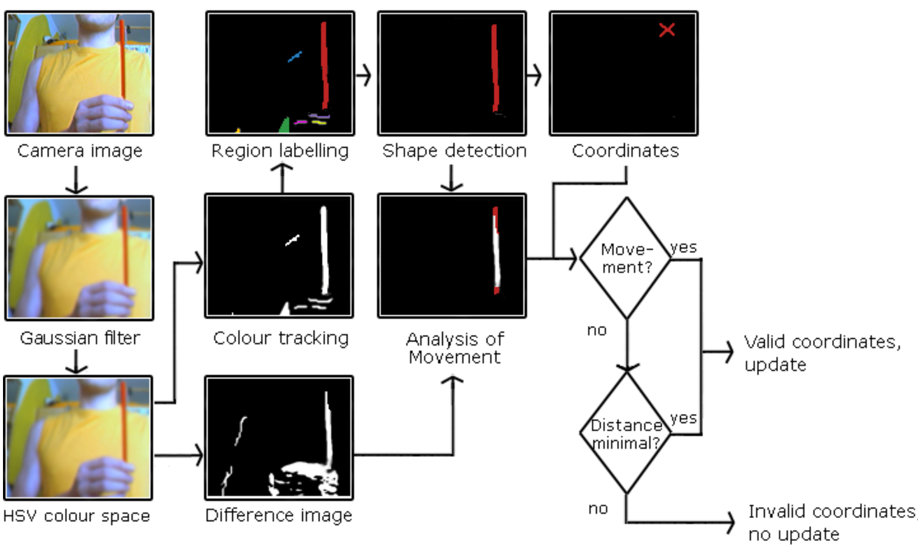


Fig. 7. The pen-tracking process

5.1 Border Areas

While moving the pen in front of the camera, it is inevitable that the pen is clipped at the border areas. The top of the pen will disappear at the side areas and the visible part becomes shorter, leading to distorted coordinates. The lower border area is most problematic, because the pen disappears more and more from the image the further down the pen is positioned. To accommodate for this fact, a smaller rectangle is defined inside the image as buffer for the border error. The whole image is used for tracking, but the calculated coordinates are translated into a rectangle yielding valid coordinates. Thus, we solve the problem of the border areas but decrease the available resolution.

5.2 Cursor Control

The valid coordinates identified from the images are projected to the actual monitor resolution. The frame rate can be too low to use the exact coordinates as control for the cursor. Thus, it is necessary to interpolate between the coordinates to achieve a fluid movement. Additionally, a prediction filter was implemented to compensate the trembling of the hand and filter out single defective measurements. The filter predicts the next position of the pen, compares it with the actual calculated values and smoothes them. In this way, a fluid and soft cursor movement was achieved.

6 Tracking with Particle Filters

The tracking introduced in the last section does not provide accurate results if we have background areas whose colour class fall into the colour class of the pen as shown in figure 8. In that case a connected area is identified during segmentation, consisting of pen and background area. This changes either the oblongness value of the segmented object so that it is not identified as pen or stretches the pen so that it is identified, but the coordinates of the top are distorted. Note that these problems can be minimized with a better distinguishable pen and camera repositioning. Still, there are frequent recognition errors because of light reflections on the surface of the pen. In the centre of a reflection, the brightness is 100%, the saturation is nearly zero and the hue value varies greatly due to the blasting of the camera. In most cases these reflections result only in holes in the segmented pen, which has no consequence on pen-identification and calculation of the coordinates, because the oblongness and top of the pen are not affected. This changes, if the reflection is at the top of the pen or is so

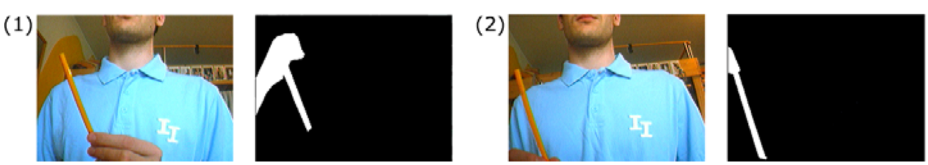


Fig. 8. Covering effects with distorted segmentation

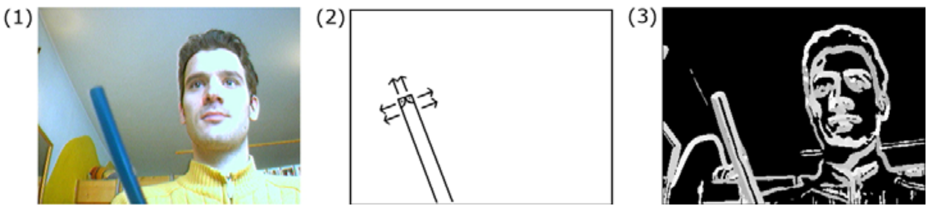


Fig. 9. Shape properties of pen-head, (1) Original image, (2) Simplified sketch with gradients, (3) Gradient-image after use of a Sobel filter. Grey values represent the angles of the gradients.

strong that it divides the segmented pen into two parts. In this case we have distortions that have to be handled.

Thus, we have implemented a *particle filter* as alternative approach to track the pen position. For the particle filter we need to identify the properties to weight the particles. Note that the oblongness of the pen, which is one of the most important properties of the last approach, cannot be used in a particle filter, because only the local environment of the particles is regarded. Therefore we use in addition to the colour the shape of the pen's top as primary property. This is the end of the pen, that is not held in the user's hand as shown in figure 9 (1), consisting of three borders which are orthogonal to each other (2).

Two of the gradients point in opposite directions, whereas the third is at a right angle to both of other two. To find this special arrangement with image-processing methods, a gradient-image is calculated with a Sobel filter as shown in figure 9 (3).

We search for the orthogonal gradients in the environment of the particles. Therefore, a gradient histogram p which counts the relative frequency of the different gradient angles is generated for each particle. With the help of the Bhattacharyya-coefficient [1]

$$BC(p,q) = \sum_{x \in \{0,5,\ldots,360°\}} \sqrt{p(x)q(x)} \in [0,1] \tag{3}$$

these histograms are compared with a static reference histogram q resulting from a model of a pen top. Note that it is sufficient to change the reference histogram in order to recognize different shapes of pen tops.

The result is a measurement for the similarity of the compared histograms. As the value is not rotational invariant, every particle is equipped with a parameter for rotation in addition to its position. To compute rotational invariant histograms, each particle together with the associated gradients is rotated to the same vertical direction before calculating the histograms.

Among the shape of the pen's top, the colour is integrated in the weight of the particles to exclude all other areas of the image but the pen. Figure 10 shows the particle filter in action.

One main advantage of particle filters over other approaches is that instead of yes/no decisions after each working-cycle probabilities that include the measurements and results of former cycles are computed. While particle filters can improve the

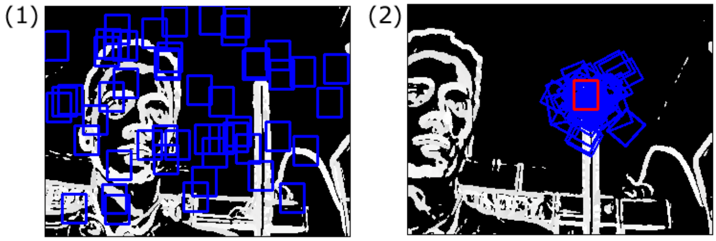


Fig. 10. Tracking with the particle filter, (1) Initial state of the particles, (2) The particles follow the pen-head

recognition rate they cannot completely compensate for the errors caused by similar colours and reflection. Thus, both segmentation based tracking and particle filter based tracking constitute a flexible and reliable pen tracking system.

7 Evaluation

A usability test with a small group of people was conducted focusing on the applicability of the pen tracking system as alternative pointing device for RSI affected users [16]. The group consisted of 15 test persons aged 15-55 where 9 persons have no discomfort during computer work, 4 persons with some discomfort and 2 persons with obvious RSI-symptoms as well as acute pain. They had to perform everyday tasks such as Internet research, data-handling and operating office software. The results are based on a questionnaire completed after the test by the users where the questions were graded from 1 (excellent) to 6 (unsatisfactory). Main focus was the perceived user satisfaction regarding pen-calibration, cursor movement, changing between pen and keyboard, aiming accuracy as well as learnability of the pen tracking system. The results are shown in figure 11.

In most cases, the users were satisfied with the calibration movement and also the movement of the cursor satisfied the expectations. However, the latency between user input and visible result was perceived as too high (”*The cursor operation is too sluggish.*”).

The change between pen and keyboard did not pose any problems. One of the most criticized points was aiming accuracy. Opinions varied between ”good” (2) and ”poor” (5). The learnability of the system was classified as easy to moderate. Pointing with a pen seems to be a very intuitive gesture.

Further questions concentrate on ergonomic effects. The left diagram in figure 12 shows the result of the questions regarding strain perceived in shoulder, elbow and wrist and the right diagram shows how the strains compare to the strains while using a typical mouse. Highest strains were felt in the shoulder and the elbow. Thus, our system cannot be recommended as long term full replacement of a mouse. The range of the answers concerning strains was extremely diverse and does not allow to draw statistically significant conclusions. However, 54% of the test persons stated that in comparison to typical mice, the strain in the body seems to spread differently using

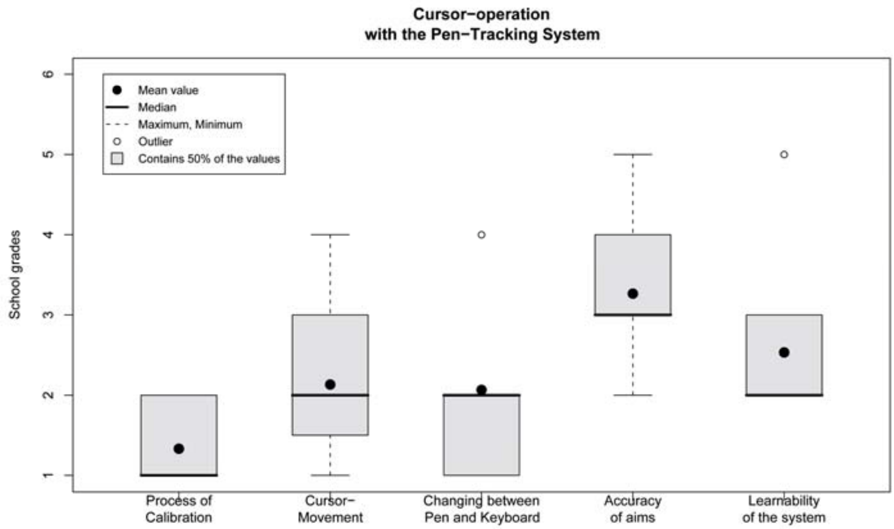


Fig. 11. Summary of cursor-operation with the pen-tracking system

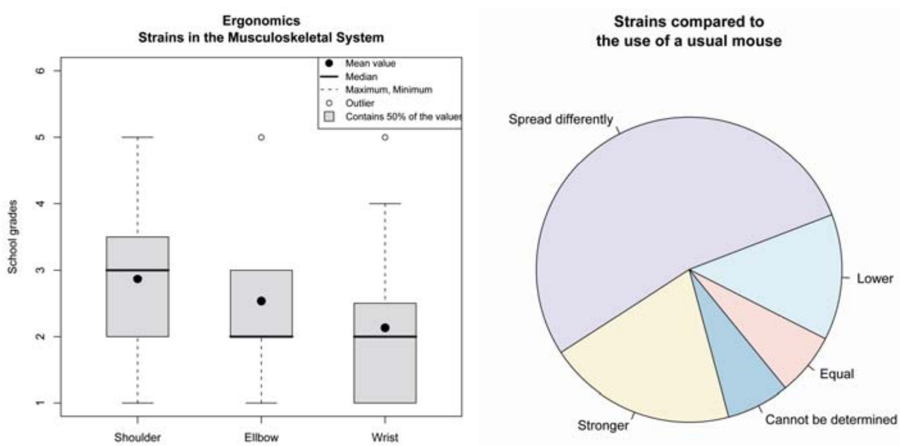


Fig. 12. Strains during the use of the pen tracking system, Left: Strains in shoulder, elbow and wrist, Right: Strains compared to the use of a usual mouse

our pen tracking system. This suggests that using the pen tracking device for short periods of time can temporarily alleviate pain. In addition the different interaction pattern compared to mouse based interaction may help to overcome remaining nociceptive impulses that became part of the motor program initially responsible for tissue damages [17].

Overall, more than half of the test group might use this kind of cursor-operation as alternative; surprisingly, also more than half of the test group without any relevant symptoms. This makes the system suitable for other scenarios such as in presentation boots and for simple games, as proposed by some test persons.

8 Conclusion

The presented pen tracking system provides an alternative input method for positioning a cursor. It employs a radically different work movement compared to the usual handling of the mouse. Furthermore, users seem to intuitively grasp the usage and accept pen based positioning as additional approach, however not as replacement.

It was possible to achieve such a solution with minimal hardware requirements using only off-the-shelf components such as a webcam and a coloured pen as well as no special environmental requirements regarding for example lighting. A demo video is available that shows the pen tracking system in action [19].

Currently we use two different approaches with different sets of characteristic properties alternatively to track the position of a pen. With just one approach it was not possible to consistently get a reliable positioning in changing environments. A joint approach combining the probabilities of the correct position has the potential to provide more stable results.

References

1. Bhattacharyya, A.: On a measure of divergence between two statistical populations defined by probability distributions. Bull. Calcutta Math. Soc. 25, 99–109 (1943)
2. Bruce, J., Balch, T., Veloso, M.: Fast and Cheap Color Image Segmentation for Interactive Robots. In: Proceedings of IROS 2000 (2000)
3. Buxton, W.: Human Input to Computer Systems Theories, Techniques and Technology, `http://www.billbuxton.com/`
4. Cheng, K., Takatsuka, M.: Estimating virtual touchscreen for fingertip interaction with large displays. In: OZCHI 2006, pp. 97–110 (2006)
5. Damany, S., Bellis, J.: It's Not Carpal Tunnel Syndrome! RSI Theory & Therapy for Computer Professionals Simax 1 (2001)
6. Hamette, P., Tröster, G.: Architecture and applications of the FingerMouse: a smart stereo camera for wearable computing HCI. Personal Ubiquitous Comput. 12, 97–110 (2008)
7. Hennessey, C., Noureddin, B., Lawrence, P.: A single camera eye-gaze tracking system with free head motion. In: ETRA 2006, pp. 87–94 (2006)
8. Hough, P.V.C.: Method and means for recognizing complex patterns. Patent (1962)
9. Jackson, R.: So how are your hands? Thoughts from a CS student with RSI. SIGCHI Bull. 28, 12–14 (1996)
10. Jähne, B.: Digital Image Processing, 6th edn. Springer, Heidelberg (2005)
11. Kohlbecher, S., Bardinst, S., Bartl, K., Schneider, E., Poitschke, T., Ablassmeier, M.: Calibration-free eye tracking by reconstruction of the pupil ellipse in 3D space. In: ETRA 2008, pp. 135–138 (2008)
12. Maat, L., Pantic, M.: Gaze-X: adaptive affective multimodal interface for single-user office scenarios. In: ICMI 2006, pp. 171–178 (2006)

13. Malik, S., Laszlo, J.: Visual touchpad: a two-handed gestural input device. In: ICMI 2004, pp. 289–296 (2004)
14. Meyers, B., Brush, A.J.B., Drucker, S., Smith, M.A., Czerwinski, M.: Dance your work away: exploring step user interfaces. In: CHI 2006, pp. 387–392 (2006)
15. Schlömer, T., Poppinga, B., Henze, N., Boll, S.: Gesture recognition with a Wii controller. In: TEI 2008, pp. 11–14 (2008)
16. Seeliger, I.: Ein optisches Pen-Tracking System zur alternativen Mauszeigersteuerung für RSI-Betroffene, Diploma-Thesis (2008)
17. Sorgatz, H.: Repetitive strain injuries, Unterarm-/Handbeschwerden aufgrund repetitiver Belastungsreaktionen des Gewebes, Orthopäde 2002 - 31, 1006–1014 (2002)
18. Pascarelli, E.: Dr. Pascarelli's Complete Guide to Repetitive Strain Injury: What You Need to Know About RSI and Carpal Tunnel Syndrom, p. 1. John Wiley and Sons, Inc, Chichester (2004)
19. http://www.medieninf.de/pentracker

Did "Minority Report" Get It Wrong? Superiority of the Mouse over 3D Input Devices in a 3D Placement Task

François Bérard[1], Jessica Ip[2], Mitchel Benovoy[2],
Dalia El-Shimy[2], Jeffrey R. Blum[2], and Jeremy R. Cooperstock[2]

[1] University of Grenoble,
Laboratoire d'Informatique de Grenoble,
BP 53 - 38041 Grenoble cedex 9, France
Francois.Berard@imag.fr
[2] McGill University,
Centre for Intelligent Machines,
Montréal, H3A 2A7, Canada
{jessicaip,benovoym,dalia,jeffbl,jer}@cim.mcgill.ca

Abstract. Numerous devices have been invented with three or more degrees of freedom (DoF) to compensate for the assumed limitations of the 2 DoF mouse in the execution of 3D tasks. Nevertheless, the mouse remains the dominant input device in desktop 3D applications, which leads us to pose the following question: is the dominance of the mouse due simply to its widespread availability and long-term user habituation, or is the mouse, in fact, more suitable than dedicated 3D input devices to an important subset of 3D tasks? In the two studies reported in this paper, we measured performance efficiency of a group of subjects in accomplishing a 3D placement task and also observed physiological indicators through biosignal measurements. Subjects used both a standard 2D mouse and three other 3 DoF input devices. Much to our surprise, the standard 2D mouse outperformed the 3D input devices in both studies.

Keywords: 3D placement, input device, stress.

1 Introduction

The growing availability of low-cost graphics rendering hardware and associated 3D applications targeting non-specialists (e.g., Google Earth, Sketchup) has focused increased attention on the problem of improved interaction paradigms for 3D tasks, for example, positioning an object in a volume. Such tasks are inherently more difficult to perform than 2D tasks because of the additional degree of freedom in which users must operate. However, we live and interact with objects easily and efficiently in the 3D physical world. Thus, it seems natural that computer systems should be able to leverage these physical-world skills for interacting with 3D virtual worlds. This intuition is expressed in motion pictures such as "Minority Report", portraying compelling scenarios of possible 3D interactions. To make this transfer of skill possible, new devices have been developed that better match the high DoF of our hands than the

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 400–414, 2009.

mouse does. Indeed, many prototypes have been proposed in the literature [1-3], some of which have evolved into commercial products (e.g., the *SpaceNavigator*™[1]). Nevertheless, despite the wealth of research and development into 3D input devices, the mouse remains the dominant input device for performing 3D tasks, at least in the case of desktop applications such as Computer Aided Design (CAD) or 3D modelling.

This paper presents an effort to understand the preference for the mouse. Specifically, we attempt to answer the fundamental question: does the mouse persist as the dominant input device for 3D tasks because of factors beyond performance, for example, cost or ingrained preferences, or is it the intrinsic qualities of the mouse that make it better suited for 3D interaction than high DoF devices? While the latter possibility appears counterintuitive, it is worthwhile to investigate because it could have significant impact on 3D interaction research.

This question, however, must first be expanded. Interaction is not defined by an input device alone, but by the combination of a device and an interaction technique. In the example of 3D object rotation, the mouse is typically used with the virtual sphere technique [4] while a free-space device is used with a direct mapping (either absolute or relative) [3]. Thus, each device must be matched with its most suitable interaction technique in making performance comparisons, rather than choosing a single interaction technique for all devices. Moreover, "3D interaction" encompasses a large spectrum of tasks, including simple object picking to more complex 6 DoF docking. Different combinations of device and interaction technique may be better suited for some of these tasks, but not for others. However, regardless of the diversity of tasks and possible choices of device and interaction technique, users have primarily settled on the mouse, so the fundamental question raised above remains valid.

In the remainder of this paper, we first review related literature and introduce our methodology for investigating this question. In particular, we justify our choices of benchmark task, set of devices and interaction techniques. We then report on two experiments comparing user performance on a 3D object placement task with a mouse and three other higher DoF devices. Finally, we offer some conclusions and suggest directions for possible further research.

2 Related Work

Several proposals for a suitable device for object placement in 3D can be found in the literature. These include the *Bat* [3], a free-space, 6 DoF device held in the hand; the *Roller mouse* [2]: a regular mouse augmented with wheels on top of the case that add one DoF for the control of cursor depth; the *Rockin'mouse* [5], another augmentation of a regular mouse that tilts around its base to provide one or two additional DoFs; and the *GlobeFish* and *GlobeMouse* [1], which decouple translation and rotation at the device level: translation is isometric, or elastic rate-control, and rotation is isotonic using a trackball. The performance of most of these devices has been the subject of formal user experiments. Unfortunately, fair comparisons between them are rare as the experiments tend to differ from each other in the performed task, and most experiments neglected to include the mouse in the comparisons.

[1] http://www.3dconnexion.com/3dmouse/spacenavigator.php

Previous studies have shown that in some situations, the mouse is less efficient than higher DoF devices. McMahan et al. [6] found that the mouse was less efficient than 6 DoF devices on a 6 DoF task (docking). Hinckley et al. [7] show that a free-space device is more efficient than mouse-based interaction for object rotation, a 3 DoF task. We chose placement rather than docking and rotation because we consider it a more fundamental and frequent task, as described in further detail in Section 3.1 below.

Balakrishnan et al. observed that users typically interact both in 2D and 3D, and 3 DoF (or greater) devices are usually poorly suited for 2D interaction. This motivated their development of the *Rockin'mouse*, a 3D device modelled after a regular mouse, but which provides an additional DoF from its tilting base. The *Rockin'mouse* was shown to be superior to the mouse [5] for placement tasks, although this result relied on the use of a non-standard interaction technique that was arguably biased toward a 3 DoF input device. We go one step further in examining the possibility that the regular 2D mouse could actually be *more efficient* for 3D object placement than higher DoF devices, regardless of the 2D context in which 3D interaction takes place. We evaluate this possibility by testing the mouse in conjunction with its best-suited interaction technique: a conventional 4-views display environment.

More recently, Teather et al. [8] tried to discover what properties of the mouse make it a *good* device for 3D placement. They found that the mouse outperformed a 3D tracker (Intersense IS900) when the tracker was used as a “mouse emulator” (i.e., constrained to a 2D surface). They hypothesized that the higher resolution of the mouse could be the key factor in its superiority. However, the tracker was used in *absolute* mode (like a tablet) while the mouse was used in its natural *relative* mode; this may have contributed to the observed difference in performance. They also found that both the mouse and the “mouse emulation” tracker outperformed the tracker when used as a free-space device, although the interaction technique was based on sliding one object over top of other objects in the scene, thus reducing the task from 3 DoF to 2 DoF. We are interested in evaluating the performance of the mouse on a true 3 DoF task without constraining the necessary operations to a 2D surface.

In the following section, we describe our methodology in trying to compare, as objectively as possible, the performance of the mouse and higher DoF devices in a 3 DoF object placement task.

3 Methodology

Our goal was to establish whether the dominance of the mouse is due to its intrinsic qualities for 3D interaction, or to *other factors*: its broad availability, the mixing of 2D and 3D tasks in user's typical work, or the long user habituation to the mouse. We conducted a set of studies comparing user performance in a 3D task while using the mouse and a set of credible, representative examples of 3 DoF input devices. We designed these studies so that pure 3D performance was isolated from the *other factors*: device availability was obviously not an issue in a controlled experiment, the experiments only involved a pure 3D task, and the experiments were designed to last sufficiently long so that the learning curve for each device could be inferred, providing an indication of device performance after user training. We also aimed at a direct measurement of device effect on stress through the recording and analysis of the participants' biosignals during the experiment. We now explain our rationale for selecting the benchmark task, the tested devices, the interaction techniques, and the choice of recorded biosignals.

3.1 The Task

One of the most frequent tasks in 2D graphical interaction is the selection of objects through spatial acquisition. This suggests, similarly, that target selection in 3D scenes would be a good candidate task. However, as the 3D scene is always perceived through 2D projections (on the screen or on the retina), target selection in 3D can be simplified to a 2D task. In the physical world, this is done by closing one eye and pointing at an object. In the digital world, the equivalent is 2D selection of the *projection* of the object on the screen, otherwise known as “picking”. There is little incentive, then, to use a high DoF device when the mouse is available and has been highly optimized for this task, especially when combined with smart selection techniques such as the bubble cursor [9], or semantic pointing [10].

Another frequent task is the placement of objects at arbitrary locations. This may involve changing the position of some object in a 3D modeling application, or positioning a marker within volumetric medical data. In either case, the system is unable to provide the assistance it can with picking tasks, because it has no clue as to where the user wants to place the object. The user is thus required to specify all three coordinates in some manner: in other words, this is a true 3D task. Indeed, there is presently no way to perform this task directly with a regular mouse. Instead, the standard solution is to decompose the task into two 2D placements, e.g., one in the *x-y* plane, and a second in the *x-z* plane. This suggests that 3 DoF devices should prove superior in performance by supporting task completion in a single action.

3D placement can be reduced to a 2 DoF task by the clever use of constraints: objects are, most of the time, placed next to each other. They are not left floating in the air. Hence, placement can be performed by sliding the moving object on top of the other (static) objects in the scene [8,11]: a 2 DoF task. While this kind of clever interaction technique illustrates the possibilities of performing 3D tasks with a 2D mouse, we want to compare the efficiency of the mouse with higher DoF devices in a true 3D task. For this reason, we chose the unconstrained positioning task in our experiment.

Other candidate tasks involving three or more DoF include object rotation, docking (a 6 DoF object placement including both a translation and a rotation), and navigation, which is equivalent to docking the camera that is viewing the scene. Our rationale for selecting placement as the benchmark task is as follows: if a number of 3D input operations are required by the application, we believe that users will choose the device that is best suited for the most fundamental and frequent of these (i.e., placement). Furthermore, previous studies [12] have shown that coordinating 6 DoF is too difficult to perform in a single action, and thus, 6 DoF tasks such as docking are actually performed by a sequence of 3 DoF translations (placements) and rotations even on 6 DoF devices.

3.2 The Devices

Our objective was to compare the mouse against devices that are considered state of the art for 3D input. In addition to the mouse, we ran the experiment with a free-space device, a *SpaceNavigator*™, and the *DepthSlider*, a mouse augmentation we designed to exploit bi-manual input. These are similar in design to systems described in previous literature, as summarized in the table below, and illustrated in Figure 1.

The DepthSlider. The *DepthSlider* is a physical slider, whose base is attached to the desk on the side of the keyboard opposite the mouse, and positioned so that its principal axis of movement is toward or away from the screen. It is inspired by the *Roller Mouse* [2] and the *Rockin'mouse* [5] in the sense that the objective is to augment the mouse to better support 3D interaction, rather than replace it. It differs from the earlier devices in that the mouse is not physically modified, and thus retains all the properties that make it suitable for 2D interaction. The *DepthSlider* is therefore used by the non-dominant hand to control depth, while the mouse is used with the dominant hand in the standard way (control of x and y values). The slider operates in *relative* mode, the same way as the mouse. The depth of an object is only controlled by the slider when the mouse "holds" it, that is, with the pointer on the object and the mouse button depressed. In consequence, the mouse controls the clutching and un-clutching of the slider. Un-clutching is necessary when the slider reaches either of its physical limits: the object is then released so that the slider can be moved away from the limit. Clutching again allows the user to resume moving the object.

Fig. 1. The devices (mouse not shown). From left to right: the *DepthSlider* alone and on its rail, the *SpaceNavigator*™, and the free-space device *Wii Remote*™ tracked by a *Vicon*™.

Table 1. Tested devices and related systems

Tested device	Properties	Related system
Mouse	2 DoF isotonic	
DepthSlider (mouse+slider)	2+1 DoF, isotonic	Roller Mouse and Rockin'mouse
SpaceNavigator™	6 DoF elastic	GlobeFish and GlobeMouse
free-space	6 DoF isotonic	Bat

The SpaceNavigator™. The *SpaceNavigator* (or simply, the *Navigator*) is a widely used, commercial device with 6 DoF. It provides elastic rate-control, which is considered optimal for coordination and minimizing fatigue [13].

The *Navigator* is intended for desktop use in the non-dominant hand so that the user does not have to switch hands to operate the mouse. In our experiments, the *Navigator* is used in this configuration, while the dominant hand rests on a mouse for target validation (pressing the right mouse button). We only use the 3 DoF translational control of the *Navigator*, activated by applying lateral forces to the device. Reducing the DoF to translation only is in fact common and supported explicitly by the device driver, as this helps avoid unintended translations resulting from rotation, and vice versa.

The free-space device. We designed a free-space device, inspired by the *Bat* [3], using the Nintendo *Wii Remote*™, with the addition of infrared reflectors to track translation and rotation with a *Vicon*™ motion capture system. The remote was selected mainly as an ergonomic hand-held device, but its buttons were also used for clutching ("B" trigger button below the remote) and to indicate task completion ("A" button on top).

3.3 Interaction Techniques

We expected performance to vary significantly depending on the choice of interaction technique. Since our study concerned the dominance of the mouse in 3D interaction, our approach was to study first whether or not the most *common interaction technique* used with the mouse was the key factor. We did this by conducting an initial experiment where all devices used the same (common 4-view) interaction technique. In a second experiment, we tried to optimize the interaction technique used with each device.

Subjects were asked to move the *object* (the white cube in Figure 2) so as to enclose the *target*, represented by the red cube, and confirm their placement by pressing a button (the right mouse button for all devices except for the free-space device, which used the *Wii Remote*™ "A" button). The target had to be completely

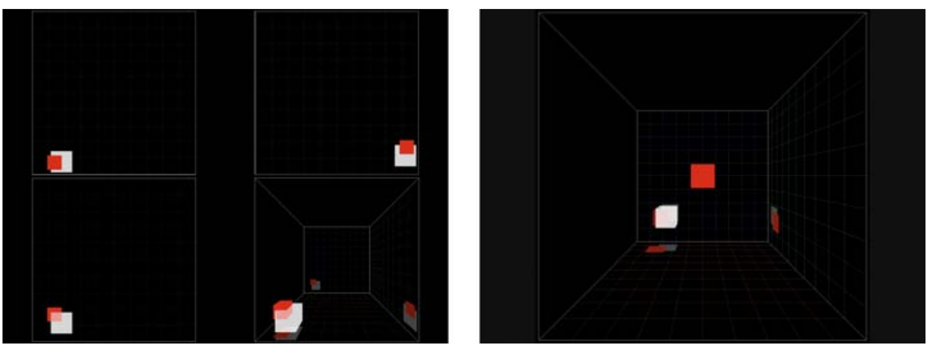


Fig. 2. Graphical display. Left: 4-view mode (top, side, front, and perspective view of the scene). Right: fullscreen perspective mode. The translucent white cube is the controlled object; the opaque red cube is the target location.

surrounded by the object for the placement to succeed. Otherwise, when the button was pressed, graphical feedback indicative of the error was provided by changing the color of the white cube to red for one second. Object and target were both represented as cubes and the object was made semi-transparent to provide the benefits of the Silk Cursor [14]. Both cubes cast "shadows" on the sides of the scene as additional location cues.

First experiment. The scene was rendered in four different views as illustrated in Figure 2, left. Three of the views were orthographic projections allowing interaction in two dimensions of the scene (front: *x-y*, top: *x-z*, and side: *y-z*). The fourth view displayed a more natural perspective rendering of the scene.

With the mouse and the *DepthSlider* (which both involve the mouse for placement), the object must first be acquired before being moved. Object acquisition is done in the classical way: the mouse *pointer* is moved on over the object, the left mouse button is depressed, and the object is then coupled with the device such that its position is controlled until the left mouse button is released. In the mouse condition, it is necessary to execute a minimum of two sequences of object acquisition and placement in two different views because each view only allows the control of 2 DoF in the 3 DoF task. In the *DepthSlider* condition, only one object acquisition is necessary because the mouse controls the x/y position of the object while the slider controls the z position. In contrast with the mouse and the *DepthSlider*, the *Navigator™* and the free-space device are always coupled with the object. No acquisition is necessary as the object moves as soon as the devices are controlled by the user. The rationale is that these devices are specialized for 3D operations; hence, they are meant to be used in conjunction with a mouse for general 2D operations such as object selection (picking).

The free-space device is used in *absolute* mode: the mapping of the physical location of the device to the virtual location of the object is fixed. The control/display (C/D) ratio is empirically chosen once for all experiments, and the origin of the mapping is re-defined before each experimental session so that participants can comfortably reach every location of the virtual scene.

Second experiment. After running the first experiment and seeing the results, we realized that the four views were not a good match for the high DoF devices. The orthographic views encourage the decomposition of motion into 2D segments, but the 3D devices should be more efficient when all dimensions of the task are solved in a coordinated manner. We thus designed the second experiment where all devices but the mouse are used with a more natural, full screen, perspective view of the scene (Figure 2, right). In addition, the view is rendered in stereo in order to improve depth perception.

Object acquisition was the same as before, but we introduce a new *relative* mode of operation for the free-space device. When used in relative mode, the object is only coupled to the object when the clutching button of the *Wii Remote*™ is depressed (the "B" trigger situated below the remote), as with the *Bat* [3]. This permits a dynamic C/D ratio: we set a strong gain during fast motion to allow quick positioning during the early phase of motion, and the gain is reduced to a small value during slow motion to facilitate accurate positioning. A dynamic C/D ratio is the main motivation for the introduction of the relative mode, as we observe in absolute mode that it is not possible to provide both fast initial positioning and accurate final positioning.

3.4 Biosignals as a Measure of Stress

By recording the biosignals of the participants while performing the task, we obtain a continuous measure of the level of stress induced over time by the input device being used. We believe that the use of biosignals to measure stress presents a novel and practical way to help gauge the usability of devices. We accomplished our stress assessment by using three measures of autonomous nervous system activation, namely galvanic skin response (GSR), blood volume pulse amplitude (BVP) and heart rate (HR).

Galvanic skin response (GSR) measures the electrical conductance of the skin. The signal can be decomposed into skin conductance responses (SCR), related to short events, and the skin conductance level (SCL), related to the underlying basal arousal activity. The GSR is often the primary psychophysiological measure used when gauging emotional activation as it responds very quickly (1-3 seconds after onset of stimulus) [15] and correlates highly with stress or anxiety. The blood volume pulse (BVP) signal is an indicator of blood flow using a photoplesthysmyography [15]. The amplitude of the blood volume pulses tends to decrease following sympathetic arousal [16,17]. The heart rate (HR) is computed from the raw BVP waveform by finding consecutive local maxima. An increase in sympathetic activity will increase the heart rate. In summary, higher stress is detected with lower BVP values and higher GSR, SCR, and HR.

4 First Experiment

4.1 Goal and Hypothesis

The goal of the initial experiment was to determine which device is the most efficient and causes the least fatigue during 3D placement. Our hypotheses were as follows:

- H1a: some of the present authors anticipated that direct specification of 3D position by hand movement in free space would outperform all other methods. The apparent naturalness of placing physical objects in a similar manner suggested that such interfaces represented an obvious improvement over devices constrained to operations in a 2D plane such as a mouse.
- H1b: other authors expected that the *DepthSlider* would be the most efficient device because they assumed that manipulating a stable device on the table surface would be key to task performance, outweighing the benefits of the natural mapping of a free-space device. They further expected that two-handed interaction could be quickly mastered, as shown in previous studies, and that the additional DoF offered by the slider would improve performance compared to the mouse, alone.
- H2: we expected to measure signs of fatigue from use of the free-space device because it does not allow the hand and the arm to rest on the table.

4.2 Task and Apparatus

Participants were asked to place a translucent white cube so as to enclose a red target cube, as described in Section 3.3. Subjects sat at a desk, facing an LCD monitor. In the free-space condition, the desk was removed, leaving the dominant hand able to move freely. A Vicon motion capture system tracked the free-space device in 6 DoF and the 1 DoF of the *DepthSlider*, using infrared-reflective markers attached to each device.

Graphical feedback (Figure 2, left) was displayed on an LCD screen (1280x1024 pixels resolution, 19" diagonal with a 4:3 aspect ratio, 60 Hz refresh rate), using custom-built software on top of the OpenSceneGraph library (http://www.openscenegraph.org/).

Physiological data was collected with Thought Technology's biofeedback system (http://www.thoughttechnology.com). All signal were measured by sensors strapped to fingertips. For consistency, the sensors were placed on each participant by the same experimenter. As a normalization step, the average of the one minute silent baseline data preceding each trial was subtracted from the “active” data of the participant. The data was then divided by the range of the participant's active data (i.e., *max* - *min*) to account for inherent physiological differences between participants.

4.3 Design and Procedure

In order to study the effect of device usage duration on performance and induced stress, trials were grouped into six consecutive *trial blocks*. We used within-subject design with independent variables *input device* (mouse M, Navigator N, DepthSlider D, and free-space Wii remote W), and *trial block* (1st, 2nd, ..., 6th). Each trial block consisted of 3 *runs*. Each run consisted of 9 targets combining 3 Fitt's indices of difficulty (ID = 1, 2, 4) with 3 target widths (7, 14, and 22 pixels). The orientation towards the target was randomized, as was the target's presentation order within blocks, but the same random sequence was used for all conditions and all participants. The order of device presentation was balanced between all participants.

Ten participants (5 male, 5 female) took part in the experiment. All were members of our laboratory, right-handed, and none had experience with the *Navigator*, the free-space device or the *DepthSlider*. The experimental design was as follows: 10 participants x 4 devices (mouse, Navigator, DepthSlider, free-space) x 6 blocks of trials x 27 targets = *6480 placements* in total. Participants took part in four sessions (one for each device). Sessions lasted approximately half an hour each and were separated by at least four hours to ensure sufficient recovery time from any fatigue induced by the previous session. At the beginning of a session, the task and the mode of operation of the tested device were explained verbally. Participants were told that they should perform the task as fast as possible while trying not to make placement errors. They familiarized with the device on the first 36 trials (4 runs) which were used for training. The actual trial blocks were then executed immediately after the training with no rest between the training and the first block, nor between the 6 blocks. The goal was to create a task that was sufficiently demanding for fatigue to be measured. The dependent variables measured were placement time, placement error, mean blood volume pulse amplitude, mean heart rate, mean galvanic skin response, and the number of SCRs.

4.4 Results

Figure 3 (top-left) represents the mean of total session time and mean of error rates across all participants for each device. An ANOVA on the data shows a strong main effect of device on placement time ($F_{3,9}$=10.7, $p < 0.001$) and error rate ($F_{3,9}$=72.9, $p < 0.001$). Post-hoc analysis shows that both the mouse and *DepthSlider* are significantly more efficient ($p < 0.01$) than the *Navigator* (30% faster) and the free-space device (40% faster), but the difference was not significant between the mouse and the *DepthSlider*, nor between the *Navigator* and the free-space device. Participants made significantly more errors in the free-space condition than in all other conditions ($p < 0.001$). The only other significant difference in error rate was between the mouse and *Navigator* conditions ($p < 0.05$).

Figure 3 (top-right) shows the evolution of trial block completion time across the six blocks. There was a significant effect of block number on completion time in the mouse condition ($F_{5,9}$=14.1, $p < 0.001$), *DepthSlider* condition ($F_{5,9}$=7.79, $p < 0.001$) and the *Navigator* condition ($F_{5,9}$=6.89, $p < 0.001$), but not the free-space condition ($F_{5,9}$=0.944). There was no significant effect of block number on number of errors.

The results of the mean physiological data for each input modality across participants is shown in Figure 3 (bottom-left). An ANOVA yielded a significant difference for the GSR ($F_{3,36}$=2.89, $p < 0.05$), #SCRs ($F_{3,36}$=7.41, $p < 0.01$), BVP ($F_{3,36}$=3.9, $p < 0.05$), HR ($F_{3,36}$=3.72, $p < 0.05$). Figure 3 (bottom-right) shows the mean GSR,

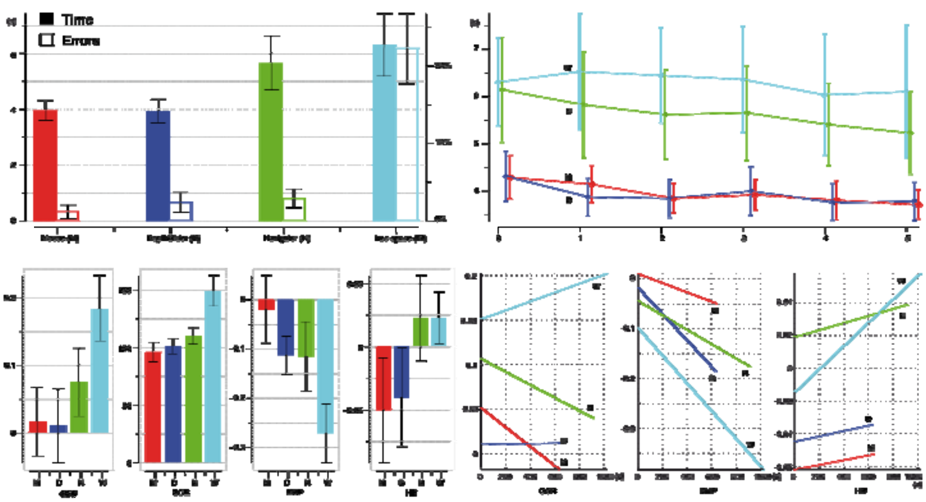


Fig. 3. Results of first experiment. All figures with 95% confidence intervals. *Top-left*: Means of average placement time (left axis, in seconds) and error rate (right axis, in %) per device accross all participants. *Top-right*: Means of average placement time (s) per condition and trial block accross all participants. *Bottom-left*: means of normalized conductance (GSR), #skin conductance responses (SCR), normalized amplitude (BVP), and normalized heart rate (HR) per device accross all participants. *Bottom-right*: Linear regressions of the mean biosignals waveform accross all participants over task completion time (in s) per device (same scales as bottom-left).

BVP, and HR data across all the participants for each input device type as it varies over the time of the experimental session. The data were first re-sampled to the average time-of-completion for the given device, then fit to a line with a robust linear regression ($r > 0.59$, $p < 0.05$ across all regressions) in order to reveal the underlying timewise trend. The most noticeable result is that the free-space device led to the greatest stress as it induced the highest mean GSR and #SCRs, and the lowest mean BVP. It also produced higher HR than the mouse and the *DepthSlider*, but for this measure, there was no significant difference with the *Navigator*. For the GSR, the mouse and *Navigator* have a downward trend, indicating that the subjects' stress may be decreasing as time proceeds while using these devices. Conversely, the free-space device seems to induce progressively higher GSR. The higher rates of decline of BVP and increase of HR indicate a more powerful effect of the free-space device on stress.

5 Second Experiment

5.1 Objectives and Variations from the First Experiment

The main goal of the second experiment was to test whether the previously observed superiority of the mouse holds when high DoF devices were used with a graphical display that better promotes coordinated input than the 4-views layout, in particular, a stereoscopic display in which depth can be observed directly. A secondary goal was to evaluate if the free-space device would perform better in *relative* than *absolute* mode. The two experiments were similar in design apart from the following:

- *Graphical display*. Graphical feedback was projected on a screen approximately 1.5 m from the participant's chair. In the mouse condition, a single projector (1024x768 resolution) provided the same feedback as in the first experiment, albeit at reduced resolution. In the other conditions, an additional 1024x768 resolution projector was used, in combination with polarized filters and glasses, in order to create a stereo rendering of a single, fullscreen, perspective view of the scene.
- *Design*. As the first experiment indicated quick learning of the new devices, mostly during training, we did not feel it necessary to reproduce the long experimental sessions, hence only three blocks of trials were used. A relative mode was added for the free-space device, as well as a higher Fitts' Index of Difficulty (ID = 6). We situated the targets along the eight diagonals from the center (initial position) to the workspace vertices to control the orientation to target. Ten right-handed participants (8 male, 2 female) took part in the experiment; all were members of our laboratory. The experimental design consisted of 10 participants x 5 devices (mouse M, DepthSlider (Stereo) DS, Navigator (Stereo) NS, free-space absolute (stereo) WS, free-space Relative (Stereo) WRS) x 3 blocks of trials x 4 indices of difficulty (1, 2, 4 and 6) x 8 diagonal orientations = *4800 placements* in total.

5.2 Results

Figure 4 describes the results of the second experiment. An ANOVA revealed a significant effect of device on task accomplishment time ($F_{4,9}$=5.83, $p < 0.01$) and error

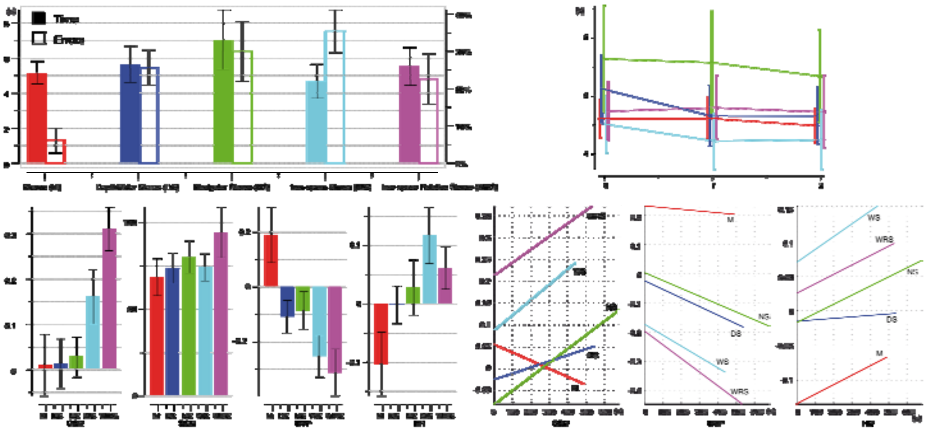


Fig. 4. Results of second experiment. Legend is the same as Figure 3.

rate ($F_{4,9}$=17.6, $p < 0.01$). The *Navigator* was significantly slower than all other devices. Contrary to the first experiment, there was no significant difference in time between the mouse and the free-space device in either absolute or relative mode. However, error rates were significantly higher with all devices compared to the mouse. The best competitor, the free-space device in relative mode, had a 22.6% error rate, more than three times higher than the mouse, at 6.04%.

An ANOVA on the physiological results revealed a significant effect in the GSR ($F_{4,45}$=6.64, $p < 0.01$), BVP ($F_{4,45}$=4.58, $p < 0.01$), and HR ($F_{4,45}$=4.01, $p < 0.01$), but not in the SCRs ($F_{4,45}$=2.82, $p > 0.05$). For the GSR, there is again a clear distinction between the tabletop devices and the two modalities for the free-space device. The latter produced significantly higher GSR, indicating higher stress, with the relative mode the highest. The BVP results are similar: the free-space device produced the lowest BVP amplitude, another indicator of stress, whereas this value was significantly higher with the mouse than the two other tabletop devices. The mouse also induced the lowest heart rate. The mean trends shown in Figure 4, bottom-right, ($r > 0.65$, $p > 0.05$ across all regressions) indicate that the stereo visualization mode may induce higher stress levels over time as all the GSR slopes except the mouse (which was not viewed in stereo) are positive.

6 Discussion

The most natural device (free-space) suffered the worst performance in terms of time (first experiment) and error rate (both experiments). The task was particularly difficult to perform with the free-space device, especially when maximal accuracy was required to position the center of the object in a 4x4x4 pixel cube. Several factors may explain this difficulty: tracking jitter, parasitic motion induced by clicking the button to terminate the placement, hand instability in free space, and fatigue. In the second experiment, more natural motion trajectories were performed, since these did not need to be decomposed into the 4-view display planes. This improved speed but decreased accuracy. Since the results contradict our hypothesis H1a, we considered whether the task might be biased against the free-space device, or whether this device is not designed for accurate placement. Indeed, more complex 3D tasks, such as docking or deformation, might yield completely different results. However, we believe that the placement task is a fundamental building block of other more complex tasks, and thus the lack of accuracy observed here will likely translate to other tasks. The adequacy of the free-space device appears to depend more on the required accuracy than on the nature of the task.

To determine whether motion tracking was the source of inaccuracy, we recorded the static position of the free-space device on a table for three minutes, then repeated the measurement with five participants attempting to hold the device stationary, without support. Analysis revealed that our Vicon system offered stability of 0.034 mm (equivalent to 740 dpi) whereas the most stable participant maintained the device within an average radius of 0.15 mm over successive time intervals of 0.5 s (the mean for the five participants was 0.48 mm). This indicates that the lack of accuracy is due mainly to the difficulty of holding the free-space device stable in the air. As such, desktop devices (mouse, *DepthSlider* and *Navigator*) benefit from the stabilizing effect of the hand resting on the desk while the fingers execute precise control. This explains the low error rate (< 3%) of desktop devices in the first experiment compared to the free-space device (22%). However, in the second experiment, the *DepthSlider* and *Navigator* desktop devices had error rates over 25% when used with a single perspective view display, whereas the mouse remained at 6% with the 4-views display. While the perspective view was more natural and encouraged coordinated (i.e., faster) motion, it appeared inferior for placement evaluation. We did not consider the level of indirection in the system. In all our experiments, placement is done through *tele-operation*, with the placed object distant from the hand performing the action. Different results may be obtained with the object co-located with the hand (e.g., through a head-mounted display), as this would allow proprioception, but would not help to stabilize the hand.

In summary, we find that both the most natural device and most natural display (i.e., most resembling properties of the physical world) are not the best suited for accurate placement tasks. On the contrary, less natural tools may offer important attributes (e.g., accuracy for the mouse, optimized viewpoints for the 4-views) that are better suited for some tasks. Naturalness may offer strong benefits for immersion and ease of use, but our study indicates that it can also decrease performance when accuracy is required. While this may seem predictable a posteriori, it contradicts the strong intuition that more “natural” interaction is systematically superior.

Surprisingly, the *DepthSlider* was not faster than the mouse, contradicting our hypothesis H1b. While participants did not demonstrate any difficulty with the two-handed interaction, the task did require significant accuracy from the non-dominant hand, which could explain the higher level of stress induced compared with the mouse.

7 Conclusion

We conducted two experiments intended to study the dominance of the mouse in desktop 3D interaction. We compared both user performance and biosignals measured on a fundamental 3D task: object placement in the scene, with the traditional 2D mouse and various 3 DoF devices. Contrary to our expectations, the experiment demonstrated that despite the lack of a third degree of freedom, the mouse was more efficient than the 3 DoF devices for accurate placement. Its accuracy and low level of induced stress appear to compensate for the requirement to decompose 3D tasks into several 2D tasks. The experiments also demonstrated, through biosignals analysis, what had been hypothesized but never shown in the literature: that the free-space devices induce significantly more stress than desktop devices. This indicates that biosignals analysis can be a powerful addition to the set of tools used in usability evaluation.

References

1. Froehlich, B., Hochstrate, J., Skuk, V., Huckauf, A.: The globefish and the globemouse: two new six degree of freedom input devices for graphics applications. In: ACM Conference on Human Factors in Computing Systems (CHI), pp. 191–199. ACM Press, New York (2006)
2. Venolia, D.: Facile 3d direct manipulation. In: ACM Conference on Human Factors in Computing Systems (CHI), pp. 31–36. ACM Press, New York (1993)
3. Ware, C.: Using hand position for virtual object placement. The Visual Computer 6(5), 245–253 (1990)
4. Chen, M., Mountford, S.J., Sellen, A.: A study in interactive 3-D rotation using 2-D control devices. ACM SIGGRAPH Computer Graphics 22(4), 121–129 (1988)
5. Balakrishnan, R., Baudel, T., Kurtenbach, G., Fitzmaurice, G.: The rockin'mouse: integral 3d manipulation on a plane. In: ACM Conference on Human Factors in Computing Systems (CHI), pp. 311–318. ACM Press, New York (1997)
6. McMahan, R.P., Gorton, D., Gresock, J., McConnell, W., Bowman, D.A.: Separating the effects of level of immersion and 3D interaction techniques. In: ACM Symposium on Virtual Reality Software and Technology, pp. 108–111. ACM Press, New York (2006)
7. Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., Kassell, N.: Usability analysis of 3d rotation techniques. In: ACM Symposium on User Interface Software and Technology (UIST), pp. 1–10. ACM Press, New York (1997)
8. Teather, R.J., Stuerzlinger, W.: Assessing the effects of orientation and device on (constrained) 3D assessing the effects of orientation and device on (constrained) 3d movement techniques. In: IEEE Symposium on 3D User Interface, pp. 43–50 (2008)
9. Grossman, T., Balakrishnan, R.: The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In: ACM Conference on Human Factors in Computing Systems (CHI), pp. 281–290. ACM Press, New York (2005)
10. Blanch, R., Guiard, Y., Beaudouin-Lafon, M.: Semantic pointing: improving target acquisition with control-display ratio adaptation. In: ACM Conference on Human Factors in Computing Systems (CHI), pp. 519–526. ACM Press, New York (2004)
11. Oh, J.Y., Stuerzlinger, W.: Moving objects with 2D input devices in CAD systems and Desktop Virtual Environments. In: Proceedings of graphics interface (GI), Canadian Human-Computer Communications Society, pp. 195–202 (2005)
12. Masliah, M.R., Milgram, P.: Measuring the allocation of control in a 6 degree-of-freedom docking experiment. In: ACM Conference on Human Factors in Computing Systems (CHI), pp. 25–32. ACM, New York (2000)

13. Zhai, S.: User performance in relation to 3D input device design. ACM Computer Graphics 32(4), 50–54 (1998)
14. Zhai, S., Buxton, W., Milgram, P.: The “Silk Cursor”: investigating transparency for 3D target acquisition. In: ACM Conference on Human Factors in Computing Systems (CHI), pp. 459–464. ACM Press, New York (1994)
15. Cacioppo, J., Tassinary, L., Berntson, G.: Handbook of Psychophysiology. Cambridge University Press, Cambridge (2007)
16. Picard, R., Vyzas, E., Healey, J.: Toward machine emotional intelligence: analysis of affective physiological state. IEEE Trans. on Pattern Analysis and Machine Intelligence (2001)
17. Shapiro, D., Jamner, L., Lane, J., Light, K., Myrtek, M., Sawada, Y., Steptoe, A.: Blood pressure publication guidelines. Psychophysiology (1996)

The MAGIC Touch: Combining MAGIC-Pointing with a Touch-Sensitive Mouse

Heiko Drewes[1] and Albrecht Schmidt[2]

[1] Media Informatics Group, LMU University of Munich,
Amalienstraße 17, 80333 München, Germany
heiko.drewes@ifi.lmu.de
[2] Pervasive Computing Group, University of Duisburg-Essen,
Schützenbahn 70, 45117 Essen, Germany
albrecht.schmidt@acm.org

Abstract. In this paper, we show how to use the combination of eye-gaze and a touch-sensitive mouse to ease pointing tasks in graphical user interfaces. A touch of the mouse positions the mouse pointer at the current gaze position of the user. Thus, the pointer is always at the position where the user expects it on the screen. This approach changes the user experience in tasks that include frequent switching between keyboard and mouse input (e.g. working with spreadsheets). In a user study, we compared the touch-sensitive mouse with a traditional mouse and observed speed improvements for pointing tasks on complex backgrounds. For pointing task on plain backgrounds, performances with both devices were similar, but users perceived the gaze-sensitive interaction of the touch-sensitive mouse as being faster and more convenient. Our results show that using a touch-sensitive mouse that positions the pointer on the user's gaze position reduces the need for mouse movements in pointing tasks enormously.

Keywords: Eye-tracking, eye-gaze pointing, touch-sensitive mouse, MAGIC pointing.

1 Introduction

Pointing and selecting are very important operations in graphical user interfaces (GUIs). Typical input devices for pointing are mouse devices, trackballs, touch pads, or touch screens. Studies on the performance of these input devices started 30 years ago [3]. Research on eye-tracking systems for controlling computers with eye-gaze started in the 1980s [2]. Using an eye-tracker as a pointing device is an obvious approach utilizing eye-gaze for computer input.

Eye-tracking technology exists now for many decades and eye-trackers matured over the last years [4]. Current video-based eye-tracker systems work unobtrusively and provide also head-tracking, so that the users can move their heads freely while working with the system. Since such eye-tracking systems consist mainly of video cameras and software, eye-tracking technology will fall in price and may be included in future systems without significant additional hardware costs. The high costs for an eye-tracker in the past made such systems only affordable in the field of accessibility.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 415–428, 2009.

Consequently, existing eye-tracker applications for computer input use eye-gaze as the only input modality. A typical example is eye-typing for quadriplegics. In contrast, we focus on the use of eye-gaze as an additional input modality for the masses. The main challenge is figuring out how to integrate eye-gaze as an additional modality and to find forms of interaction, which are appropriate and useful. Although Bolt [2] gave a vision of gaze-aware multimodal interfaces already in 1981, an interface working similar to human-human interaction using gaze and speech is still a topic of research [17] and far away from application in real life.

Our research starts with the simple idea to position the mouse pointer at the point where the user looks at when she or he touches the mouse. We enhance Zhai's MAGIC (Mouse And Gaze Input Cascaded) pointing [16] by combining it with a touch-sensitive mouse device, similar to the mouse presented by Hinckley [7].

1.1 Eye-Gaze Interfaces

Jacob [8] did the first systematical research on eye-tracking as a pointing device in GUIs. His research investigated basic gaze interaction techniques like selection, scrolling text, and choosing menu items from pop-up menus. He identified the problem of accidental invocation by inspection and called it the Midas-Touch problem. Eye-tracking interfaces typically use the dwell-time method to avoid the Midas Touch problem. In this method, the user has to look for a certain time at a particular position on the screen to activate a function. This approach has been proven to be very useful for disabled users [10]. However, normal users work more efficiently with standard interfaces, e.g. mouse and keyboard. The dwell time required to activate the function nullifies the speed benefits resulting from the proverbial fast eye movements. Typically, dwell time adds about 500 to 1000 milliseconds to the overall interaction time [11], and results in pointing tasks that take longer than with a traditional mouse. Shortening the dwell time is not a viable solution, since it results in an increase in accidental function invocations by just looking at the screen. Another possibility to solve the problem is the use of a gaze key. Pressing the gaze key triggers the action belonging to the GUI object the user is looking at. As shown in [15], the use of a gaze key also speeds up pointing tasks.

When using eye-tracking systems, pointing accuracy is a central issue. The eye-tracker system used in our experiments offers an accuracy of ±0.5° visual angle, which results in minimum target sizes of about 36 pixels on current desktop systems. In contrast, current window-based GUIs typically require a precision better than 16 pixels to select a menu item and an even higher precision to position a text cursor between two letters. The accuracy of the gaze position is not only a question of the eye-tracker quality, but also of the intrinsic nature of eye movement. The size of the fovea (the high resolution area of the retina) is about one degree and to see something clearly the eyes move so that the projection of the object of interest falls inside the fovea but not necessarily in the center of it. Hence, better hardware or algorithms cannot easily improve the accuracy of eye trackers. As a result, user interfaces designed for gaze interaction have to accommodate a larger threshold of pointing inaccuracy. Recent efforts have tried to solve the accuracy problem by adding intelligence to gaze positioning [14]. An intelligent algorithm positions the mouse pointer within the radius of inaccuracy at the target the user intended to activate. Other approaches use expanding targets [12] or enlargement of the region for eye-gaze input [9].

A further challenge for eye-gaze interfaces is the provision of feedback. In general, it is difficult to work with a system that does not provide visual feedback. A gaze pointer, analogous to a mouse pointer, does not work well because even little inaccuracies (e.g. calibration shift, intrinsic inaccuracy) set the gaze pointer next to the position where the eye is looking. When the eye tries to look at the gaze pointer the pointer moves again and this ends up in chasing the pointer across the screen. Typically, eye-trackers do not report small shifts of the eye gaze position as a consequence of noise filtering. This avoids chasing the pointer but leads to a jumping pointer hopping across the targets.

1.2 Pointing Devices

Currently used input technologies, i.e. keyboard and mouse, require mechanical movement and hence cause problems for some users. One common problem is that people after prolonged usage can suffer from the carpal tunnel syndrome. With the current trend towards bigger displays, the mouse movement distances will also increase and consequently the physical load on the hand muscles.

Additionally, people can have problems finding the position of the mouse cursor on the display [1]. There are several solutions to this problem found in commercial GUIs. One solution is a highlighting of the mouse pointer; another solution is to position the mouse cursor at the position where the user expects it. For example, the common Windows GUI offers two options in the control panel: (1) Show location of pointer when I press the CTRL key, (2) Automatically move pointer to the default button in a dialog box, and (3) the API knows the DS_CENTERMOUSE parameter for creation of dialogs. These mechanisms are designed to ease the use of the mouse and show that there is a potential area for improvement.

One approach that inspired our work is Zhai's MAGIC (Mouse And Gaze Input Cascaded) pointing [16]. The basic idea of the MAGIC pointing is positioning the mouse cursor at the gaze position on the first mouse move after some time of inactivity. The concept of MAGIC pointing is to use the gaze position for coarse positioning and the mouse for fine positioning. One problem of MAGIC pointing is the effect of overshooting the target because the hand and mouse are already in motion at the moment of pointer positioning. Zhai et al. suggests compensating this problem by taking the distance vector to the target and the initial motion vector of the mouse move into consideration when calculating the pointer position. In contrast, our research started with adding a touch sensor to the mouse and using the mouse to detect touch instead of movement. This approach does not require any compensation techniques, as the mouse does not move at the moment of pointer positioning. It also has the advantage that the user can invoke the gaze positioning at any time instead of waiting for the mouse inactivity timeout to expire.

2 The Touch-Sensitive Mouse

The work of Hinckley [7], which uses a touch-sensitive mouse to hide and display menus, inspired the initial design of the touch mouse. Our first naïve idea was to position the mouse pointer at the moment the users touches the mouse. Consequently, in

Fig. 1. The pictures show the first and second prototype of the touch-sensitive mouse

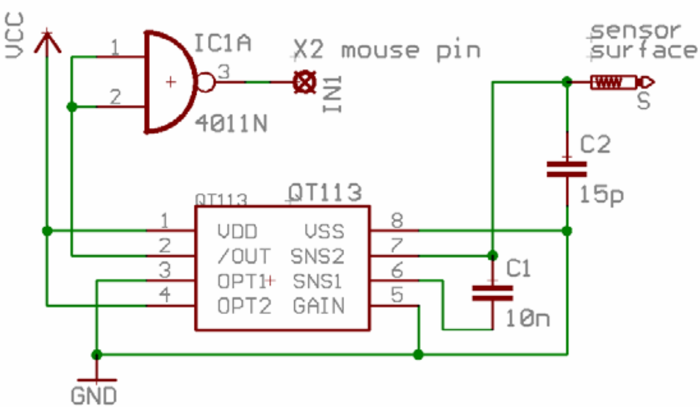


Fig. 2. Schematic of the circuit integrated into the mouse to detect touch

our first prototype the entire mouse was touch-sensitive. Observation during a pilot study showed that users tend to leave their hand on the mouse while looking for the next target to click. Therefore, we created a prototype where only the left mouse button is touch-sensitive (see Fig. 1). When the finger contacts the mouse key (rising edge of the touch signal), the system sets the mouse pointer at the gaze position reported from the eye tracker. This gives the user the possibility to check the pointer position before performing the click. As long as the finger stays on the touch sensor, the mouse behaves like an ordinary mouse.

The sensor for detecting touch uses capacitive sensing. The recognition time for the touch is critical, because the time interval between touching the sensor and performing a click can be very short. The first sensor chip we used turned out to be to slow as it reported the touch signal after the mouse click. Therefore, the sensor chip used in the project is a capacitive sensor (QT 113 [13], see Fig. 2 for the schematic) with a very fast response time of 30 milliseconds. We connected the output of the touch sensor chip to the X2 button of the mouse. The X2 button has an assigned meaning in the Windows operating system and generates an event by the interrupt-driven mouse driver. This helps to speed up the positioning process, which is critical for the user experience. The overall reaction for touch positioning of the pointer is in the order of 60 ms (XP OS time slice 10 ms, touch detection 30 ms, displaying on TFT screen 20 ms), which appears to the user as an "immediate" reaction.

3 The User Study

In a user study, we compared the pointing performance for a traditional mouse and the touch-sensitive mouse with gaze positioning. We used the commercial eye tracker ERICA [5] and software developed for the study.

We recruited 10 participants aged 23 to 46 years old. All users were regular computer users and familiar with a mouse device.

We utilized a within-subject design with two independent variables: input modality (traditional mouse or gaze position using touch-sensitive mouse) and background image (plain background or complex background, see Fig. 3). The dependent variable was the duration of a pointing task. The condition of a complex background was introduced, because it is reasonably good simulation of a situation where the user cannot find the mouse pointer. We used a fixed target size of 100 pixels (3°) for the experiment. Pilot studies showed that Zhai's concept of raw and fine positioning also works with the touch-sensitive mouse, and we wanted to focus on the triggering by touch and not go into a discussion of Fitts' law for the eye. The distances between the starting position of the pointer and the target were chosen at random with an average distance of 400 pixels. In each condition, we asked the participants to do 50 pointing tasks. For each task, they had to move the pointer over the target and make a left click.

At the beginning of the experiment, users received a one-minute explanation of the basic function of the system and then had some time (~1 min) to tryout the system and familiarize themselves with its functions. Then the participants carried out the pointing tasks for all four conditions, with short breaks in between. The first task was on a white background using a traditional mouse. In the second task, the participants used the touch-sensitive mouse. Afterwards, we asked the users with which input method they performed faster and which one is more convenient. The next two runs were a repetition of the first two runs but on complex background. After the experiment, we asked the participants again, which method they perceived as being faster and more convenient. We did not randomize the order of the experiments because we wanted to

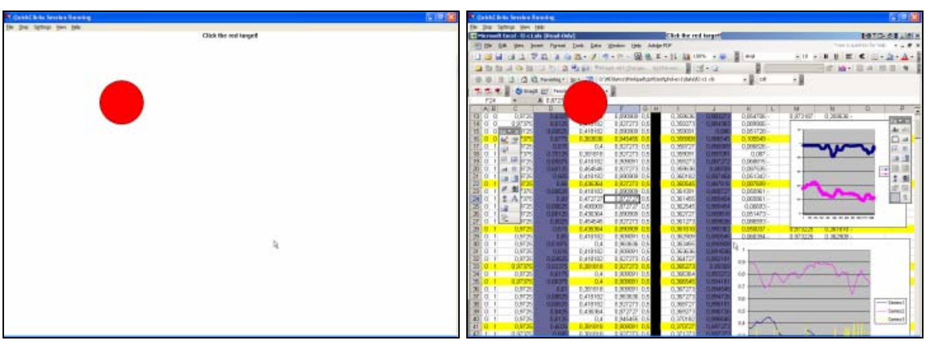


Fig. 3. The two background conditions. On the complex background, the mouse pointer is more difficult to find.

have the answers for the blank background condition first (see 4.1). As shown in the discussion later, the picture for the interaction speed with the touch-sensitive mouse is quite clear. We do not expect to get further insights by a repetition of the user study with randomized task order.

4 Results and Discussion

We took the medians from the 50 click of a task to avoid problems with outliers. Table 1 summarizes the results. It is easy to see that there is no speed benefit for gaze positioning on blank background but for complex background.

The values in Table 2 are the results of paired Student's t-test for four different combinations and give the probabilities that the compared data sets are from a distribution with the same mean value. There is strong significance that gaze positioning is faster when using a background, i.e. the user is not aware of the mouse pointer position. The values also show that mouse positioning takes longer on a complex background. The effect of better performance for gaze positioning on a background compared to no background (929 v. 1032 ms) is most likely an effect of learning due to not randomizing task order.

Table 1. Medians for the total time in milliseconds for each of the participants on blank and complex background and the resulting arithmetic mean over the entire group

Participant	blank background		complex background	
	mouse positioning	gaze positioning	mouse positioning	gaze positioning
P1	1097.0	751.0	1382.0	826.5
P2	971.5	1007.0	1066.0	806.0
P3	1121.0	716.0	1276.5	791.0
P4	1096.5	1066.5	1352.0	896.5
P5	932.0	871.0	1191.0	696.0
P6	926.5	1256.5	1066.0	1121.0
P7	1111.0	957.0	1217.0	891.0
P8	1211.5	1327.0	1412.0	1327.0
P9	1062.0	1482.0	1266.5	1277.0
P10	976.0	881.0	1101.0	656.0
Mean	**1051**	**1032**	**1233**	**929**
Std. Dev	94	253	128	234

Table 2. T-tests for four different task combinations

gaze vs. mouse pointing		background vs. blank background	
blank background	with background	gaze positioning	mouse positioning
0.823686	**0.002036**	**0.020738**	**0.000014**

4.1 The Subjective Speed Benefit on Blank Background

In the post-study interview, all but two users felt they were faster with the touch-sensitive device on the blank background, and all users felt they were faster with the touch-sensitive device on the complex background. However, the data show that their perceptions are only valid for the complex background. For the blank background, their completion times were equal.

When setting up the user study we experienced the phenomenon ourselves and for this reason, we put the question on subjective speed in the interview. Because we did not want to bias the participants with the task on complex background, where the gaze positioning brings a real speed benefit, we decided to start with the task on blank background for all participants.

Fig. 4 illustrates a potential explanation for users' perception that the touch-sensitive mouse is faster. It shows the steps for task completion on a timeline for the condition with a blank background. Even though the overall time to complete the task is the same for classical pointing and gaze positioning, the time it takes users to move their hands is shorter when using gaze. This suggests that the users do not perceive the time required moving the eye as time needed to perform the task. Thus, users experience the use of eye-gaze as faster than a traditional mouse.

4.2 A Real Speed Benefit on Blank Background

Fig. 5 plots the time measured for both pointing task against the index of difficulty ($\log_2$(distance/target size)), as typically done for the evaluation of Fitts' law [6]. As

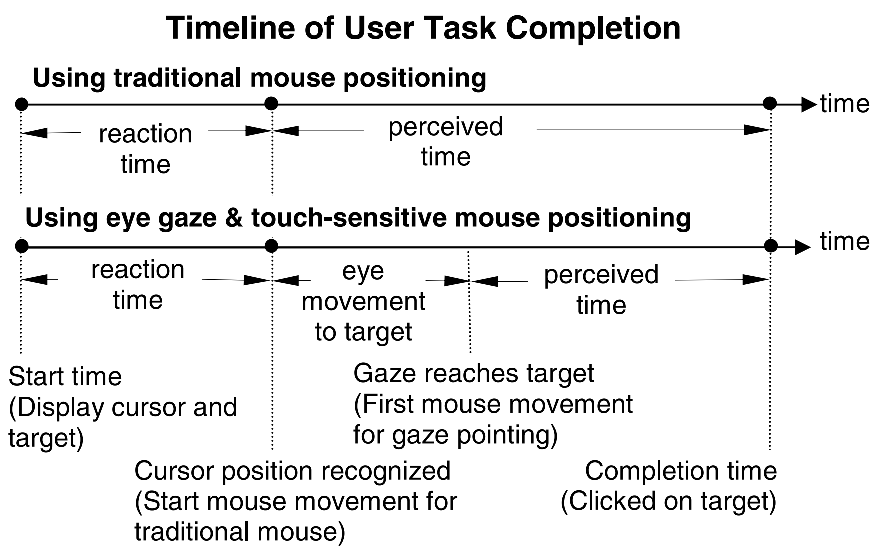


Fig. 4. Explanation of perceived speed-up with a touch-sensitive mouse and gaze positioning on a blank background

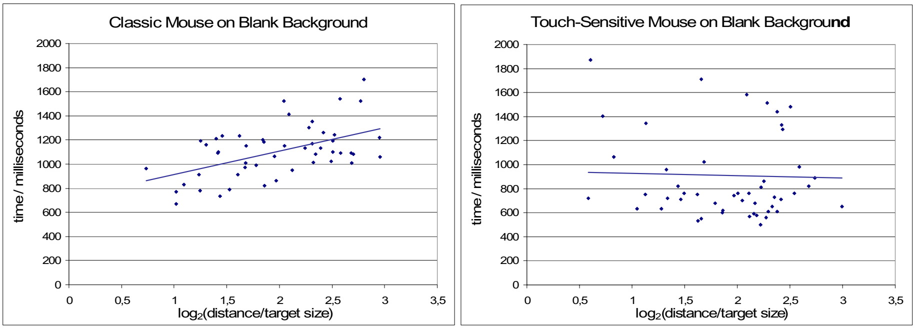


Fig. 5. Typical plots for the time to click the target against the index of difficulty for classical mouse and gaze positioning

expected, the plot for the classical mouse task shows a dependency of time from the distance between target and mouse pointer. This is in accordance to the experiments of Card [3]. In contrast, there is no dependency on the distance for the positioning by eye gaze. The explanation here is that the only task for the eye is to find the target and the only task for the hand is to click – the position of the mouse pointer does not matter for completing the task. The average distance of pointer and target of about 400 pixels (12°), given by the dimensions of the display used, was the reason that the completion time for the task with blank background was the same for classical mouse and touch mouse. This is the distance where both methods need the same time. It is the crossover-point of the constant function for the touch-sensitive device and the monotonic increasing function (Fitts' law) of the classical device. For shorter distances, the time to check whether the touch of the sensor positioned the mouse pointer correctly exceeds the time of a mouse move. For bigger distances, moving the mouse takes longer than the constant time of gaze positioning. As a consequence, we expect that there will be a speed benefit for the touch-gaze setup for longer distances, e.g. on very large screens or multi-screen setups.

4.3 The Speed Benefit on Complex Background

Every click-task started with displaying target and pointer. On blank background, the participants spotted both objects with pre-attentive perception. The data show mouse pointer and gaze move independently towards the target where they meet. Fig. 6 depicts a mouse trail and a gaze path. The dotted gray lines connect points for the same time on mouse and gaze trail. It is easy to see that the gaze is already at the target when the mouse only just started moving towards the target.

Fig. 7 shows the corresponding plot for the distance of gaze and mouse pointer to the center of the target over time. The gray area visualizes the size of the target.

On a complex background, it is possible to spot the big red target but not the mouse pointer. All participants started to stir the mouse to create a visual stimulus for finding the pointer. No participant moved the focus of the gaze onto the pointer;

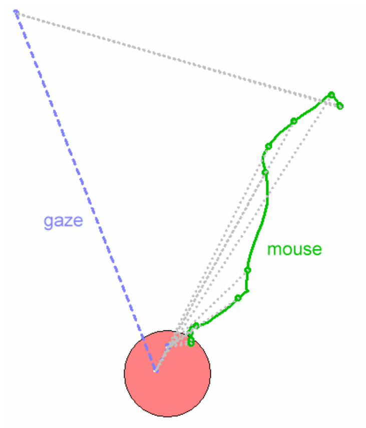


Fig. 6. Gaze (dashed) and mouse trail (solid) for the classical mouse task without background. The dotted grey lines connect points of same time.

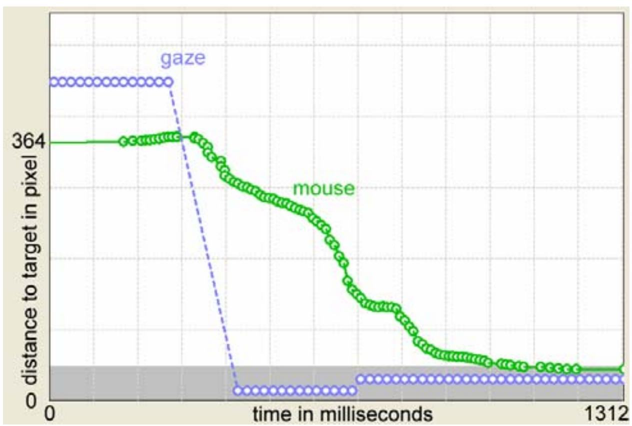


Fig. 7. Gaze (dashed) and mouse trail (solid) for the classical mouse task without background as a plot of distance to target over time

peripheral vision of movement is sufficient to locate and direct the mouse pointer. When using gaze pointing with the touch sensitive mouse the position of the mouse pointer has no relevance and consequently, it is not necessary to stir the mouse. The speed benefit for gaze positioning on complex background results from saving the time to stir the mouse.

Fig. 8 shows a typical example of pointing on a complex background and Fig. 9 gives the corresponding plot for the distance over time.

It is worth to mention that for both background conditions the gaze arrives earlier at the target than the mouse. Nevertheless, with modern mouse accelerator techniques the mouse can be temporarily faster than the eye (see slopes in Fig. 9).

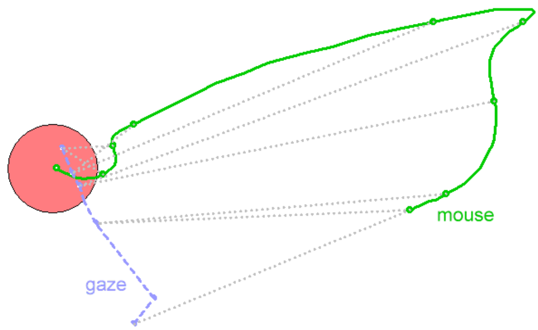


Fig. 8. Gaze (dashed) and mouse trail (solid) for the classical mouse task on complex background. The user stirs the mouse to detect its position by movement.

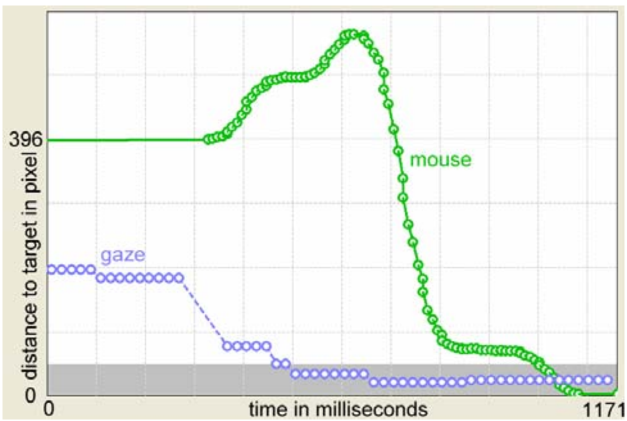


Fig. 9. Gaze (dashed) and mouse trail (solid) for the classical mouse task on complex background as a plot of distance to target over time

4.4 Reduction of Mouse Movements and Workload for the Eyes

It is obvious that gaze positioning saves an enormous amount of mouse moves. It was a surprising observation during the user study that the participants did not move the mouse at all. In case of missing the target, they did not move the mouse. Instead, they lifted the finger and repositioned the mouse pointer again by gaze.

It seems also obvious that using gaze positioning does not put extra workload to the eyes because the eyes move to the target anyway. It is hard to imagine that we can click a target without looking at it. However, a closer look to the data for the classical mouse tasks revealed that in approximately 20% of the recordings the eye gaze did not hit the target, but only moved close to it. In other studies, we observed that, when offering big targets, the eye gaze does not move to the center of the target, but already stops after slightly crossing the edge of the target. However, the end point of the gaze trail in Fig. 10 is definitely too far away from the target edge for an explanation with accuracy and calibration errors. Still in 6.7% of all classical mouse click tasks the eye

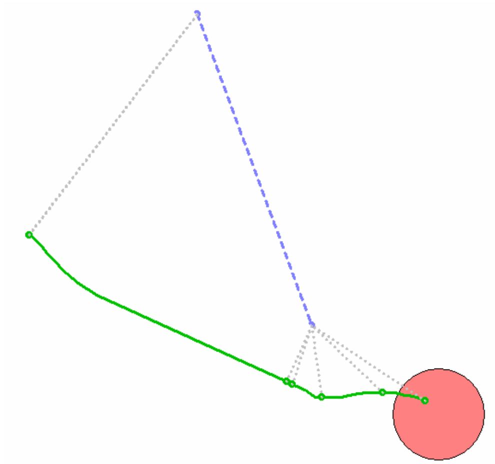


Fig. 10. Gaze (dashed) and mouse trail (solid) for the classical mouse task without background. Here the eye gaze does not hit the target.

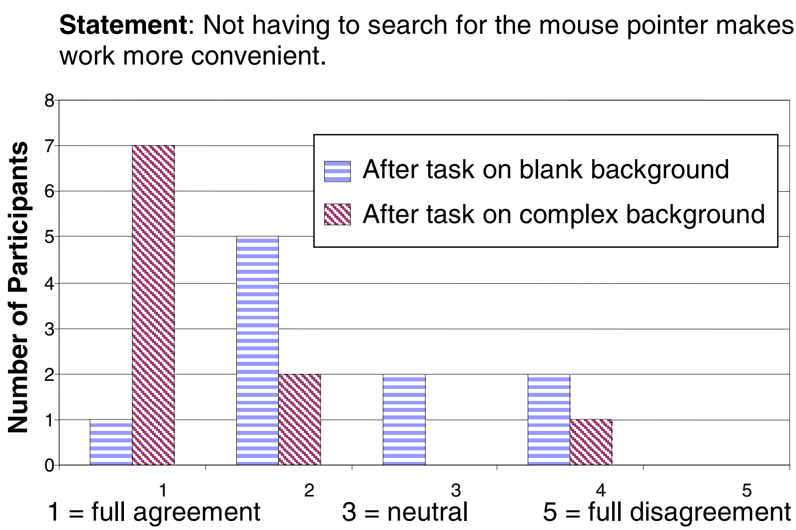


Fig. 11. After the task on the complex background the participants agreed even more that it is convenient not having to search for the mouse pointer

gaze moved less than 90% of the distance from the initial gaze position to the target edge. This leads to the conclusion that it is possible to hit a target without looking at it.

Attempts to reproduce such data intentionally showed that this is indeed possible. The target used is big enough and its red color is striking, which allows easy detection by peripheral vision. The position of the standard mouse pointer is less obvious but can be easily located with peripheral vision because of its movement. As a consequence, it is possible to click on the target without directly looking at it. Nevertheless, such cases do not really contribute to the distance covered by the gaze and therefore it is correct to state that gaze positioning does not require extra gaze activity.

4.5 User Experience

After the two runs on blank background and after the two runs on complex background, we asked the participants to respond to the statement: “It is convenient not having to search for the mouse pointer.” The participants answered using a Likert scale from 1 (completely agree) to 5 (completely disagree). Fig. 11 presents the results. After the tasks on the complex background, agreement shifted towards “completely agree”. In the discussion after the experiments, the majority of the users mentioned that the appearance of the mouse pointer at the where you are looking seems very natural and intuitive.

5 Conclusions

The use of a gaze pointing technique can save a big portion of mouse movements without incurring additional load to the eye. The use of eye gaze pointing will therefore not fatigue the eye muscles more than with traditional mouse pointing. With the common trend to bigger displays, the distances covered by the mouse will become bigger. Gaze positioning in general seems to be a desirable input technology because it reduces the necessary mouse movements and less space is required for the mouse.

Gaze positioning with a touch sensitive mouse can bring a speed benefit compared to the classical mouse, especially when the user is not aware of the pointer position on the screen or when the distance to the target is very far. The reason why the touch sensitive mouse key does not produce the same speed as the gaze key introduced by Ware and Mikaelian [15] lies in the time for checking whether the pointer is in the correct position after the touch. The gaze key triggers an action immediately without feedback on the gaze position. Gaze positioning with a touch-sensitive mouse gives feedback and the possibility to correct the position to achieve the accuracy demanded by the sizes of the GUI elements.

Combining the MAGIC pointing with a touch-sensitive mouse brings several advantages. The MAGIC pointing positions the mouse pointer at a first mouse movement and therefore needs compensation techniques to compensate the movement of the mouse. The MAGIC touch does not need such compensation, as the mouse does not move at the moment of gaze positioning. The concept of a first mouse move in MAGIC pointing implies a timeout and it is not clear how long this time should be. A user has to know whether the timeout expired already to predict the behavior of the system. The concept of touching a key makes the system easier to understand for the user. It also gives more control to the user because it offers also the possibility to reposition the mouse pointer immediately by lifting the finger and touching the key again.

Using the gaze as the only input modality has many severe challenges, but the combination with a touch sensitive mouse key solves some of them – the Midas Touch problem, the feedback issue and issues related to inaccuracy. The combination of an eye-tracker with a touch sensitive mouse key offers a way to utilize the gaze as an additional pointing device without changes in the behavior of the GUI.

There is only a little speed benefit in the case that the user is not aware of the mouse pointer position, but the savings in distance covered by the mouse are enormous. During the interviews, we got very positive feedback for the gaze positioning. The possibility to position the mouse pointer at the position of the gaze just with a little movement of the finger tip feels really pleasant. As a consequence of all, the use of gaze pointing is desirable.

Acknowledgements

The work has been conducted in the context of the research project Embedded Interaction ('Eingebettete Interaktion') and was partly funded by the DFG ('Deutsche Forschungsgemeinschaft').

References

1. Ashdown, M., Oka, K., Sato, Y.: Combining head tracking and mouse input for a GUI on multiple monitors. In: Extended Abstracts on Human Factors in Computing Systems. CHI 2005, pp. 1188–1191. ACM Press, New York (2005)
2. Bolt, R.A.: Gaze-orchestrated dynamic windows. In: Proceedings of the 8th Annual Conference on Computer Graphics and interactive Techniques. SIGGRAPH 1981, pp. 109–119. ACM Press, New York (1981)
3. Card, S.K., English, W.K., Burr, B.J.: Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. Ergonomics 21, 601–613 (1978)
4. Duchowski, A.: Eye Tracking Methodology: Theory & Practice. Springer, Heidelberg (2003)
5. ERICA eye tracker product description, `http://www.eyeresponse.com`
6. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47, 381–391 (1954)
7. Hinckley, K., Sinclair, M.: Touch-sensing input devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI 1999, pp. 223–230. ACM Press, New York (1999)
8. Jacob, R.J.: The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get. ACM Trans. Inf. Syst 9(2), 152–169 (1991)
9. Kumar, M., Paepcke, A., Winograd, T.: EyePoint: practical pointing and selection using gaze and keyboard. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI 2007, pp. 421–430. ACM Press, New York (2007)
10. Majaranta, P., Räihä, K.: Twenty years of eye typing: systems and design issues. In: Proceedings of the, Symposium on Eye Tracking Research & Applications ETRA 2002, pp. 15–22. ACM Press, New York (2002)
11. Majaranta, P., Aula, A., Räihä, K.: Effects of feedback on eye typing with a short dwell time. In: Proceedings of the, Symposium on Eye Tracking Research & Applications ETRA 2004, pp. 139–146. ACM Press, New York (2004)
12. Miniotas, D., Špakov, O., MacKenzie, I.S.: Eye gaze interaction with expanding targets. In: Extended Abstracts on Human Factors in Computing Systems CHI 2004, pp. 1255–1258. ACM Press, New York (2004)

13. Qprox capacitive sensor QT 113 product description, http://www.qprox.com/
14. Salvucci, D.D., Anderson, J.R.: Intelligent gaze-added interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI 2000, pp. 273–280. ACM Press, New York (2000)
15. Ware, C., Mikaelian, H.H.: An evaluation of an eye tracker as a device for computer input. In: Proceedings of the CHI + GI 1987, pp. 183–188. ACM Press, New York (1987)
16. Zhai, S., Morimoto, C., Ihde, S.: Manual and gaze input cascaded (MAGIC) pointing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI 1999, pp. 246–253. ACM Press, New York (1999)
17. Zhang, Q., Imamiya, A., Go, K., Mao, X.: Resolving ambiguities of a gaze and speech interface. In: Proceedings of the, Symposium on Eye Tracking Research & Applications ETRA 2004, pp. 85–92. ACM Press, New York (2004)

Honeycomb: Visual Analysis of Large Scale Social Networks

Frank van Ham[1], Hans-Jörg Schulz[2], and Joan M. Dimicco[1]

[1] IBM TJ Watson Research Center, Cambridge, MA 02142, USA
fvanham@us.ibm.com, joan.dimicco@us.ibm.com
[2] University of Rostock, 18051 Rostock, Germany
hjschulz@informatik.uni-rostock.de

Abstract. The rise in the use of social network sites allows us to collect large amounts of user reported data on social structures and analysis of this data could provide useful insights for many of the social sciences. This analysis is typically the domain of Social Network Analysis, and visualization of these structures often proves invaluable in understanding them. However, currently available visual analysis tools are not very well suited to handle the massive scale of this network data, and often resolve to displaying small ego networks or heavily abstracted networks. In this paper, we present Honeycomb, a visualization tool that is able to deal with much larger scale data (with millions of connections), which we illustrate by using a large scale corporate social networking site as an example. Additionally, we introduce a new probability based network metric to guide users to potentially interesting or anomalous patterns and discuss lessons learned during design and implementation.

1 Introduction

Social networks are structures that model the interconnections between people. Analyzing these social networks provides insight into complex phenomena such as organizational behavior, social organization, and remote collaboration. Typically, social network analysts use a combination of metrics and visualizations to determine central actors, important ties and clusters in relatively small networks of at most a few hundred nodes. With the advent of social network sites on the internet, we now have access to user generated networks that are orders of magnitude larger. For example, Facebook.com currently claims over 100 million active users, which generate billions of network connections. Traditional methods of social network analysis would break down at a fraction of this scale. Non-visual approaches have been proposed to address this problem of scale for general network analysis, such as (hierarchical) clustering, sampling, modeling or approximation. All of these methods reduce the amount of data by creating abstractions of the network. Although these abstractions might give us partial insight into higher level structures, many of these approaches abstract the data to the point where it is very hard to relate high level features to real world phenomena. Another complimentary tool in social network analysis is visualization, which uses interactive graphics to depict the structure of the social network. These diagrams can help scientists understand the over all structure of a particular network and expose

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 429–442, 2009.

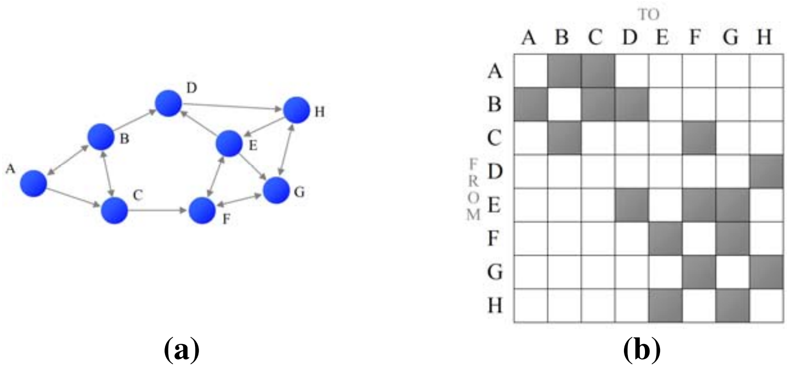


Fig. 1. Social network visualizations. **(a)** A node link representation of a very simple 8 node network **(b)** An adjacency matrix representation of the same network, note that the number of grey cells is equal to the number of links in **(a)**.

patterns that they might not have been aware of previously, or might not have even considered as a possibility. In effect, visualization allows users to perform exploratory analysis on the network and quickly generate and verify hypotheses. One of the strengths of visualizing social networks in particular is that it supports hypothesis generation about an organization's structure and interactions, something of great interest to the organizational behavior and CSCW communities [14]. But typical graph visualizations are limited in their ability to support end users in interpretation of the information and the sheer scale of the networks makes a visual mapping challenging. Social scientists have noted that "…it would seem useful for management to map the social capital ties that are relevant to the various tasks the organization faces. This mapping poses a considerable challenge: from a purely technical point of view it is far easier to map a small number of ego networks than to generate an intelligible socio-centric, whole-network map of a large, complex organisation" [2]. In this paper we present a matrix based tool for exploring very large social networks in their entirety. In contrast with other existing tools, it has the following capabilities:

- **scalability**: Although the social applications we analyzed are accessible to company employees only, and their size pales in comparison to the many millions of members of some online sites, the network we are dealing with in this paper has over 37,000 users and over 400,000 connections. Our tool handles up to a million connections with ease and we have been able to navigate synthetic datasets of up to a few million edges.
- **ability to perform temporal analysis**: Social network analysis is often challenging, but grasping the changes in a social network over time is especially difficult. These dynamic aspects may give us insight in the spread of a network over time and allow us to create better network models and predictions.
- **displaying information that is absent**: While most traditional information visualization tools can show individual information elements and the patterns emanating from them, it is much harder to show the absence of information where it was expected. Concretely, in social network analysis it might be desirable to see between which groups connections are absent or very minimal.

The main contribution of this paper is to illustrate how we can use a matrix based visualization in combination with algorithmic metrics to help social scientists grasp large social network datasets and potentially guide them to interesting features. We illustrate the usefulness of this approach by providing real world samples. In the next section we describe related work in this area. We then give a brief description of our sample dataset, followed by an outline of the visualization method. The main part of this paper describes the visualization and edge metrics as well as some of the insights we have gained in applying them. Finally, we discuss implications for the design of these types of tools.

2 Related Work

The analysis of social networks with statistical and graph theoretical means has had a major influence on social sciences ever since it was first proposed [7]. From the beginning, sociograms and sociomatrices were used as representations of the analyzed social networks. Sociomatrices cross-tabulate a particular connection metric over a number of actors (nodes on the network) and sociograms are traditional node-link representations of the network. Both of these representations have been used as a basis for interactive analysis tools.

The node-link representations in sociograms (Fig. 1a) have been echoed in many social network visualization tools from different research communities. From social network analysis, tools like UCINET/Netdraw and Krackplot [13] offer support for advanced metrics and offer basic node-link visualizations. From the graph drawing community, tools like Visone [4] offer a combination of metrics and advanced layout algorithms. From the visualization/HCI community tools like Vizster [9], the more advanced SocialAction [17] allows users to interactively examine visual representations of (social) networks. All of the above approaches can generate very readable diagrams of social networks of a few thousand nodes at most but have the disadvantage that they cannot display very large networks, especially if these networks are dense. Among the most advanced workbenches for performing social network analysis on large networks are Netminer and Pajek [3] which both offer a large selection of analysis algorithms as well as matrix and node-link visualizations. Users can use the supplied algorithms to extract and visualize meaningful subgraphs from a large network. However, integrating these different views on the network in a single mental model can be a daunting task.

Alternatively, sociomatrices can very easily be transformed into adjacency matrices, where each cell *(i,j)* stores the strength of the connection between actors *i* and *j*. We can then render this matrix by coloring a cell if there is a connection present and, optionally, visually mapping the strength of the connection (Fig. 1b). Note that any colored cells on the diagonal indicate connections of nodes to themselves. Reorderable matrices have been used as an analysis tool since the early 80's and have a number of distinct advantages over node link diagrams, especially if the network is dense [8]. They are impervious to clutter and overlap, scale much better and allow quick verification of the existence of a connection. When the columns are reordered properly [16] they can also be used to identify clusters in the network. On the downside, matrices are not great at visualizing paths between multiple nodes and are not nearly as intuitive as node-link diagrams, which may explain why they are currently under-represented in social network analysis. Recently, interactive adjacency matrices have been advocated for medium scale social network visualization in [10] and a sophisticated hybrid approach has been proposed in [11].

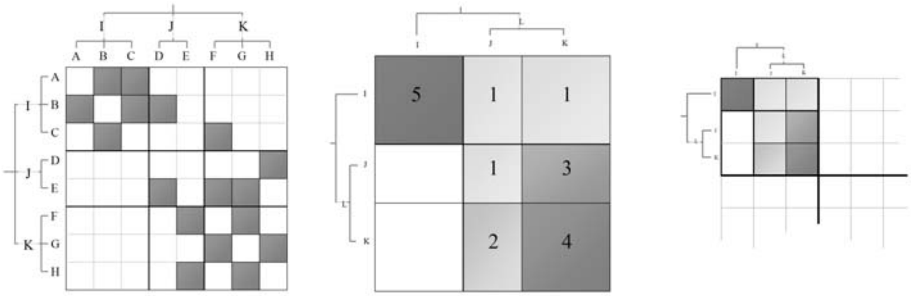
Outside the social networks analysis community, matrix visualizations have been used to analyze different types of networks, however the problem of scalability remains. One potential solution is to group the nodes in the network into clusters and render the aggregated network instead. This can then be repeated at multiple levels of scale, depending on the size of the input graph. This hierarchical grouping can then be used to create a single interactive visualization that allows the user to browse the network at different levels of scale [18, 1, 6]. The major problem with these prior approaches is that, without methods of automatically highlighting potentially interesting or anomalous data points, the user can spend a significant amount of time browsing these representations at multiple levels of scale without learning anything new.

3 Data

Our main data set is taken from a social network site running internally at a large multinational company. The site is an opt-in site that people use for connecting with other employees, through the sharing of photos, lists and events [5]. Part of the process of connecting is to directly link to other employees, as is typical on any social network site. When a user on the site adds someone as a connection, the other user is not required to reciprocate the connection, so employees can connect to anyone inside the company, without the need for the other person to join the site first or reciprocate back afterwards. The data set we use as a sample is a snapshot that was taken mid-July 2008. At that time, 37,000 employees had joined the site and they had formed approximately 300,000 connections.

4 Visualization

Given the scale of our target dataset rendering a single 37,000 by 37,000 matrix is impractical from both a resource and user interaction perspective. We therefore follow the approach taken by [18] and use a hierarchy on the nodeset to reduce the size of the adjacency matrix. Fig. 2 shows how we can collapse an 8 by 8 to a 3 by 3 matrix by using a predefined hierarchy to aggregate cells. The resulting collapsed matrix color codes the total number of edges for each submatrix below (darker colors indicating more edges) and still maintains many of the features of the original network. For example from Fig. 2b we can still easily see that there are no connections from groups J and K to group I and that there are relatively many connection among nodes in group I. As a final step, we normalize the cell sizes so that we can repeat the same process with the collapsed matrix. Note that, even though the relative difference in cell sizes in Fig. 2b accurately portrays the difference in the number of leaf nodes, we found adjacency matrices with irregular cell sizes much harder to comprehend when they grow larger, especially when the difference in cell sizes is large. For that reason we will use a regular subdivision in the samples in this paper, even though our



(a) **(b)** **(c)**

Fig. 2. Collapsing an 8 by 8 adjacency matrix to a smaller 3 by 3 matrix **(a)** original matrix with node hierarchy on both sides **(b)** collapsed version of the matrix with lowest level of the hierarchy eliminated and edge counts aggregated **(c)** This collapsed version itself forms a small section of a higher level adjacency matrix

visualization tool allows users to choose either option. The actual hierarchy used to drive the above process is variable and depends on the interest of the users of the visualization. In the samples in this paper we have used two distinct hierarchies. One uses the management hierarchy to correlate connection behavior with organizational structure, while the other one uses a geographical hierarchy based on the user's working location (i.e. continent - country - state - city - building) to correlate connection behavior with geographical location. In practice, we can also use different hierarchies or construct a hierarchy ourselves by using other node attribute information.

In terms of interactivity our tool is very similar to its predecessor described in [18] but it is more memory efficient and allows for pluggable metrics. The user is initially presented with an adjacency matrix that displays connections at the highest level of abstraction (e.g. in the case of the geographical hierarchy connections between employees in different continents). By left clicking on a cell *(X,Y)* the user can indicate he or she wants to examine that particular connection in more detail and the visualization then displays the matrix that shows the connections between the direct children of *X* and *Y*. A simple right click brings the user back to the cell he or she came from. The transition between these two matrices is animated to help the user understand the relationship between the two representations. Dynamic labels help the user understand what relation they are looking at and a popup menu provides details on demand.

To deal with the issue of visual scalability we have used the hierarchy to reduce the matrix to a more manageable size. Computational scalability is obtained by using a semi-external memory approach, that is, we keep the entire nodeset and the hierarchy of the network in RAM while a relational database stores the actual connections between the nodes in the network. When a user requests a higher level view of the network, aggregation of edges in the database is done on the fly using a fast lookup algorithm. Our current prototype is implemented in Java and uses OpenGL for graphics output. We have successfully loaded and navigated synthetic graphs up to 5 million edges using only 200MB of RAM.

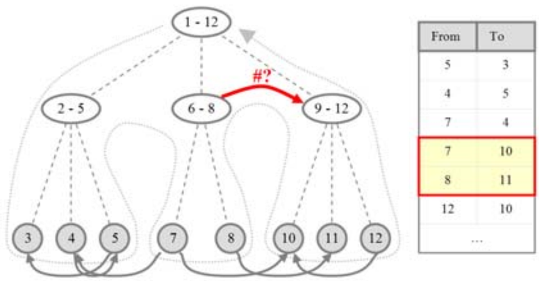


Fig. 3. Schematic representation of the network (grey nodes and curved connections) and the aggregation hierarchy. By numbering the nodes in the aggregation hierarchy in a depth first manner (dotted line) and keeping track of the minimum and maximum values encountered during this traversal we can determine the number of edges connecting groups (6-8) and (9-12) by running the query: SELECT COUNT * FROM EDGES WHERE 6 'From' 8 AND 9 'To' 12;

5 Metrics

Since every edge in the network is represented by a cell in the matrix, we can use the space in this cell to communicate information about the connection it represents. Although some structural edge metrics exist [15] they are mainly focused at identifying individual edges that connect different communities and rarely take edge weights into account. This means that although they are able to identify communities at the level of individual actors, running them on aggregated data produces less reliable results and in some cases might be too computationally expensive. Previous approaches [18, 1] have exclusively mapped the number of aggregated edges, or the *connection count*. We can express the connection count *CC* for an edge *(X,Y)* as:

$$CC(X,Y) = \|(x,y) \in E : x \in desc(X)\ and\ y \in desc(Y)\|$$

with *desc(X)* indicating the set of descendant leaves of a node in the hierarchy. To be able to compute connection counts between groups efficiently we employ a special numbering scheme on the nodes in the network. We traverse the hierarchy on the nodes in the network in a depth first traversal starting at the root, and incrementally number the encountered nodes. For each node in the hierarchy we keep track of the minimum and maximum value encountered for that node. We can then use these Depth First Search numbers to query the edgelist stored in a database whenever we need the total number of connections between two groups (see Fig. 3 for details).

Combined with caching of the top level matrices this keeps our tool memory-efficient and fast. Alternatively, a solution where all edge counts are precomputed and stored in a disk based index can be used [6], although this makes it harder to switch between multiple hierarchies.

Although the connection count is probably the most straightforward metric one can think of, it has a number of drawbacks. Firstly, at higher abstraction levels the adjacency matrices grow increasingly dense, making it harder to detect patterns based on the presence of connections. The use of a color scale to indicate the number of connections

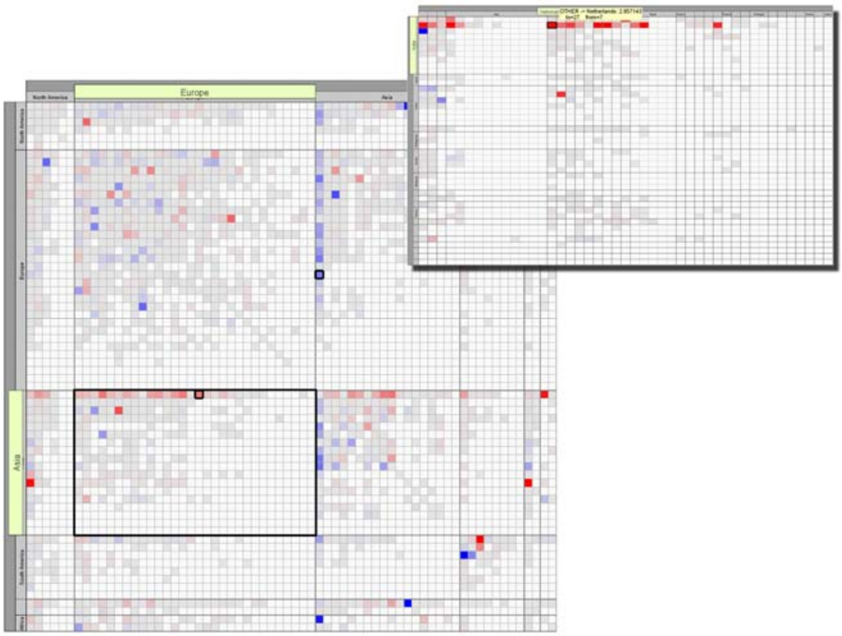


Fig. 4. Shows the difference between incoming and outgoing connections, with India showing up as an outlier. The horizontal red band and vertical blue bands indicate there are many unreciprocated connections from Indian employees. The inset shows a zoomed view of the outlined cell with the connections between Asia and Europe, the red horizontal band represents connections from mobile employees (with location OTHER) in India to people in Europe.

might help somewhat, but the large number of connections on the diagonal in social networks often drown out subtler off-diagonal patterns.

Finally, one key problem in all previous matrix-oriented visualization approaches is that it is very hard for a user to identify what connection patterns are potentially interesting to examine further. In many instances the user is simply presented with a matrix of aggregated connection counts and is expected to identify anomalous patterns by interactively browsing the data. Although the user can now potentially look at interesting patterns, this ability is useless if the user has to spend a large amount of time inspecting abstract matrices at different levels of detail to find these interesting patterns. There is a definitive need for edge-oriented metrics that go beyond simple edge counts, yet can still be computed efficiently. In the next two sections we will propose two of these metrics and use our sample dataset to highlight their usefulness.

5.1 Asymmetry

Since connection behavior does not need to be reciprocated (i.e. X can connect to Y but Y does not have to connect back to X) it might be interesting to look for patterns that involve asymmetric connections. For example, for each edge (X,Y) we can define the asymmetry as

$$ASYM(X,Y) = CC(X,Y)/CC(Y,X)$$

Fig. 4 shows the asymmetry in connections, using the geographical hierarchy to aggregate. Small imbalances in asymmetries can distract from detecting large imbalances, so we only plot asymmetry values above a small cutoff. One pattern that immediately emerges from this image is the large horizontal (red) and vertical (blue) bars in the row and column representing India. This indicates India overall has more connections going out to another country than it has coming back. The team that had developed the social networking application was aware of the fact that a lot of US users were getting friend requests from users in India they did not know personally, but did not know that this pattern persisted over different countries. To determine where this behavior was coming from, we drilled down on the connections from India to Europe (inset) and observed that most of the people responsible for these asymmetric connections report their work location as 'OTHER', which in practice means 'non-traditional office'. Employees working in non-traditional offices (typically drop-in stations or from home) may be using the social network site with a particular focus on meeting new people, because they do not have regular face-to-face contact with co-workers.

In a separate survey of 2000 users of this social network service, we asked about different reasons why they were motivated to use the site. Comparing the responses between countries, users in India reported to a significantly higher degree than other countries that they were using the site for getting to know people they would not otherwise meet at the company, using the site to find experts and using the site to discover people with similar interests. All three of these activities involve reaching out and connecting across organizational and cultural boundaries, and these motivations offer a partial explanation of why this particular cluster of outward links may exist. Asymmetry-type metrics can also be useful in wider contexts, especially if the connection's weights can be significantly different for each direction. Examples of such networks include trade, financial networks or communication networks.

5.2 Deviation from Expected

One of the major disadvantages of rendering absolute metrics, whether they are absolute connection counts or the absolute number of unreciprocated connections, is that groups with the largest number of users are very likely to have the highest number of connections. Indeed absolute connection counts allow us to make observations such as 'Users in the US have 3024 connections to users in India' but there is no way to estimate how valuable that observation actually is. Is 3024 connections a lot, given the characteristics of the network and the size of the two countries, or did we expect to see more?

One way of estimating the usefulness of an observation is to try and determine what part of this observation might be explained by pure chance. In other words, if we had distributed the E edges in the graph completely arbitrarily, how many edges can we expect to fall between the US and India? This problem is similar to a chi-square analysis of a set of observations, using the matrix M of connection counts as the contingency matrix. Note that the total number of connections in a row X of M is equal to the total number of edges X_{out} that have their startpoint in X, and the total number of connections in a column Y of M is equal to the number of edges Y_{in} that have their endpoint in Y. The probability of having an edge connecting *X* and *Y* if the data were randomly distributed (i.e. the choice of start and endpoint of an edge are independent) is then equal to $P(X,Y) = \frac{X_{out}}{E} * \frac{Y_{in}}{E}$ with expectation:

$$EXP(X,Y) = E * P(X,Y) = \frac{X_{out} \cdot Y_{in}}{E}$$

We can then compute the chi square metric for the total adjacency matrix, which tells us if there are significant correlations between the clusters we have defined on the nodeset. If that is the case (and in our particular case it was) we can look at the value of the individual cells and try to establish where these correlations might originate from.

As an alternative derivation, we can look at the problem of obtaining a graph with exactly *E* edges, of which *k* connect nodes *X* and *Y* as a discrete probability problem: Suppose we have a 2 dimensional *R* by *C* grid *M* (our adjacency matrix) in which we have to assign edges to cells. In order to be able to generate all possible assignments we have a jar of $R \times C - W$ black and *W* white edges. White edges are to be placed in a specially designated subgrid of *M* of size *X* by *Y*. Black edges have to be placed outside of this subgrid. If we sample *E* edges from the jar *without replacement* and the probability of pulling a white edge from the jar is *P(X,Y)* as defined above, what is the probability that exactly *k* edges will be white? This problem can be modeled as a hypergeometric distribution which has the same expectation value $E * \frac{RC*P(X,Y)}{RC} = EXP(X,Y)$ and variance:

$$VAR(X,Y) = EXP(X,Y) * (1 - P(X,Y)) * \frac{R*C-E}{R*C-1}$$

The remaining free variables in the above expression are the values for *R* and *C*. These determine in relation to what part of the matrix we compute the variance and expectancy. Typically $R \times C = N^2$ if we want to look at the distribution of the edges over the whole graph, but we can also compute the expected values given the characteristics of a particular subcell (typically an ancestor of the cell we are looking at) in the matrix. The metric outlined above is in effect a more advanced version of a basic density metric, which estimates the number of expected edges in a matrix cell solely based on cell size, i.e. $P_{density}(X,Y) = \frac{X}{R}\frac{Y}{C}$ The above metric instead uses the in- and out degree distribution of the graph under consideration to estimate the number of expected edges, similar to the p_1 network model [12].

A good test case for our deviation metric would be to see if we can correlate the company's business units to their respective geographies. As an example, we would expect employees of the German business unit to have most of their contacts with employees in Germany. Other business units span different geographies however, and our metric should also point this out. Although up to now we have only shown matrices that use the same hierarchy on both sides, the visualization can also support different hierarchies on the two sides of the matrix. Fig. 5a shows part of the matrix that displays the links from the managerial hierarchy (rows) to the geographical hierarchy (columns) with manual annotations. Color represents the deviation from the expected value with blue cells indicating fewer connections than expected and red cells indicating more connections than expected. The size of the glyph in each cell gives an indication of the significance of the deviation from the expected number as cells that

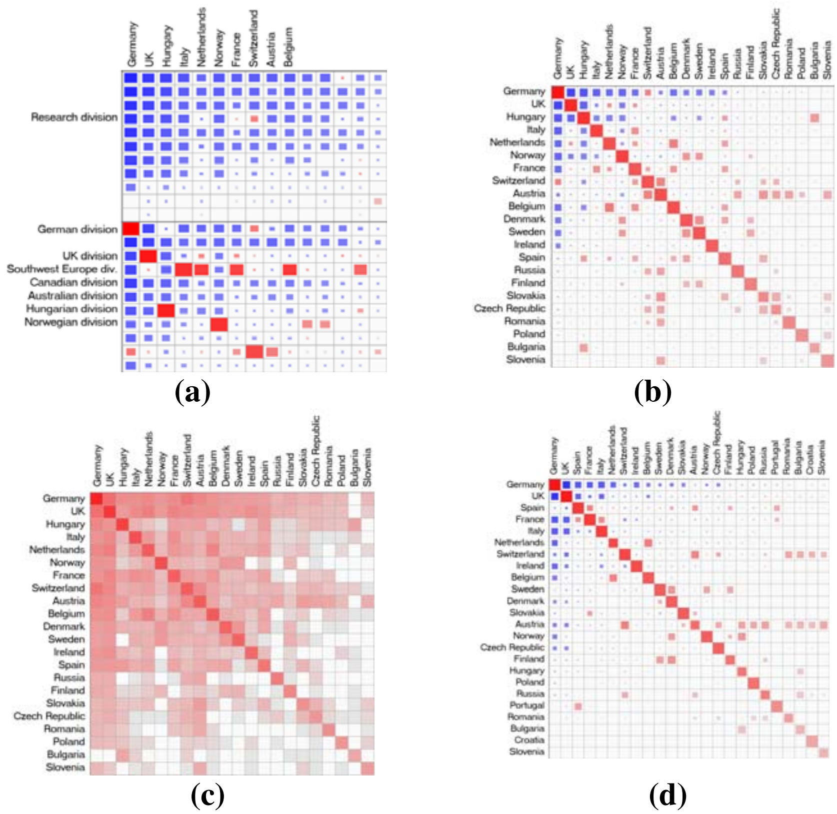


Fig. 5. Mapping the difference between the observed number of connections and the number of connections we would expect if the distribution of the edges were random. Red cells indicate a higher number of connections than expected, blue cells indicate a less than expected number of connections. The size of each square indicates the significance of the deviation. **(a)** connections between divisions and countries **(b)** connections among different European countries. **(c)** same view as **(b)** but using the connection count metric. **(d)** filesharing activity between different European countries.

deviate only in a small amount from the expected value are within the bounds of normality and should not be visually highlighted. The expected patterns show up very clearly, but we also observed connections that might not have been so obvious initially, with employees in research connecting to employees in Switzerland. This is most likely due to a large corporate research laboratory in Switzerland, but the ability to isolate this fact is encouraging.

We also applied the same metric to look at connections within Europe (Fig. 5b) and found a number of interesting patterns. One fairly obvious one is the tendency to connect to people in the same country, indicated by the bright red diagonal. Other observed patterns have to do with language and geographical proximity (for example, Switzerland connects more than expected to Germany, Austria and France) or with organizational structures (Sweden, Norway, Finland and Denmark are all part of a single organizational structure). One pattern we did not expect to find were the larger than expected number of connections between Austria and many of the countries in Eastern Europe. We were subsequently able to verify that the company, like many other multi-national firms, handles a significant part of its Eastern European business from Vienna.

As a second check on the validity of these observations we explored a different set of connections. The company also uses internal online document sharing services where people can upload a single document and share it with multiple colleagues, which saves the user from having to mass-mail the document. We visualized a set of 75,000 document sharing relationships, where two actors are related if one of them has shared a document with the other. Fig. 5d shows document sharing relationships over different countries in Europe. Again, Austria appears to connect significantly more to Eastern Europe than other countries. The asymmetry is expected since sharing a file is an activity that is usually not reciprocated by the receiving user. Still, the rough similarities between the images in Fig. 5b and Fig. 5d verify that the patterns seen in both the social network service and the file-sharing service reflect meaningful organizational patterns in the company.

The practice of computing statistics for different network decompositions is an established procedure in the social network analysis community. However, the results are often displayed in simple statistical tables, which makes interpretation not very intuitive. The added value here is that we can display the resulting deviations from the expected values in context, use a comprehensive display to look for patterns and easily compare them with other network metrics at different levels of granularity to investigate. Within the context of a corporation, social network analysis is often touted as a mechanism for revealing the invisible power structures outside of the management hierarchy. One of the unique strengths of the visualization is that the structure of the social network is overlaid onto the formal hierarchy of the company (divisional or geographical) allowing comparison of the formal and informal hierarchies in the company.

5.3 Time Varying Networks

Another distinct advantage that matrix based visualizations have over node link based ones is their ability to deal with data that varies over time. Many other approaches have used animated node-link diagrams to display changing networks but this approach has a major drawback. If node positions are kept static over time, the topology of the network will not accurately reflect the data at that point in time. Yet if we animate the node positions over time, the user is confronted with an animation that both changes the geometry and the connectivity pattern simultaneously, which in practice is very hard to understand. In matrix visualizations, each edge (whether realized or not) is already allocated a section of screen space, which allows us to animate the growth of connections in a stable visual representation. Depending on user interest we could for example choose to display the cumulative edge count over time, a heatmap of the new connections made in the last month or highlight the most volatile connections. In our prototype we have implemented a slider that lets us visualize any time interval between the launch of the site and the date of the snapshot and apply any of the metrics to the connections that were made in that time interval.

6 Lessons Learned

In this paper we have shown that interactive adjacency matrices are a viable alternative and, in many cases, preferable to node-link diagrams when it comes to analyzing large scale social networks. Large scale formal evaluation on these types of tools is difficult because exploratory analysis inherently involves tasks which are ill-defined. That is, the user does not know what they are looking for exactly (yet). We have illustrated their usefulness by performing exploratory analysis on real-world large social networks obtained from an internal company research project. This section discusses some lessons learned during design and implementation.

6.1 Integrate Multiple Metrics in a Single Mental Model

Scientists in the area of social network analysis have come up with a number of useful metrics to analyze the connection patterns in social networks. However, understanding the structure of a social network requires analysts to understand how the different metrics in the social networks interrelate. Many existing tools allow the user to run an analysis and then visualize the results [3] but integrating all these separate perspectives into a single, coherent mental model is often left up to the user. Having a single consistent mental model of a complex data structure allows user to incrementally build up knowledge by allowing easier correlation of newly observed facts with previously observed facts. As an example, we observed that Austria has strong ties to other Eastern European countries. When we looked to verify this observation using a different network of the same individuals, we could easily identify a similar pattern because connections from Austria were visualization in the same row and formed the same horizontal pattern. Had we tried to do this analysis with node-link diagrams the layouts of the two networks would have varied wildly, making direct comparison much harder.

6.2 Use Concrete Hierarchies to Drive the Analysis

One of the prerequisites of this visualization technique is a suitable hierarchy on the set of nodes in the network. Here, we have used readily available hierarchical decompositions such as the management hierarchy. This allows us to interpret connections between higher level aggregations directly, because they relate to a concrete and meaningful grouping. Previous approaches [1, 6] have used network clustering or matching algorithms to automatically generate a hierarchy on a network when none was available. Although these approaches offer the similar scalability and corresponding interaction model, interpretation of the resulting visualization is often problematic because it is much harder to put a meaningful label on an individual cluster. Recursive clustering algorithms are more problematic as we do not know what the common attributes of nodes in a group might be. This makes it much harder to ground the interpretation of the resulting structure in existing knowledge. If no natural hierarchy is available for the network under consideration we recommend using ordinal node attributes to create a hierarchical partitioning if the number of categories. That is, we can recursively aggregate nodes in the network by grouping them if they have the same value for a particular attribute. This approach is in a sense a hierarchical version of the idea behind PivotGraph [19].

6.3 Absent Data Is Also Information

One of the chief advantages matrices have over node-link diagrams is their ability to highlight missing connections. For denser networks, the absence of a connection where one was expected might be just as informative as the presence of a connection where it was not expected. Matrix views explicitly represent absent connections as an empty cell, which allows us to look for patterns that involve empty cells. Sample patterns may include sparsely filled rows or columns in an otherwise dense matrix or (almost) empty aggregate cells. Similarly, computing the difference between a statistically expected value and the observed value allows us to examine a higher order of absence: given a suitable baseline model we can determine how many connections we would have expected and highlight the difference. Many current information techniques only focus on highlighting patterns that are present, often neglecting patterns that are notably absent.

7 Conclusions and Future Work

We have shown how matrix based visualization tools have significant advantages over node link diagrams when it comes to analyzing very large networks in general and large social networks in specific. They deal better with denser networks, offer ample screenspace to display metrics for each edge and can more easily display change over time. We have also shown how we can use statistical measures to estimate the usefulness of a particular observation. The one major disadvantage that matrices have over traditional node-link diagrams is that they are unable to display structural features such as shortest or multi-segments paths in the network in an intuitive manner. However, aggregating networks greatly increases their density and in most cases the high level overviews generated here are too dense to display in node-link form.

The approach outlined here opens up a number of viable routes for further work. Having a common, consistent 'data space' available may offer possibilities for collaborative analytics where multiple users analyze a potentially large network dataset. One can think of highlighting areas that have already been examined to guide users to previously unexplored sections of the matrix or allowing users to collaboratively annotate the visualization for example. In terms of metrics, our expectancy metric provides promising initial results, but still overemphasizes the tendency of actors to connect within their own community. Better random network models that take this into account may provide a better assessment of what datapoints (e.g. connections) are interesting and which are not. Alternatively, secondary network data (for example, obtained from email exchanges or other known organizational patterns) might be overlaid to see where users' connections deviate. In terms of visualization one could imagine implementing a flexible hierarchy based on node attributes, where users can quickly try out partitions to see where correlations with structure might lie. Also, our current visualization does not allow users to view connections between groups at different levels in the hierarchy simultaneously and asymmetric zooming might be helpful here. Finally, application of this type of visualization techniques to graphs from different application areas may give researchers insight in much larger networks than previously possible.

References

1. Abello, J., van Ham, F.: Matrix zoom: A visual interface to semi-external graphs. In: Proceedings of the IEEE InfoVis 2004, pp. 183–190 (2004)
2. Adler, P.S., Kwon, S.-W.: Social capital: Prospects for a new concept. The Academy of Management Review 27(1), 17 (2002)
3. Batagelj, V., Mrvar, A.: Pajek - program for large network analysis. Connections 21, 47–57 (1998)
4. Brandes, U., Wagner, D.: visone – analysis and visualization of social networks. In: Jünger, M., Mutzel, P. (eds.) Graph Drawing Software, pp. 321–340. Springer, Heidelberg (2004)
5. DiMicco, J.M., Millen, D.R., Geyer, W., Dugan, C., Brownholtz, B., Muller, M.: Motivations for social networking at work. In: Proceedings CSCW 2008 (2008)
6. Elmqvist, N., Do, T.-N., Goodell, H., Henry, N., Fekete, J.-D.: Zame: Interactive large-scale graph visualization. In: Proceedings of the IEEE Pacific Visualization Symposium 2008, pp. 215–222 (2008)
7. Freeman, L.C.: The Development of Social Network Analysis: A Study in the Sociology of Science. Empirical Press (2004)
8. Ghoniem, M., Fekete, J.-D., Castagliola, P.: On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. Information Visualization 4(2), 114–135 (2005)
9. Heer, J., Boyd, D.: Vizster: Visualizing online social networks. In: Proceedings of the IEEE InfoVis 2005, pp. 32–39 (2005)
10. Henry, N., Fekete, J.-D.: Matlink: Enhanced matrix visualization for analyzing social networks. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4663, pp. 288–302. Springer, Heidelberg (2007)
11. Henry, N., Fekete, J.-D., McGuffin, M.J.: Nodetrix: a hybrid visualization. IEEE Transactions on Visualization and Computer Graphics 13(6), 1302–1309 (2007)
12. Holland, P., Leinhardt, S.: An exponential family of probability distributions for directed graphs. Journal of the American Statistical Association 77, 33–50 (1981)
13. Krackhardt, D., Blythe, J., McGrath, C.: Krackplot 3.0: An improved network drawing program. Connections 17(2), 53–55 (1994)
14. McDonald, D., Fisher, S.F.D.: Workshop on social networks for design and analysis: Using network information in CSCW. In: Proceedings of CSCW 2004 (2004)
15. Melançon, G., Sallaberry, A.: Edge metrics for visual graph analytics: A comparative study. In: 12th International Conference on Information Visualisation (IV 2008), pp. 610–615 (2008)
16. Mueller, C., Martin, B., Lumsdaine, A.: A comparison of vertex ordering algorithms for large graph visualization. In: International Asia-Pacific Symposium on Visualization, pp. 141–148 (2007)
17. Perer, A., Shneiderman, B.: Integrating statistics and visualization: case studies of gaining clarity during exploratory data analysis. In: Proceedings of CHI 2008, pp. 265–274 (2008)
18. van Ham, F.: Using multilevel call matrices in large software projects. In: Proceedings of the IEEE Symposium on Information Visualization, pp. 227–232 (2003)
19. Wattenberg, M.: Visual exploration of multivariate graphs. In: Proceedings CHI 2006, pp. 811–819. ACM Press, New York (2006)

Simulating Social Networks of Online Communities: Simulation as a Method for Sociability Design

Chee Siang Ang[1] and Panayiotis Zaphiris[2]

[1] Centre for HCI Design, City University London
jimmybbq@gmail.com
[2] Department of Multimedia and Graphics Art, Cyprus University of Technology

Abstract. We propose the use of social simulations to study and support the design of online communities. In this paper, we developed an Agent-Based Model (ABM) to simulate and study the formation of social networks in a Massively Multiplayer Online Role Playing Game (MMORPG) guild community. We first analyzed the activities and the social network (who-interacts-with-whom) of an existing guild community to identify its interaction patterns and characteristics. Then, based on the empirical results, we derived and formalized the interaction rules, which were implemented in our simulation. Using the simulation, we reproduced the observed social network of the guild community as a means of validation. The simulation was then used to examine how various parameters of the community (e.g. the level of activity, the number of neighbors of each agent, etc) could potentially influence the characteristic of the social networks.

Keywords: Agent-based model, online community, simulation, sociability, social network.

1 Introduction

Developing online communities involves usability design, which supports consistent, controllable and predictable Human-Computer Interaction (HCI), as well as sociability design, which focuses on human-human interaction (i.e. social interaction) and social policies [24]. The majority of research on sociability design has so far been revolved around conventional HCI methods such as query-based techniques (e.g. questionnaires, interviews), observation and content analysis to identify sociability requirements of online communities. For instance, using these methods, several studies have already analyzed the content of messages that people post in online communities [26, 13, 11, 4, 27, 25, 14, 22].

When developing online communities, we believe that it is useful to be able to evaluate and compare design alternatives. For instance, how different social policies could potentially lead to the expansion of the community; how specific interface design elements could sustain the growth of the online social networks. We might also want to find out how certain changes in an existing community could affect its future development.

One possible way of sociability design evaluation and comparison is through simulations. Suppose we have an online game community and we notice that there are a lot

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 443–456, 2009.

of off-topic discussions going on. We want to introduce a new policy/new design that could reduce these discussions but we are unsure whether this will be beneficial or detrimental to the community. Often it is impractical to run large-scale experiments on such a massive online community. Therefore, we are proposing a way of testing such situations through the use of simulations. In this case, we could develop a social simulation that mimics the interaction behavior of the existing community and conduct "virtual" experiment using the simulation.

Our aim was to develop an Agent Based Model (ABM) to simulate and study social networks of online communities and investigate the usefulness of simulations in sociability design. As a case study, we chose to analyze a Massively Multiplayer Online Role Playing Game (MMORPG) guild community (a virtual association with formalized membership and rank assignments that encourage participation). To achieve this, we:

- observed and analyzed the social interaction in an existing guild community
- developed an ABM which was founded on this empirical observation to simulate the community
- conducted "virtual experiments" with the simulation to demonstrate its usefulness in sociability design.

2 Agent Based Simulation

In term of social simulations [10], an Agent Based Model (ABM) [33] is a computational model for studying the process of the social system as a whole through simulating the actions and interactions of autonomous individuals, known as agents. Agents can have data-gathering and decision making abilities, sometimes equipped with sophisticated learning capabilities and they can be adaptive to the environment. Traditionally, sociologists have understood social life as a system of institutions and norms that shape individual behavior from top down. When dealing with linear systems, the behavior of the whole system corresponds exactly with the sum of its constituting parts (such as multiple regression models). In non-linear complex systems (such as online communities), even if the observer has a good understanding of how each part works, it will not be possible to understand the system as a whole [28]. Instead of starting from the system as a whole and decompose it (top-down), it will be more fruitful to start from its constitutive parts (bottom-up). In such systems, coherent system behaviors not defined *a priori* may spontaneously emerge from the aggregate dynamics of its components (in our case agents).

ABM simulation can work like a virtual laboratory through which the researcher can test hypotheses, manipulating variables and constraints of the model and observe the outcome. By generating many simulation runs with different conditions, large data sets of online communities can be accumulated. Research using ABMs has demonstrated the ability to "grow" societies "*in silica*" [9]. Various social phenomena have been studied using ABM techniques, including the simulation of digital markets [18], violence and conflict modeling [8, 12], trust and cooperation [19], archaeological simulation [6], ecosystem management [2], etc.

The main strength of ABMs lies in its focus on simulating relationships between agents. Therefore, not surprisingly, ABMs have been applied to simulate interaction between agents in social networks. In most cases, ABMs are used to study two situations in social networks. Firstly, simulations can be developed to examine agents' interaction and the aggregate interaction dynamics within specific topology of social networks. Wilhite [32] for instance examined trading interaction, i.e. search, negotiation and exchange, within four types of social network topology: global network, local disconnected network, local connected network and small-world network. Instead of exploring the process of network formation, he was interested in understanding how different types of social network configurations affect trading behaviors.

More common in the research of social network simulation is the generation of social networks that arise from agents' interactions. For example, some work has been carried out to explore the network formation based on the simulation of innovation creation and knowledge diffusion in industry [1, 11, 23]. In another study, simulations were used to replicate the structural characteristics of an online discussion forum, and evaluated how different types of expert ranking algorithms may perform in communities with different characteristics [34, 35]. ABMs have also been used increasingly in studying network evolution [5, 20]. The study reported in this paper focused on social network generation. The ABM described in this paper simulated the interaction behaviors of individual members of a MMORPG community to understand the dynamics of the emerging social network in that community. Through the simulation we explored how various interventions could affect community performance.

3 Methods

The study was carried out in four phases: modeling process, simulation development, parameter setting and validation, as well as virtual experiments.

3.1 Modelling Process

At this stage, we analyzed a guild community of World of Warcraft (WoW) [3], a MMORPG, to understand the interaction patterns and social networks of the community. We collected player chat interaction data from a relatively large guild (with 76 members at the time of observation). Using the in-game chat-log function, we kept a record of guild members' chat activities throughout the study. A total of 1944 guild messages were collected in 30 hours of observation.

We carried out qualitative analysis of the chat interaction patterns in the guild community in order to formulate the interaction rules of the agents in the simulation. We categorized the messages using thematic and content analysis [16], focusing on three interaction types: "help provision", "help seeking" and "friendly chat". Then, we tabulated the messages from the empirical data into socio-matrices for social network analysis by identifying "who talked to whom" relationships from the chat log. Three social-matrices representing help seeking, help provision and friendly chat interaction, as well as a social-matrix representing the aggregate interaction of the three interaction types were generated for quantitative Social Network Analysis

(SNA) [29]. Based on the results from this empirical study, we derived the “rules” of interaction patterns and formalized them into an ABM. Figure 1 shows the conceptual framework of the simulation. Agents interact with each other, resulting in the formation of ties and thus social networks (simulation output). The interaction is defined by five sets of interaction rules:

- initiation rules that determine whether an agent will initiate an interaction in a simulation time cycle
- interaction type decision rules that decide what types of interaction an agent will initiate
- help seeking, help provision and friendly chat rules: the behavior of the agents engaging in these three interaction types.

The simulation parameters (simulation input) are extrinsic factors that have an impact on all of the agents and therefore their interaction. In the context of sociability design, we can see parameters as external factors imposed on the community that are caused by design decisions or policy interventions. We explored the following five parameters and their influence on the formation of social networks in the community:

- **Social budget**: how does the maximum number of immediate friends an agent can have affect the social network of the whole community?
- **Activeness factor:** the level of activity; what will happen to the social network if all the agents of the community become more active in initiating an interaction?
- **Cohesion factor**: what if the agents tend to interact mostly with their friends in the social network and not with anyone outside their friends network?
- **Help threshold:** the ability of the agents to help someone who needs help
- **Chat threshold:** the tendency of the agents to engage in chatting interactions.

These parameters were derived from literature [4, 11, 13] in order to explore the most prominent factors that affect online social networks. All parameters were scaled from 0 to 10. Note that these are relative values for the purpose of comparison. We believe that understanding the influence of these parameters on the social network will answer practical design questions about online communities. We explain later how we used practical scenarios to explore this further.

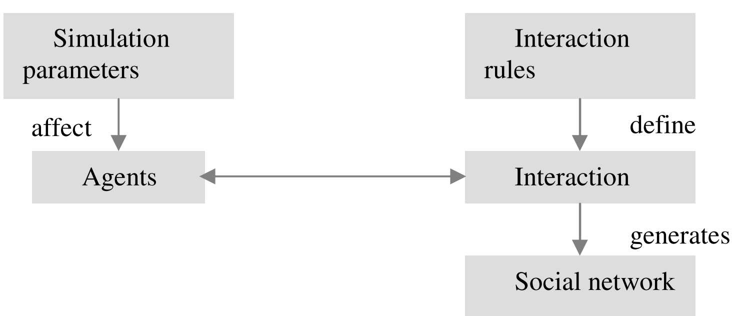


Fig. 1. The conceptual framework for the simulation

3.2 Simulation Development

Figure 2 shows a snapshot of our simulation developed using Netlogo cross-platform multi-agent programmable modeling environment [31]. This snapshot shows the formation of a social network.

There are four major types of components in the interface:

- The parameters (simulation inputs): researchers can tweak the parameters of the simulation to control the process of the simulation.
- The network visualization (simulation outputs): researchers can examine the visual patterns of the social network being created; the size of the node indicates the centrality and the color brightness indicates the size of the friend network.
- Network analysis result displays (simulation outputs): includes a general network measure report, density, in- and out-degree centralization and reciprocity [29]. These results can be automatically calculated when the social network evolves or can be calculated at any point of the simulation.
- The control panel: allows researchers to start, pause and stop the simulation; to xport the social network matrices generated by the simulation for further in depth analysis with standard SNA software packages such as UCINET (http://www.analytictech.com/).

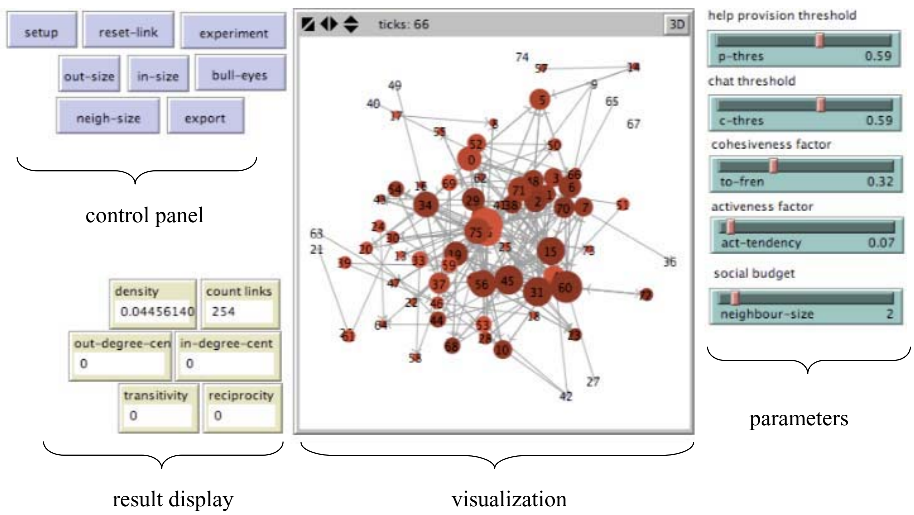


Fig. 2. The implementation of the simulation

3.3 Parameter Setting/Calibration and Validation

Once the simulation was programmed and verified (bugs removed) the next step was to set the initial values of the parameters of the social network to be simulated and to validate the simulation.

Much recent work on modeling of complex networks in social, biological and technological domains has focused on replicating one or more aggregate characteristics of real world networks, such as degree centrality, clustering, path lengths [21].

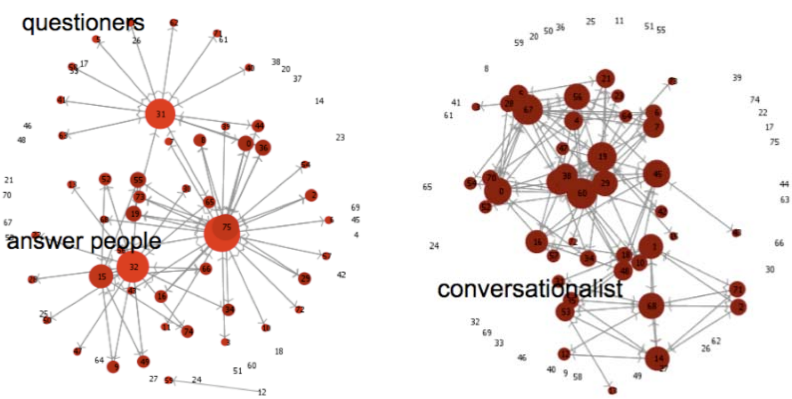
We followed this research tradition. We validated our model by producing social networks that match the aggregate characteristics of the social network of the observed WoW guild community. To achieve this, the initial values of the parameters had to be set. It was an iterative systematic process in which we compared four SNA measures (density, in- and out- degree centralization and reciprocity [29]) of simulated and observed (the aggregate network of the three interaction types as mentioned earlier) networks. We would like to emphasize that at this stage of the study, we only attempted to model this specific guild community we observed, with no intention to generalize it to other online communities.

At each iteration, we adjusted the parameters, generated 30 social networks and calculated the means of the SNA measures. Using confidence interval validation techniques, we examined the error/discrepancy between the simulated and the empirical results. We re-adjusted the parameters and repeated the same process until the error was within the tolerate level (refer table 1). This iterative process not only helped us set the initial values of the parameters, but also augmented our understanding of the parameters and the underlying structure of the model.

The validity of our simulation is demonstrated by the fact that at the final iteration of the parameter setting, the worse case error of all SNA measures is lower than the tolerate error (table 1). Furthermore, once the parameters were set and the simulation was validated to match empirical data, we demonstrated that the simulation was able to reproduce social networks observed and described in past research [30]. For

Table 1. Confident interval validation of the simulation

	density	Out-degree	In-degree	Reciprocity
Observed Mean	0.0449	0.2247	0.1977	0.6230
Tolerance Error	0.0050	0.0500	0.0500	0.0500
Simulated mean	0.0448	0.2416	0.1836	0.6533
Standard Deviation	0.0006	0.0351	0.0372	0.0324
Worst case error	0.0004	0.0300	0.0280	0.0425



(a) (b)

The bigger the size of the note, the higher the in-degree centrality The darker the color the bigger the local friend network

Fig. 3. (a) A typical simulation run with only "seek help" interaction activated. (b)A typical simulation run with only "friendly chat" action activated.

instance, a star structure network emerged when we only activated "provide help" interaction rules (figure 3(a)). On the other hand, a more "cliquey" and strongly clustered network was observed when we activated only the "friendly chat" interaction rules (figure 3(b)). These results are in agreement with [30].

3.4 Virtual Experiments

Finally, we carried out virtual experiments with the simulation to identify the relationship between the parameters and the social network characteristics. This was done through changing the value of a specific parameter and observing the emerging social network.

The value of the five parameters was varied at 11 levels independently and 30 social networks were generated for each level. For instance, to explore the influence of the cohesion factor on the social network, we generated 30 networks for each 11 level of cohesion factor (from 1 to 10 on step 1) holding all other parameters constant. The average of the SNA measures was calculated for each set of the 30 social networks. A total of 30 x 11 x 5 (parameters) = 1650 social networks were produced for the main analysis.

4 Results and Design Implication

In this section, we present the results of the virtual experiments and discuss how the results could be practically useful in various sociability design scenarios. Note that the results were exploratory. Instead of using the simulation to confirm hypotheses, the simulation was used to explore the dynamics of the community in relation to the parameters.

4.1 Sociability Design Scenario 1: Friend Network

One benefit of participating in a guild community is to make and maintain friends. We believe that the number of friends people have on an individual level might influence the structure of the whole social network. Therefore a sociability consideration of online communities is to design and facilitate friend networks. To explore this scenario, we analyzed the causality of the parameter "social budget" and the social network formation of the community. Social budget defines the maximum number of immediate friend(s) an agent could have (note that the value of parameters is relative, i.e social budget of 4 does not literally means 4 friends). The higher the social budget, the bigger the agent's friend network could be.

Figure 4 shows that when the agents had 0 social budget (i.e. no friend at all), the network density was relatively high. That was because the simulated agents were interacting randomly with everyone in the community, resulting in the creation of more ties and thus a dense network. The density dropped quite drastically when the agents had only one social budget.

As the agents were allowed more friends (higher social budget), the density rose gradually since the increase of social budget resulted in the expansion of their friend network size. However, even at social budge of 10, the density never reached a level as high as when we had random interaction (when social budget = 0). This was

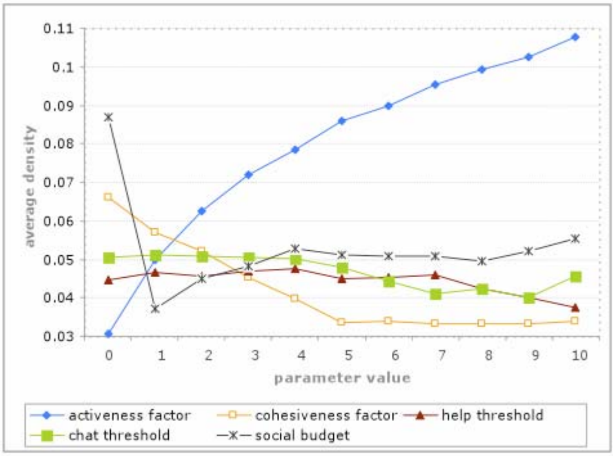


Fig. 4. The affect of the parameters on the average density of the social network

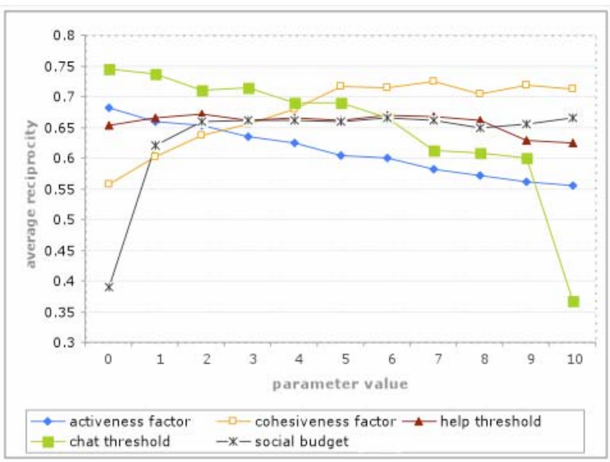


Fig. 5. The affect of the parameters on the average reciprocity

confirmed by 30 additional simulation runs from which we found that when social budget was 8, the average size of friend networks was only 5.416.

Hence, we can safely claim that once an agent had built up a friend network, the agent would be rarely interacting with strangers (non-friends). Therefore, the friend network stopped growing. This explains why the density remained stable even when we kept increasing the social budget.

Contrary to the density, the reciprocity of the network saw a jump as social budget was increased from 0 to 1 (Figure 5). We believe that this was due to the formation of strong ties within the agent's friend network. On the other hand, when the agent was interacting randomly only with strangers (social budget = 0), the chance of building reciprocated ties was low.

Therefore, we can infer from Figure 4 and 5 that although more ties were formed when social budget was 0, these ties were mainly unreciprocated and thus weaker ties. Once a friend network was formed, the agents interacted more frequently with their friends, resulting in more reciprocated ties. However, increasing the social budget further did not increase the reciprocity of the whole network because as previously shown, the friend networks stopped expanding.

From this analysis of the relationship between social budget and social networks, we can conclude that there is a need of a design mechanism that facilitates the formation of friend networks that promote strong ties. Moreover, the results suggest that once the agents formed a small friend network they stopped interacting with new people. Although this resulted in stronger ties, the density was compensated. Therefore, we could introduce a design that could support not only the *formation* but also the *expansion* of friend networks. For instance, a friend suggestion functions could be implemented to recommend other players who might be doing similar quests and have complementary skills.

4.2 Sociability Design Scenario 2: Level of Activity

One common belief is that active participation ensures the success of an online community. Therefore, one major sociability design issue is to design to motivate users to act. We explored this scenario with the parameter "activeness factor" that defines the level of activity of the community overall. The higher the activeness factor, the more likely the agents will initiate an interaction.

From all the graphs (figure 4-7), we can see that the activeness factor had a profound impact on all aspects of the social network. Firstly, we observed a continuous increase in the network density as the activeness factor was increased (Figure 4). This result was expected, since when more agents initiated an interaction, more ties were created in the network.

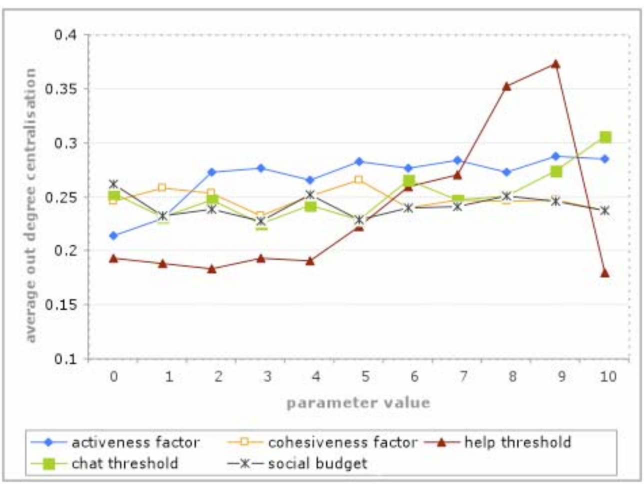


Fig. 6. The affect of the parameters on the average out-degree centralization

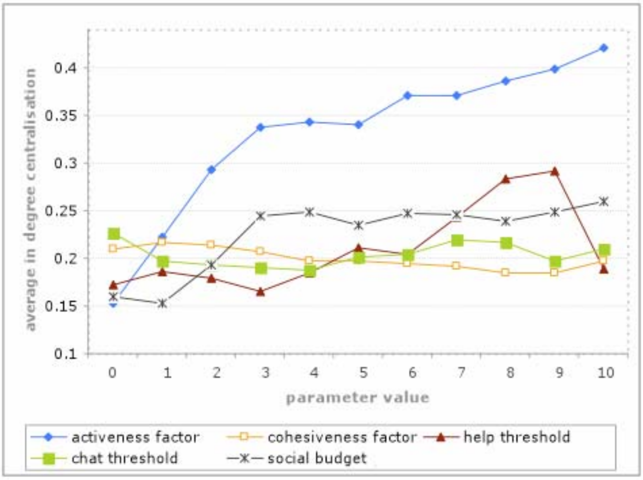


Fig. 7. The affect of the parameters on the average in-degree centralization

From Figure 6 and 7, we noticed that in-degree centralization rose much more than out-degree centralization as the activeness factor was increased. Out-degree centralization seemed to level off at 2 while in-degree centralization continued to rise.

Degree centralization can be used as a measure of equality of the network. The higher the degree centralization, the less equal the network is. This is caused by the emergence of a small "elite" group who are either popular or who dominate the interaction. The network could be unequal in term of the number of messages sent out (out-degree) and the number messages received (in-degree). Out-degree centralization can therefore be related to *domination*, while in-degree centralization implies *popularity*.

Since density increased as activeness factor went up, it is reasonable to observe a rise as well in in-degree centralization. The increase in network density (thus the number of ties formed) increased the number of messages being sent to these popular agents, amplifying the inequality in the social network.

The affect of activeness factor on out-degree centralization was rather counter-intuitive because one would have expected a decrease in out-degree centralization because higher activeness factor motivated more agents to initiate an interaction. In other words, more agents would have a chance to initiate an interaction, thus lowering inequality in message sending. However the results showed an increase in out-degree centralization, suggesting otherwise.

To understand this, we examined the increase in the number of messages sent by each agent. We ran additional 60 simulations to collect data at individual level. We found that by increasing the activeness factor from 0 to 2, the number of messages sent by each agent increased at different individual rates. Correlation analysis showed a positive relationship between the number of messages sent when activeness factor = 0 and the increase in messages sent when activeness factor increased from 0 to 2, $r = 0.330$, $p < 0.05$. It means that active agents who already dominated the community sent out even more messages when the activeness factor was increased. The increase impacted inactive agents to a less extent. This created a more unequal situation in the social network, widening the activity gap between the active and inactive agents.

As for reciprocity, we observed a gradual drop as the activeness factor increased. This was due to the fact that the increase in density was caused by an increase in the number of out-going ties instead of in-coming ties, i.e. as the agents were likely to initiate more interaction while the likelihood of reciprocating a tie stayed relatively constant.

What do these results tell us about sociability design? We found that higher activeness factor caused the formation of more ties in the community. This finding conforms to the idea that one should attempt to encourage the community members to become more active in participation. This is often achieved through the introducing ranking system that rewards all active behaviors. However, as shown in our analysis, high levels of activity could result in inequality in the community. Higher in-degree centralization suggests that only a small group of guild members benefited (through receiving more messages) from the high level of activity in the community. Depending on the types of community we are designing, this might not be desirable because inequality can often lead to social isolation and exclusion. Moreover, high level of activity also resulted in formation of weaker ties (low reciprocity). When designing a social support or empathy community, we might prefer a social network with reciprocated but fewer ties over a network with a lot of unreciprocated ties. In a collaborative writing community (such as Wiki) however we might prefer a social network with a lot of ties even if they are not reciprocated.

5 Discussion and Conclusion

In this paper, we proposed a new method for sociability design and research through the use of simulation. Due to space limitation, we were not able to discuss in detail the impact of all the parameters to social network formation but we hope we have demonstrated the strength of ABM in studying online communities. Our simulation was empirically based in two ways. Firstly, the interaction rules were formalized and drawn from the analysis of the user interaction of a WoW guild community. Secondly the simulation was validated through SNA of the interaction network of the existing community. We would also like to reiterate that the main aim of this study was to demonstrate the strengths/weakness of social simulations in sociability design and the results were largely exploratory.

As aforementioned, since we are unable to directly modify the community or its users, we need an alternative way to study how the underlying characteristic of the community influence the social network that arises. We believe that simulations allow us to understand this causal relationship by generating data which may not be possible to obtain empirically. This understanding can also enable us to design better ways to transform and expand these communities with minimal risks. Another huge advantage of ABM is that it is easy to conduct simulation runs and generate a large number of data. Through simulations, we can relate the simulation parameters to design decisions or social policy reinforcements.

Therefore, one obvious question is of course "how do we practically relate design/policy decisions to the change in local level interaction (the simulation parameters)", for instance "how do we know that introducing a specific reward system will lead to the increase/decrease of the activeness of the community?" We believe this can be done through conventional HCI methods. Existing HCI methods that focus on empirical studies/user testing can be easily integrated into simulations. Experiments can be carried out to examine the impact of a policy change or modification of graphical user interface elements on interaction behaviors at individual/micro level. For instance, we could conduct small-scale experiments with 30 people and estimate to what degree a particular “reward system” (a design decision) increases or decreases their initiative to send a message (a simulation parameter). We can then incorporate this empirical result into the simulation to predict the social network that might arise through the introduction of this reward system.

Apart from the potential strength in practical sociability design of online communities, simulations could also allow us to examine fundamental research question regarding user interaction behavior and social networks. Through the process of modeling and simulation development, we gain a deeper understanding of the mechanism of the community because the interaction rules need to be formulated very precisely to be programmable into ABMs. Therefore a possible use of simulation in sociability research is to construct theory-driven simulations. Using simulations, we can formalize social theories into precise program codes and examine the theories in the context of online communication.

The simulation presented in this paper focused on how social interactions affect network generation. A reverse way of social network simulation is to impose specific network structures and observe how these affect social interactions. For instance, we could examine the difference between a “star” and “clique” structure network in term of information diffusion in an online community. Finally, more future work needs to be done to enhance the validity of the simulation in order to use the simulation as a valid tool for sociability design. Therefore, more data needs to be collected and analyzed (e.g. include more communities, analyze the virtual actions of the players instead of just the chat log, etc)

As a conclusion, simulations are a powerful technique for understanding sociability in online communities. Our simulation was not only capable of generating results which were in accordance with what was observed in the real community, but also allowed us to explore various what-if questions that helped us to understand the sociability design tradeoffs of online communities.

References

1. Ahrweiler, P., Pyka, A., Gilbert, N.: Simulating knowledge dynamics in innovation networks (SKIN). In: Leombruni, R., Richiardi, M. (eds.) Industry and labor dynamics: The agent-based computational economics approach. World Scientific Press, Singapore
2. Antona, M., Bommel, P., Bousquet, F., Page, C.L.: Interactions and organization in ecosystem management: The use of multi-agent systems to simulate incentive environmental policies. Paper presented at the 3rd Workshop on Agent-Based Simulation., Ghent, Belgium (2002)
3. Blizzard Entertainment, World of Warcraft (2004), `http://www.worldofwarcraft.com/` (Last retrieved, 24 July 2008)
4. Burnett, G.: Information exchange in virtual communities: A typology. Information Research: An International Electronic Journal, 5(4) (2000)

5. Carley, K., Lee, J., Krackhardt, D.: Destabilizing Networks. Connections 24(3), 31–44 (2001)
6. Doran, J., Palmer, M., Gilbert, N., Mellars, P.: The EOS project: Modelling upper paleolithic social change. In: Gilbert, G.N., Doran, J. (eds.) Simulating societies: The computer simulation of social phenomena. UCL Press, London (1994)
7. Ducheneaut, N., Moore, R.J., Nickell, E.: Designing for sociability in massively multiplayer games: An examination of the third places of SWG. In: Other Players conference (2004)
8. Epstein, J.: Modeling civil violence: An agent-based computational approach. In: Proceedings of the National Academy of Sciences of the USA, vol. 99(3) (2002)
9. Epstein, J., Axtell, R.: Growing artificial societies. MIT Press, Boston (1997)
10. Gilbert, N., Troitzsch, K.G.: Simulation for the social scientist. Open University Press (2005)
11. Golder, S.A., Donath, J.: Socaial roles in electronic communities. Paper presented at the Association of Internet Researchers (AoIR) conference Internet Research 5.0, Brighton, England (2004)
12. Jager, W., Popping, R., Sande, H., v. d.: Clustering and fighting in two-party crowds: Simulating the approach-avoidance conflict. Journal of Artificial Societies and Social Simulation 4(3) (2001)
13. Klemm, P., Hurst, M., Dearholt, S.L., Trone, S.R.: Gender differences on internet cancer support groups. Computer Nursing 17(2), 65–72 (1999)
14. Kollock, P.: The economies of online cooperation: Gifts and public goods in cyberspace. In: Smith, M.A., Kollock, P. (eds.) Communities in cyberspace, London. Routledge (1999)
15. Kolo, C., Baur, T.: Living a virtual life: Social dynamics of online gaming. The International Journal of Computer Game Research 4(1) (2004)
16. Krippendorff, K.: Content analysis: An introduction to its methodology. Sage Publications, Newbury Park (1980)
17. Lamieri, M., Ietri, D.: Innovation creation and diffusion in a social network: An agent based approach (2004), http://ssrn.com/abstract=937255 (Last retrieved, 09 September 2008)
18. López-Sánchez, M., Noria, X., Rodríquez-Aguilar, J.A., Gilbert, N., Shuster, S.: Simulation of digital content distribution using a multi-agent simulation approach. In: Vitria, J., Radeva, P., Aguilo, I. (eds.) Recent advances in artificial intelligence research and development, pp. 341–348. IOS Press, Amsterdam (2004)
19. Macy, M.W., Sato, Y.: Trust, cooperation, and market formation in the U.S. and Japan. Paper presented at the Proceedings of the National Academy of Sciences of the United States of America (2002)
20. MaxTsvetovat, Carley, K.: Knowing the enemy: A simulation of terrorist organizations and counter-terrorism strategies. Paper presented at the CASOS Conference, Pittsburgh, PA (2002)
21. Newman, M.: The structure and function of complex networks. Siam Review 45(2), 167–256 (2003)
22. Niedner, S., Hertel, G., Hermann, S.: Motivation in open source projects: An empirical study among linux developers (2000), http://www.i2.psychologie.uni-wuerzburg.de/ao/research/linux_study.php (Last retrieved, 24 April 2008)
23. Özman, M.: Network formation and strategic firm behaviour to explore and exploit. Journal of Artificial Societies and Social Simulation 11(17) (2007)

24. Preece, J.: Online communities: Designing usability, supporting sociability. John Wiley and Sons, Chichester (2000)
25. Raymond, E.: The cathedral and the bazaar: Musings on Linux and open source from an accidental revolutionary. O'Reilly and Associates, Sebastopol (1999)
26. Savicki, V., Lingenfelter, D., Kelley, D.: Gender language style and group composition in internet discussion groups. Journal of Computer-Mediated Communication 2(3) (1996)
27. Turner, T.C., Smith, M.A., Fisher, D., Welser, H.T.: Picturing USENET: Mapping computer-mediated collective action. Journal of Computer-Mediated Communication 10(4) (2005)
28. Waldrop, M.: Complexity: The emerging science at the edge of chaos. Simon and Schuster, New York (1992)
29. Wasserman, S., Faust, K.: Social network analysis: Methods and applications. Cambridge University Press, Cambridge (1994)
30. Welser, H.T., Gleave, E., Fisher, D., Smith, M.: Visualizing the signatures of social roles in online discussion groups. Journal of Social Structure 8 (2007)
31. Wilensky, U.: Netlogo: Centre for Connected Learning and Computer Based Modelling Northwestern University, Evanston, IL (1999)
32. Wilhite, A.: Bilateral trade and "small-world" networks 18, 49–64 (2001)
33. Wooldridge, M., Jennings, N.R.: Intelligent agents: Theory and practice. Knowledge Engineering Review 10(2), 115–152 (1995)
34. Zhang, J., Ackerman, M., Adamic, L.: Communitynetsimulator: Using simulations to study online community networks. Paper presented at the C&T2007, Lansing, MI (2007a)
35. Zhang, J., Ackerman, M., Adamic, L.: Expertise networks in online communities: Structure and algor. Paper presented at the WWW 2007, Canada (2007b)

Designing Interaction for Local Communications: An Urban Screen Study

Fiona Redhead and Margot Brereton

Queensland University of Technology, Australasian CRC for Interaction Design, Queensland, Australia
f.redhead@student.qut.edu.au, m.brereton@qut.edu.au

Abstract. This paper discusses the ongoing design and use of a digital community noticeboard situated in a suburban hub. The design intention is to engage residents, collect and display local information and communications, and spark discussion. A key contribution is an understanding of Situated Display navigation that aids retrieval from a long-term collection created by and for suburban community, and engaging qualities of this collection.

Keywords: Situated Display, Internet Technologies, Interaction, Urban Screen.

1 Introduction

This paper presents formative evaluation of a long-term study of a digital community noticeboard (figure 1) located in a general store in an outer suburb of Brisbane, Australia (figure 2). Other evaluations of publicly accessible Situated Displays have focused on observing interaction with either predetermined content or particular media [1-4] and present limited exploration of how the engaging qualities of Situated Displays can be applied to build a mesh of grassroots and sustainable communications focused around the display locality. This research is evolving over time in a local community context and through use [5], and aims to evaluate capacity for Situated Displays to support diverse local urban communications. The noticeboard is a starting point and visible focus of our efforts to find better ways to facilitate local communication and build cohesion in a suburb that is in some respects a "dormitory" or commuter suburb. Our prior research observed the difficulty residents of such suburbs face in communicating with each other around issues of importance and identified the desire for greater community interaction [6].

A key data feed into the design process is the participation of our users, which has been increasing with time. Our focus is to grow participation in order to increase the capacity for and habit of community communication. The strategy for growth is to make the noticeboard as useful, useable and engaging as possible. We understand from beta development some needs people have in regard to interacting with a Situated Display in a local urban context, and we plan to continue this relationship with our participants through supporting and inviting further participation. Decreasing barriers to participation and discovering interactions that engage people and mesh with their everyday experiences is key to growing the capacity for use.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 457–460, 2009.

Fig. 1. The noticeboard in use

Fig. 2. The general store

The noticeboard is situated in a local general store to allow rapid perusal of community content in a place that is passed by many local people frequently. The store is located opposite a primary school creating a community hub that people use for everyday interactions, both casual and more structured (i.e. fundraising activity).

The name of the noticeboard is ‘Nnub’ (an abbreviation of neighbourhood nub). Nnub is the coupling of a Situated Display and Internet technologies and operates through a Situated Display (noticeboard) interface on a touch enabled 40" panel and a Web interface. While the interfaces hold the same content for viewing, they are designed to serve different purposes. The noticeboard interface is directed at engaging people, browsing, and quick notice creation and distribution using the touch display and mobile devices (mobile interaction is planned for future iterations). The Web interface is directed at follow up interaction, searching, and creation and distribution of notices using files typically stored on a home computer. A desktop computer is installed at the shop to provide Web interface access. Anyone can post to Nnub.

2 Design and Use

As argued in the Community Informatics literature, grassroots engagement is required to create and sustain effective technologies for community building activity [7]. However participation in community building activity that is not conceived from grassroots motivation requires some level of invitation either through direct invitation or exposure [8, 9]. This research applies an exploratory prototype as a means to invite participation and inspire design iterations. The exploratory prototype is a central artefact in the Reflective Agile Iterative Design (RAID) framework [10, 11].

RAID is a framework for evolving social software through reflective and timely response, by the practitioner, to use. The primary contributing framework is Action Research. RAID is a cyclic process of design, feedback, and reflection focused around the development of an exploratory prototype. The RAID framework is applicable to our research as we evolve the design in a dialogue with how people use the interface and in a time frame that indicates the process is active and responsive.

Although Situated Displays have potential to engage people, barriers to participation can hinder this potential [12]. To move people to direct interaction, Brignull and Rogers [12] suggest the interface should clearly convey low commitment interaction that is quick and enjoyable. Brignull and Rogers made this observation at a single event. Our challenge is to decrease barriers to interaction (through both the noticeboard and associated Web interface) in order to grow participation and the capacity to support community communication in the long-term.

Our process involves iteratively changing our design in response to use, while also extending the functionality to seek further feedback with the aim of discovering interactions that are in varying degrees quick and enjoyable, yet serve to collect and support local communications. We have observed that people prefer to interact with the notices as objects on the noticeboard rather than through text based navigation, and prefer to interact with image notices rather than text focused notices. These observations led us to improve the interface and discover other means of engaging people in creating noticeboard content of interest to the community. Logging statistics reported below were collected from June 2008 to January 2009.

The navigation for an initial noticeboard layout was used very little, with people by far preferring to directly touch a notice in view (to look at in full size) than to use the primary or secondary navigation. Primary level navigation was conducted through the top menu bar (category, author, title, date) and side panel (photos or notices) and accounted for 13% of total touches. Secondary level navigation, which selected among categories, accounted for 6% of total touches. Similar findings have been observed on a Situated Display designed to display local photos at Wray Village [1]. Logging data and observations indicated people did not pay attention to navigation buttons and interacted directly with images in view.

However building a collection of information and communications for long-term reference requires useful navigation. This led us to implement open tagging. Rather than assigning prescribed categories, tags are added by the community of Nnub users (as contributors) to describe the growing body of content. In addition we removed all other forms of navigation leaving tags as the means to find particular notices. Of the total touches, touching a tag to view notices tagged with that word accounts for 30%. This indicates that open tagging is more meaningful to the community than the previous generic metadata and categories we had specified.

Related to this tendency to interact directly with the notices rather than text based navigation is the difference in habit of what people like to view and what they post. Viewing image notices (no text) at full size accounts for 44.9% (of the total touches including navigation) while viewing text focused notices at full size accounts for 7.9% (of the total touches including navigation). However, only 10% of uploads from registered users are image notices while 43% are text focused notices.

Based on these observations we have evolved the noticeboard as an exploratory prototype to increase direct interaction with images on the Situated Display. We have implemented ‘Scribbles’ [2]. Scribbles are created by drawing or writing directly on the touch screen and have been extremely popular with children and have rapidly increased the visuals on the noticeboard. This growth of Scribbles brings new considerations for interface design both in terms of extending possibilities around successful engagement, and preserving the core design intentions. Through Flickr we invited 16 people to join a group created around the locality of Nnub. The images added to the group are fed to Nnub. Of the 16 people invited, 10 people added images. This use of other Web 2.0 interfaces to grow participation extends our reach to find relevant content. We plan to publicise statistics about use to the users to allow reflective use

[10, 11]. This will illustrate the popularity of image notices and may increase the creation of image and image focused notices.

A key contribution of this work is the design and long term community deployment of an instrumented noticeboard that has enabled us to understand barriers to interaction with a Situated Display in a local urban context, in addition to exploring possibilities for creating content that people like to engage with. Future work will continue to respond to use to evolve solutions that reflect the range of information and communications that shape the local suburb.

Acknowledgements. This research/work was partly conducted within the Australasian CRC for Interaction Design, which is established and supported under the Australian Government's Cooperative Research Centres Programme. We would like to thank the general store for their generosity and enthusiasm.

References

1. Taylor, N., Cheverst, K., Fitton, D., Race, N., Rouncefield, M., Graham, C.: Probing Communities: Study of a Village Photo Display. In: OZCHI 2007, pp. 17–24. ACM, New York (2007)
2. Churchill, E., Nelson, L.: Interactive Community Bulletin Boards as Conversational Hubs and Sites for Playful Visual Repartee. In: HICSS-40 2007, pp. 76–85. IEEE Press, New York (2007)
3. Peltonen, P., Salovaara, A., Jacucci, G., Ilmonen, T., Ardito, C., Saarikko, P., Batra, V.: Extending Large-Scale Event Participation with User-Created Mobile Media on a Public Display. In: MUM 2007, pp. 131–138. ACM, New York (2007)
4. Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., Oulasvirta, A., Saarikko, P.: "It's Mine, Don't Touch!": Interactions at a Large Multi-Touch Display in a City Centre. In: CHI 2008, pp. 1285–1294. ACM, New York (2008)
5. Brereton, M.: Designing from Somewhere: A located, relational and transformative view of design. In: Löwgren, J., Malmborg, L., Binder, T. (eds.) (Re-) searching a digital Bauhaus, pp. 99–119. Springer, London (2009)
6. Redhead, F., Brereton, M.: A Qualitative Analysis of Local Community Communications. In: OZCHI 2006, pp. 361–364. ACM, New York (2006)
7. Gurstein, M.: Effective Use: A Community informatics strategy beyond the Digital Divide. First Monday 8 (2003)
8. Schuler, D.: The Seattle Community Network: Anomaly or Replicable Model? In: van den Besselaar, P., Koizumi, S. (eds.) Digital Cities III, pp. 17–42. Springer, Heidelberg (2005)
9. Carroll, J.M.: The Blacksburg Electronic Village: A Study in Community Computing. In: van den Besselaar, P., Koizumi, S. (eds.) Digital Cities III, pp. 43–65. Springer, Heidelberg (2005)
10. Heyer, C., Brereton, M.: Reflective Agile Iterative Design. In: SIMTech Workshop on Social Interaction with Mundane Technologies (November 2008)
11. Heyer, C., Brereton, M., Viller, S.: Cross-channel mobile social software: an empirical study. In: CHI 2008, pp. 1525–1534. ACM Press, New York (2008)
12. Brignull, H., Rogers, Y.: Enticing People to Interact with Large Public Displays in Public Spaces. In: INTERACT 2003, pp. 17–24. IOS Press, Amsterdam (2003)

WidSets: A Usability Study of Widget Sharing

Kristiina Karvonen[1], Theofanis Kilinkaridis[1], and Olli Immonen[2]

[1] Helsinki Institute for Information Technology HIIT,
Helsinki University of Technology
{kristiina.karvonen,theofanis.kilinkaridis}@hiit.fi
[2] Nokia
olli.immonen@nokia.com

Abstract. In this paper, we describe a study conducted to better understand and improve the usability of the reputation elements on an online widget sharing site called WidSets. With a series of interviews and an online questionnaire with users of WidSets, we seek to understand the motivations and instigators of the actual usage, the level of understanding, and acceptance of current reputation elements. We analyse the impact of these factors on WidSets users' willingness to download and use the publicly available widgets. The outcomes show the visual elements of the user interface to predominate as the source for information for the decision to trust and download. However, the accuracy of the interpretation of these elements remains ambiguous. There also seems to be little awareness of the underlying security issues related to possible malicious widgets and developers.

Keywords: Widget, reputation, usability, user study, questionnaire, trust.

1 Introduction

Web widgets are PC or mobile micro-applications performing a dedicated task such as showing news headlines, a weather forecast or a Wikipedia article. Widgets are typically freeware distributed through widget repositories[1]. In some cases, downloading a widget may involve a risk of abuse of user's private information such as contacts and location. Widget repositories typically do not guarantee the quality or trustworthiness of all widgets. In WidSets, some widgets have been verified by the staff. Users need to make downloading decisions based on widget description, logo, developer history and data reflecting other users' satisfaction with the widget, such as number of downloads, rating and comments.

In this paper, we describe the results of a study regarding usability of the reputation elements on the WidSets site. The study was performed as an interview and a web-based questionnaire among current WidSets users (n=38).

Our results indicate that while only a small part of the users find it difficult to use the web site or the mobile UI, the information related to trustworthiness and

[1] E. g. widsets.com, www.apple.com/downloads/dashboard, widgets.yahoo.com

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 461–464, 2009.

reputation is either poorly understood or underutilised. The most important elements to affect downloading decision were 1) the *description*, either 2a) *rating* (in web site) or 2b) *number of users* (in mobile UI), and 3) *widget logo*. The results suggest that visual appearance greatly influences users' decision making. Current users mostly trust the WidSets site while at the same time almost half consider the risk to be confronted by harmful content to be real.

2 Background

Reputation systems are used in internet services where users need to make trust decision concerning people they do not know beforehand. Reputation guides users' decision making. In eBay high reputation can lead to price premiums [1].

Online trust formation has many ingredients and has been widely researched on from various viewpoints, including technical, legal, social, psychological and philosophical [2]. For our work, the research on trust from a psychological viewpoint, as well as the social aspects of trust, especially work on collaboration and trust in the context of social networking seems most relevant.

WidSets is a combination of a service and a mobile Java MIDP2 client software. Users can browse the widget catalog with the phone client or with a PC on the WidSets web site (Fig. 1). "Picked" widgets are then downloaded and used in the client. Users can rate widgets, write comments and ask questions. Widgets can be developed and submitted to the site easily, e.g. by defining a logo and an RSS feed. More advanced widgets can be developed using the WidSets SDK and a scripting language, or web technologies (HTML, XHTML, JavaScript and CSS). Web technology widgets have features that are not available to regular web applications such as access to contacts, GPS and camera.

Fig. 1. Screenshot of WidSets mobile UI (on the left) and web site widsets.com (on the right)

3 The Study

We began by interviewing current widget users to gain an initial understanding on how widgets are currently perceived and used. A web-based questionnaire[2] was built on basis of the interview outcomes to gather feedback from more users. The questions in the questionnaire consisted of a modified version of the interview structure and topical areas. We had in all 38 users with 34 males and 3 females[3]. 58 % had at least a bachelor's degree. All interviewees (n=8) were Finnish and located in Finland, whereas six of the questionnaire respondents (n=30) represented also various other nationalities. Most popular types of widgets were related to News, Weather, and Transportation. Over 75 % users had been using the WidSets service for over a year. 10 % had also developed and published their own widget, besides downloading.
Next, we analyse the outcomes in more detail.

Meaning of widgets. Users saw the service primarily as a mobile service, to make their mobile phone connected to the Internet: "It's a nice graphical user interface for accessing the Internet"; "I see it as a mobile service where you can get easily information"; "For me, it is a 'Web to Mobile' service".

Usage. The motivation for the usage varied: Some had come across the site at random on the Internet; some had learnt about the site on an online discussion board, when others had had their friends' recommendation of the site. Currently, the primary motivation behind widgets usage was to "get information", with "entertainment" coming second. The information-centricity was emphasized by other major motivations such as "work" and "organizing my day". Social pressure was one motivation: "[I use it] because my friends do".

User experience. Most users ranked both the website and the mobile user interface of the WidSets service as *easy* or *very easy to use* (58 % for website, 75 % for mobile UI), very few finding it *difficult to use* (5 % for website, 9 % for mobile UI). Almost half of the users preferred the mobile UI over the website, 30 % preferred the website and about 20 % could not distinguish between the two. As the mobile usage dominates, it is natural that mobile UI becomes easier to use, through habituation. Another factor contributing to the experienced ease of use is probably the evident simplicity of the mobile UI, in comparison to the website

Reputation elements. To find out which reputation elements were used when deciding to download a particular widget, we showed users screenshots of the download phase for five widgets from both the WidSets website and mobile UI. Users were asked to select the three most and least important reputation elements.

For the website, users most looked at *Description* (27%); *Rating* (21%) and the *Widget logo* (12%), while they ignored *Date* (2.33%), *Developer info* (2%) and *Using this Widget* (0.5%). For the mobile version, users looked at *Description* (50%), *Number of users* (30%) and *Widget logo* (23%) and ignored *Widget status* (10%), *Similar Widgets* (9%) and *Developer info* (3%).

The actual reputation information available remained for the most underutilised, with only the *Rating* (for website) and *Number of users* (mobile) among the 'looked

[2] http://survey.hiit.fi/index.php?sid=52912&newtest=Y

[3] One questionnaire respondent did not state gender.

at' elements. Instead, the *Description* was predominating for both UIs. *Developer info* was among least attractive elements. A partial reason for this may be how this information is currently shown to the users: behind an additional mouse click.

Trust. Overall, users expressed a trusting attitude towards widget sharing, with 76 % trusting the service "Absolutely" or "Pretty much", and only 5 % "not really" trusting the service. The likelihood of being confronted by harmful content on the WidSets site was considered to be a real risk by as high as 44 % of the users, whereas 46 % stated that this was unlikely. Only 5 % of the users expressed a careless attitude, "not caring" if there is the possibility of downloading harmful content from the site.

4 Discussion and Conclusions

The outcomes of our study give evidence of the prominence of the UI visual elements as the source for information behind the downloading – trust – decision. However, the accuracy of the interpretation of these elements and their interplay remain ambiguous. There seems to be little awareness of the underlying security issues related to possible malicious widgets and developers. In the current state with no practical risk this has not been a real issue but new type of widgets may change the situation in the future.

We found users consider the most important elements affecting download decisions to be the description, widget logo (reflecting also 'branding effect') and either rating (in web site) or number of users (in mobile UI). Why rating was found so important in the web UI may be due to its highly visible appearance. Various community features like discussion boards, ratings, comments and presentation of developer history, are hardly used. A friend's network does not appear to be utilized.

The importance of visual appearance suggests that important reputation elements, such as developer reputation, should be presented in a visually prominent manner. An important question is, should the system display the user an aggregation of various reputation data, or should the reputation elements be presented separately, allowing the user to make his or her conclusions in a more independent fashion.

Acknowledgements. This work was supported by TEKES as part of the Future Internet program of the ICT cluster of the Finnish Strategic Centres for Science, Technology and Innovation. The users in the study, and the efforts made by Patrik Floréen, Petteri Nurmi, Yiyun Shen and Sini Ruohomaa in commenting the questionnaire are thankfully acknowledged.

References

1. Resnick, P., Zeckhauser, R., Swanson, J., Lockwood, K.: The Value of Reputation on eBay: A Controlled Experiment. Experimental Economics 9(2), 79–101 (2006)
2. Riegelsberger, J., Sasse, M.A., McCarthy, J.D.: The mechanics of trust: A framework for research and design. International J. of Human-Computer Studies 62(3), 381–422 (2005)

A Model for Steering with Haptic-Force Guidance

Xing-Dong Yang[1], Pourang Irani[2], Pierre Boulanger[1], and Walter F. Bischof[1]

[1] Department of Computing Science,
University of Alberta
Edmonton, AB, Canada
{xingdong,pierreb,wfb}@cs.ualberta.ca
[2] Department of Computer Science,
University of Manitoba
Winnipeg, MB, Canada
irani@cs.umanitoba.ca

Abstract. Trajectory-based tasks are common in many applications and have been widely studied. Recently, researchers have shown that even very simple tasks, such as selecting items from cascading menus, can benefit from haptic-force guidance. Haptic guidance is also of significant value in many applications such as medical training, handwriting learning, and in applications requiring precise manipulations. There are, however, only very few guiding principles for selecting parameters that are best suited for proper force guiding. In this paper, we present a model, derived from the steering law that relates movement time to the essential components of a tunneling task in the presence of haptic-force guidance. Results of an experiment show that our model is highly accurate for predicting performance times in force-enhanced tunneling tasks.

Keywords: Haptic guidance, steering task, Steering law, Fitts' law.

1 Introduction

Steering through straight "tunnels" of different width and length is a common task when interacting with computer systems. Examples include navigating and selecting items in menus, selecting a line or sentence, and tracing to learn patterns. Researchers have introduced force-guidance as a means to assist in the task of steering through tunnels. Users' hand movements are restricted within a tunnel of specific width by imposing forces orthogonal to the movement. Examples include haptic-guided menu selection [4], hand-writing training [15], and surgical training [16]. The guiding force can be either generated by a force-feedback device, such as a Phantom haptic device [14], or simulated by software [11, 13]. Force-guidance is deemed to be a good method for facilitating steering tasks as it can not only reduce errors made by accidently exiting the tunnel, but also improve user performance [4, 6, 7].

As haptic guidance is becoming more popular in assisting steering tasks, a model is required that can predict the movement time of steering tasks in the presence of a guiding force. However, only a few studies have been carried out on this topic. Earlier work with force-guidance uses parameters best suited to the context or the task at hand [4, 6, 7, 17]. A more generic model is required that identifies the relationships

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 465–478, 2009.

governing tunneling with force-guidance and can better guide designers in selecting appropriate parameters for a force-guided tunneling task. Such a model could provide answers to many important questions: Does larger guiding forces always lead to better performance? Does force-guided steering extend from the steering law or does a new law dictate the relationship between parameters involved in force-guided tunneling? Should force-guided tunneling take into consideration the intensity of the guiding force? Does a given device have adequate stiffness to guide a user? Without an answer to these questions, the design of force-enhanced steering interfaces is difficult and will at best remain a task that involves significant trial-and-error work for the practitioner.

In this paper we introduce a new haptic-steering model, which is derived from the Accot-Zhai steering law [1] with the inclusion of factors that influence haptically guided tunneling tasks.

2 Related Work

We discuss prior work describing systems that improve steering tasks with haptic or tactile feedback and also discuss existing models for pointing and steering.

2.1 Models for Pointing and Steering

Fitts' law is a governing pointing and aiming [9]. It serves as a fundamental benchmark model from which other laws have been derived [1, 3]. Fitts' law models the movement time (*MT*) of the selection of a one-dimensional target of width *W* at a distance of *A*:

$$MT = a + b\log_2\left(A/W + 1\right), \tag{1}$$

where *a* and *b* are empirically determined constants. The logarithmic term is called the *index of difficulty* (ID). The larger the index of difficulty, the longer it takes to complete the task, and vice-versa.

Fitts' law was adapted by Mackenzie and Buxton [10] for selecting two-dimensional targets. The target was a rectangle of width *W* and height *H*. They found that Fitts' law can be adapted by replacing the *W* in Equation 1 by the smaller of two sides of the target. Accot and Zhai [3] demonstrated that the equation below provides a more precise prediction as it considers the effects of both width and height of the target:

$$MT = a + b\log_2\left(\sqrt{\left(A/W\right)^2 + \eta\left(A/H\right)^2} + 1\right), \tag{2}$$

where η is a coefficient to weigh the relative contributions of width and height.

In addition to pointing and selection tasks, Fitts' law can also be extended to predict movement time for goal crossing tasks [1, 2], where, the distance *A* and the height *H* of a goal determine the *index of difficulty* for the task (see Fig. 1 middle). Accot and Zhai stated that the task of navigating through a constrained path or tunnel can be considered as a series of crossing tasks, in which a crossing occurs at small intervals.

Based on the model of goal crossing, they introduced the steering law [1] for modeling performance time of steering through a constrained path, as follows:

$$MT = a + b\int_C ds / W(s), \tag{3}$$

where C is the constrained path, and W(s) is the width of the path at point s. It has been demonstrated that the steering law is sufficiently general, and, in addition to modeling common trajectory-based interactive tasks such as navigation and selection of cascading menus [4], it can also be applied to locomotion tasks in virtual reality environments [19].

The steering law is valid for paths of any shape with variable length and width. Equation 4 gives a simplified version, which captures the relationship between the variables along a straight path with a constant width. This version of the steering law is commonly used in modeling novel interactive techniques in GUI environments, as most of the trajectory-based tasks in GUI environments involve steering through straight tunnels [4, 8].

$$MT = a + b\left(A / W\right). \tag{4}$$

Recently, Kattinakere et al. [8] extended Equation 2 for tasks requiring steering in three dimensional tunnels. They showed that, all three dimensions of the tunnel affect task completion time. Their model is expressed as follows:

$$MT = a + b\sqrt{\left(A / W\right)^2 + \eta\left(A / T\right)^2}, \tag{5}$$

where A is the length of the tunnel, W is width of the tunnel, T is the height of the tunnel, and η is an empirically determined constant determining the relative contributions of W and T.

2.2 Force-Enhanced Steering

Research has shown that the performance of steering tasks can be improved by providing users with tactile feedback during the movement along a path. Campbell et al. [5] demonstrated that, while navigating through a circular tunnel, the chance of accidently steering outside of the tunnel can be significantly reduced by providing tactile feedback at the sides of the tunnel. They also showed that, if tactile feedback can be provided along the center of the path, users can perform the task faster. Forsyth and MacLean [7] also showed that, task completion time of a steering task can be improved significantly by providing a guiding force towards the center of a path.

A study by Dennerlein et al. [6] involved moving a mouse cursor using a regular mouse or a force-feedback mouse through a small tunnel, which was placed either horizontally or vertically. The force-feedback mouse generated a guiding force that physically pulled the cursor to the center of the tunnel. They showed that task completion time can be improved significantly with force-guidance. They also revealed that, without force-guidance, it took longer to steer through vertical tunnels than horizontal tunnels, whereas with force-guidance the speed difference between vertical and horizontal tunneling was reduced. They also demonstrated that the steering task assisted by force-guidance can be modeled using Equation 4.

Similarly, Ahlström [4] showed that the task of navigating a force-enhanced cascading menu can be modeled using Equation 4. The guiding force was simulated in software. For items with no submenu, the guiding force pushed the mouse cursor to the center of the path. For the items with a submenu, the force pushed the mouse cursor towards the submenu. Ahlström's study showed that force fields can decrease menu selection times by 18% on average.

Although [4] and [6] show that force-enhanced steering tasks can be captured using Equation 4, they do not include relevant parameters such as the stiffness of the haptic device or the strength of the guidance force. If these are to be used in haptic guidance systems, a model is necessary that includes these parameters. In the next Section, we describe the derivation of a haptic-steering model based on the steering law.

3 Model of Force-Enhanced Steering Tasks

Force-guidance is based on the premise that deviations from a central path are pulled back to the center of the tunnel. This reduces inadvertent deviations from the tunnel. Furthermore, users are required to spend extra effort to get out of the tunnel. Conceptually, this form of force-guidance has an effect similar to increasing the width of the tunnel. Therefore we propose an initial model that is inspired by the steering model discussed above. In our model, the guiding force is assumed to increase the width of the tunnel. For the steering task, we first propose a model for the *goal crossing* task. Our model for force-enhanced goal crossing task is expressed as follows:

$$MT = a + b\log_2\left(A/\left(W + \eta \times F\right) + 1\right), \tag{6}$$

where F is the intensity of guiding force, and η is an empirically determined constant to adjust the effect of F, which is relatively small compared to A and W. While this derivation seems straightforward, there are several elements that do not make this model practical. For instance, in most force-guidance systems, the magnitude of the guiding force is variable: It varies as a function of deviation distance from an ideal path (possibly the center of the tunnel) [17]. To make the model practical, one needs to have a priori knowledge of F, which is difficult to obtain as it varies the magnitude of the force, as applied by the user using the system. We thus replace force magnitude by the spring stiffness S to represent the intensity of the guiding force. For a given deviation a higher S results in a stronger guiding force. The *index of difficulty* for the haptic guidance system can be formulated as:

$$ID = \log_2\left(A/\left(W + \eta \times S\right) + 1\right), \tag{7}$$

where S is the stiffness of a virtual spring. Similar to [1, 8], we break the tunneling tasks into a series of N goal crossing tasks. For each goal crossing task, the width of the goal stays the same while the distance between goals is A/N. As N grows infinitely, the task becomes a steering task. Using a first-order Taylor series to approximate the *index of difficulty* for the N goal crossing tasks, we obtain

$$ID = \lim_{N \to \infty} ID_N = \frac{1}{\ln(2)} A / (W + \eta \times S). \tag{8}$$

This leads to our model of the *force-enhanced tunneling* task as

$$MT = a + b\left(A / (W + \eta \times S)\right). \tag{9}$$

In the remainder of the paper, we describe a series of experiments to validate the proposed model. Through these studies, we also gain insight into user performance patterns with respect to the intensity of guidance force, an aspect that has been unexplored in previous work.

4 Apparatus

The experiments were run on Intel Duo Core 1.66 GHz CPU with 1 GB RAM. Force-guidance was provided by a PHANToM Omni force-feedback device. The experimental interface was displayed on a 17-inch 1280×800 LCD monitor. Participants placed their dominant hand on an armrest. The height of the armrest was sufficient to raise participants' wrists to a comfortable height for manipulating the stylus. Participants were asked to hold the stylus of the PHANToM device like a pen and to perform the crossing and the steering task while the pen was allowed to slide on a plastic panel (see Fig. 1 left). The buttons on the stylus were programmed for participants to control the experimental procedure with the index finger. The test system was developed in C++ using OpenGL. The haptic guidance was implemented using the Open Haptics toolkit from SensAble Technologies [14].

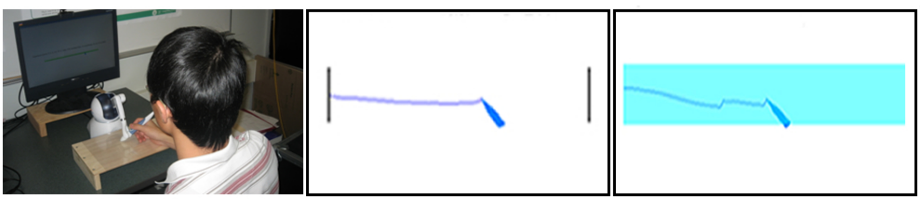


Fig. 1. Left: Experiment settings. Middle: 2D crossing task. Right: 2D tunneling task.

5 Haptic Guidance

Force-guidance was provided by the PHANToM Omni device in a passive constraint manner. The guiding force was triggered when the stylus end-effector deviated from the ideal trajectory, i.e. the middle of the tunnel, and the end-effector was dragged back to the ideal path. As in most passive force-guidance systems [17], the magnitude of the guiding force varied proportional to the distance between the end-effector and the ideal path, i.e. the force magnitude increases as the distance from the center of the path increased.

The direction of the correction force was calculated by projecting the position of the end-effector onto the central path of a tunnel. The goal was to guide the user's hand to the nearest point on the central path. Force magnitude was computed by applying Hooke's Law, *Magnitude (F)* = $S \times D$, where S is the stiffness of virtual spring, and D is the distance between end-effector position and the target position on the central path. We also added a damping system to mitigate buzzing [17].

6 Experiment 1: 2D Goal Crossing with Haptic Guidance

We first ran a pilot study to validate that Equation 6 works for a crossing task (see [8]) while haptic guidance is provided. If it is valid then we generalize it to the haptic steering model. The task required participants to move a cursor to cross two vertical goals of various heights. The goals were rendered with a width of two pixels. We provided a guidance force throughout the entire trial so that the guidance was available not only when the cursor crossed the goals but also when it is occurred at any position during a trial.

6.1 Participants

Three volunteers participated in this study. The participants were male and between the ages of 20 and 23. All were right-handed. Only one had previous experience with the PHANToM Omni device.

6.2 Procedure

Participants were asked to hold the stylus of the PHANToM Omni like a pen, and to move the stylus from left to right to move the cursor, which was rendered as a pen (see Fig. 1 middle), to cross the two bars. User trajectories between the two goals were displayed with a line.

To start a trial, participants were required to place the cursor inside a start region located to left of the start bar (left bar). The region rendered in red, was 5mm wide, and had the same height as the piecewise goals. Once the cursor was in the start region, participants clicked the barrel button on the stylus to indicate they were ready to proceed. The start region then turned green suggesting that the participants could begin. For a successful trial, both bars needed to be crossed from left to right. For instance, a participant who overshot a bar without crossing it, had to back-track to the left of the bar and repeat the crossing for that goal.

A trial started as soon as the start bar was crossed, and ended as soon as the end goal was crossed. Trials were presented with different heights and distances. We also varied the magnitude of the guiding force at four levels. Since humans have a relatively poor sensitivity to changes of force magnitude [18], we displayed the levels of guiding force both haptically and visually. When displayed visually, the strength of the guiding force was represented by number of stars in a way that the minimum guidance was represented by "*", and the maximum guidance was represented by "* * * *".

Participants practiced prior to the experiment. The experiment lasted 45 minutes. Participants were offered breaks between trials.

6.3 Experimental Design

The study employed a 4×4×4 within-subject factorial design, The independent variables are the width of the goals, W (5, 10, 15, and 20 pixels, corresponding to 5, 10, 15, 20 mm), the distance or amplitude between the goals, A (26, 80, 133, and 186 pixels, corresponding to 50, 150, 250, and 350 mm), and the intensity of the haptic guidance, which is represented by spring stiffness, S (0.05, 0.2, 0.35, 0.5). The combination of the independent variables resulted in 64 conditions, with the corresponding IDs calculated by Equation 7.

The experiment was organized into 3 blocks. Each block contained 5 threads, and each thread consisted of 64 trials each representing one $A \times W \times S$ combination. This resulted in a total of 960 trials. Trials were presented randomly within each thread.

6.4 Results

Movement time (*MT*) was recorded between the times the cursor crossed the left and right bars. Similar to [8], we removed outliers that were more than 3 standard deviations from the group mean (less than 1% of the data). A repeated-measures analysis of variance of movement times yielded a main effect of S ($F_{3,6} = 8.05$, $p < 0.02$), W ($F_{3,6} = 36.8$, $p < 0.001$), and A ($F_{3,6} = 64.04$, $p < 0.001$). There were also significant interactions A by W ($F_{9,18} = 11.91$, $p < 0.001$), S by W ($F_{9,18} = 26.7$, $p < 0.001$), and S by W by A ($F_{27,54} = 2.49$, $p < 0.01$).

MT decreased as the intensity of the force-guidance increased. Movement time for S = 0.05, 0.2, 0.35, and 0.5 were 411, 359, 343, and 334ms. A post-hoc Tukey-Kramer analysis showed that movement time for $S = 0.05$ was significantly longer than the others (all $p < 0.001$), while the others did not differ significantly from each other (all $p > 0.05$).

With S increasing from 0.05 to 0.2, from 0.2 to 0.35, and from 0.35 to 0.5, the average completion time was reduced by 52, 16, and 9ms, suggesting that, beyond a certain threshold, the improvement on movement time decreased and leveled out as the intensity of guidance force increased. We can not confirm this statistically because of a small sample size. This question is further studied in the next experiment.

Using a non-linear regression method, we estimated the value of η to be 28.274. In order to be consistent with [1], the *index of difficulty* (ID) was measured in pixels. Fig. 2 shows the movement time as a function of the proposed *index of difficulty*. Linear regression analysis showed that the data fit the model with an R^2 of 0.92. The equation for *MT* is given by:

$$MT = -57.1 + 158.83 \times \log_2 \left(A / \left(W + 28.274 \times S \right) + 1 \right). \tag{10}$$

This pilot study validated our proposed model for predicting movement time for a force-enhanced goal-crossing task. In the next experiment, we validate the proposed model (Equation 9) for the force-enhanced tunneling task.

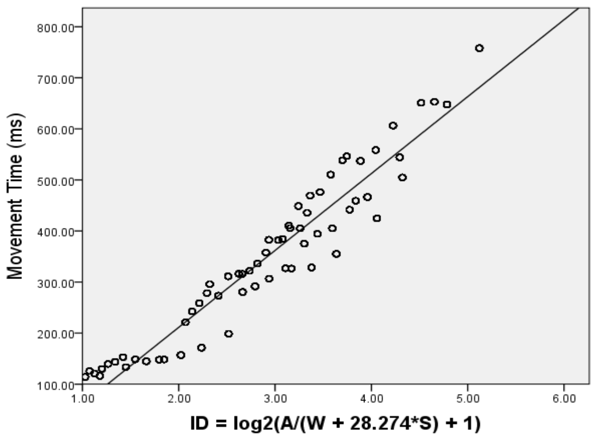


Fig. 2. Movement times by ID

7 Experiment 2: Tunneling with Haptic Guidance

The first experiment demonstrated that, when force-guidance is provided, our model of crossing can be used to predict movement time as a function of task difficulty. In this experiment, we studied haptically enhanced steering tasks. The task involved steering a cursor through prescribed tunnels. Force-guidance was provided to limit participants' hand movements within the tunnels. Similar to the first experiment, four levels of guiding force were tested. We investigated whether the proposed model (Equation 9) can be used to predict movement time for this task. We expected to determine how user performance varies with the intensity of the guiding force.

7.1 Participants

Fifteen university students participated in this study. The group consisted of 3 women and 12 men between the ages of 20 and 30. None had participated in the previous experiment. All but one were right-handed. Seven participants had experience with the PHANToM Omni device. The participants received $10 dollars for participation.

7.2 Hypotheses

Based on prior work and on the first experiment, our hypotheses were as follows:

H1: The effects of force-guidance on movement time reduce when the intensity of the force-guidance exceeds a threshold.
H2: The effects of haptic guidance on error rate reduce when the intensity of force-guidance reaches a certain threshold.
H3: Weak force-guidance result in more errors with longer and/or narrower tunnels. With strong force-guidance, the length and width of a tunnel has a weaker effect on performance.

7.3 Procedure

The procedure was similar to the first experiment. Participants were asked to steer a cursor through prescribed tunnels of different length and width (see Fig. 1 right). Force-guidance was provided to constrain participants' movements inside the tunnels. A successful trial required the participants to steer the cursor from the left end of the tunnel to the right end of the tunnel. A trial finished as soon as the cursor exited the tunnel's right end. An error occurred if the cursor exited the tunnel before the right end was reached. The tunnel was rendered in light green and turned to red if an error occurred. Similar to the first experiment, the intensity of the guiding force was displayed visually to the participants. We also displayed the error rate. Participants were asked to balance speed and accuracy so that their error rate did not exceed 3%.

A warm-up session was given to participants prior to the experiment. The entire experiments lasted 45 minutes. Breaks were encouraged during the experiment.

7.4 Experimental Design

The design of this experiment was the same as for the first experiment. A fully crossed design resulted in 64 IDs based on Equation 9. The experiment consisted of three blocks, each with 64 conditions that were presented in random order. Each condition was repeated 5 times. In total, there were 3 (blocks) × 5 (repetitions) × 64 (conditions) = 960 trials per participant for a total of 14400 trials.

7.5 Results

Movement time (*MT*) was recorded from the time the participants entered the tunnel on the left to the time of exit on the right. Similar to the previous experiment, we removed trials containing errors (1.3%). We also removed outliers more than 3 standard deviations from the group mean (1.8%). In the results reported below, Greenhouse-Geisser adjusted F-Ratios were used whenever appropriate.

A repeated-measures analysis of variance yielded a main effect of S ($F_{3, 42} = 104.72$, $p < 0.001$), W ($F_{3, 42} = 93.55$, $p < 0.001$) and A ($F_{3, 42} = 237.04$, $p < 0.001$). As expected, there were also significant interactions S by A ($F_{9, 126} = 86.2$, $p < 0.001$), S by W ($F_{9, 126} = 100.83$, $p < 0.001$), A by W ($F_{9, 126} = 103.9$, $p < 0.001$), and S by A by W ($F_{27, 378} = 45.95$, $p < 0.001$).

Hypothesis H1: The average movement times for $S = 0.05$, 0.2, 0.35, and 0.5 were 651, 494, 419, and 395ms. With S increasing from 0.05 to 0.2, from 0.2 to 0.35, and from 0.35 to 0.5, the average completion time was reduced by 157, 75, and 24ms. Post-hoc Tukey-Kramer analysis showed that movement times decreased significantly with each increase of S (all $p < 0.001$). Because of the design of this study, the results did not show the movement time improvement as the intensity of guiding force was increased. Therefore we do not know at what value the amount of improvement starts to decrease. Nevertheless, the result showed clearly that such improvements are less significant as the guiding force increases. This supports our hypothesis. Furthermore, we expected a leveling-off effect in terms of improvements to our performance after the guiding force reaches a certain threshold value.

Hypothesis H2: The number of errors for $S = 0.05$, 0.2, 0.35, and 0.5 was 108, 35, 18, and 17. The number of errors reduced by 73, 17, and 1 as S increased from 0.05 to 0.2, 0.2 to 0.35, and 0.35 to 0.5. This clearly supports our hypothesis that the reduction of the error rate reduced as the intensity of guiding force exceeded a threshold value. Unfortunately, the results do not allow us to identify the threshold. Similar to the improvement of task completion time, we expect the improvement of error rate to level off once the intensity of guiding force exceeds a certain threshold.

Hypothesis H3: Fig. 3 shows that the participants made fewer errors as the intensity of the haptic guidance increased. Eighty percent of the errors were made for $S = 0.05$ and $S = 0.2$. For these values of S, the variables A and W significantly affected error number. This effect was, however, noticeably weaker once S increased to 0.35 and 0.5. This supports our hypothesis that, with stronger haptic guidance, the characteristics (A and W) of a tunnel have less influence on the number of errors.

For some applications, for example tele-haptic surgery, driving a car [19], or learning handwriting [15], it is desirable for users to keep their movement as close as possible to the center of a path. We thus measure the effects of S, A, and W on the deviation distance between the ideal central path and user trajectory.

A repeated-measures analysis of variance revealed a significant effect of S ($F_{3,\,13831} = 1689$, $p < 0.001$). Interestingly, there was neither an effect of W ($F_{3,\,13831} = 4.1$, $p > 0.05$) nor of A ($F_{3,\,13831} = 2$, $p > 0.05$). This indicates that stiffness is the only factor that influences deviation (see Fig. 4). A post-hoc Tukey-Kramer analysis showed that deviations decreased significantly as the intensity of haptic guidance increased (all $p < 0.001$). Similar to the improvements with movement time and error rate, our results showed that the improvement of deviation decreased as the force intensity increased.

Finally, to validate the proposed model in Equation 9, we used a non-linear regression to estimate the value of η. We found it to be 19.619. The *index of difficulty* (ID) was again measured in pixels. Fig. 5 shows the movement time as a function of the proposed ID. A linear regression analysis gave a model fit with $R^2 = 0.98$. The equation for MT is given by:

$$MT = 32.59 + 64.69 \times A / \left(W + 19.619 \times S\right) \tag{11}$$

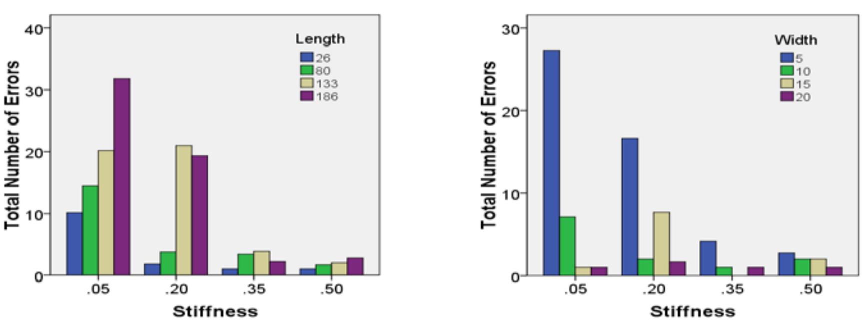


Fig. 3. Number of errors as a function of guiding forces

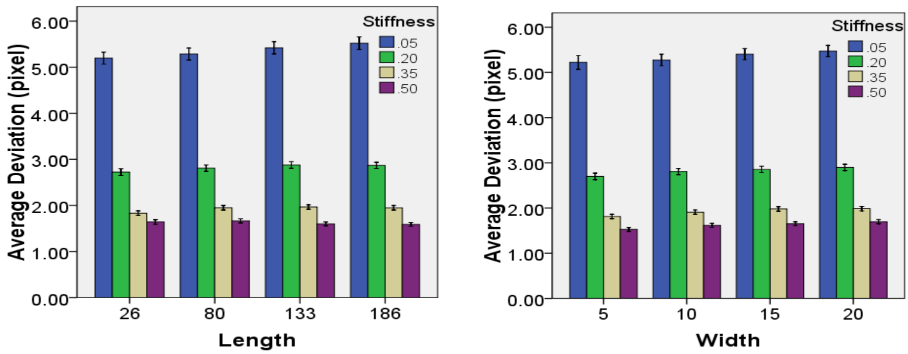


Fig. 4. Average deviation distance as a function of guiding force. Standard errors are also shown.

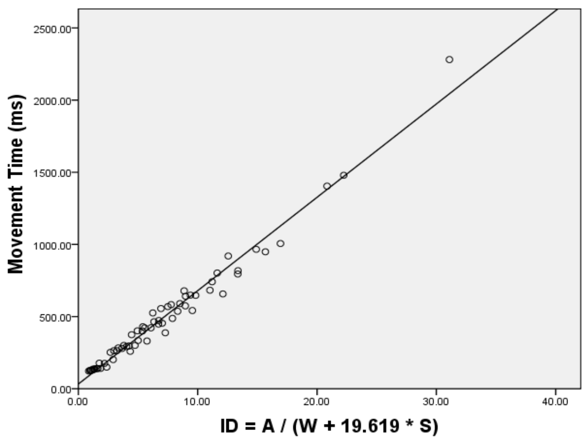


Fig. 5. Movement times by ID

The average guiding force for S = 0.05, 0.2, 0.35, and 0.5 were 0.27N, 0.56N, 0.67N, and 0.81N. Paired t-tests showed that the force magnitude increased significantly as S increased (all $p < 0.001$). We found it is also possible to model the movement time by replacing S by its corresponding force magnitude. A non-linear regression method produced an estimated value for η of 12.769. The equation for MT is as follows ($R^2 = 0.94$):

$$MT = -16.548 + 84.462 \times A / \left(W + 12.769 \times F\right) \tag{12}$$

8 Discussion

8.1 Leveling-Off Threshold

The results of our study showed that the intensity of force-guidance affects the speed of a steering task such that the stronger the guiding force the faster users can perform

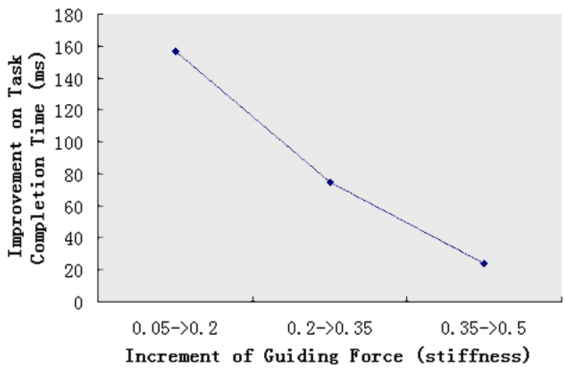


Fig. 6. Improvement on task completion time decreases as the strength of guiding force increases

the task. However, as expected, speed improvement does not grow linearly with the intensity of guiding force. It decreases as the force intensity exceeds a threshold value (see Fig. 6). This indicates that a certain a guiding force *G* is "sufficient" for assisting users to perform quickly on a haptic steering task. Increasing the magnitude of the guiding force beyond *G* does not necessarily improve task completion time. In real-world applications, we commonly see that designers employ the maximum output of the force feedback device as the force-guidance in trajectory-based tasks [17]. Our results, however, demonstrate that this is not necessary for achieving optimal performance. Therefore, being aware of the leveling-off at *G* is beneficial for designing effective force enhanced interfaces. This will be further investigated in our future work.

Note that *G* may vary based on the difficulty of a steering task. In general, a steering task is composed of several sub-steering tasks. Hence the difficulty of the task can be considered a combination of the difficulties of the sub-tasks. For instance, the difficulty of steering through an L-shape path can be considered as the combination of the difficulty of steering through a corner and steering through two straight paths. Given that the difficulties of steering through a corner and steering through horizontal and virtual paths are different [6, 12], a single *G* does not work for this case. For this reason, it is necessary to be aware of the values of *G* at different parts of a steering task, and to adjust the intensity of guiding force accordingly to achieve the best performance.

8.2 Errors and Deviations

As expected, error rates to decrease as the guiding force gets stronger. With a sufficiently strong force, errors can thus be reduced. This explains the finding that, with strong haptic guidance, the length and width of a tunnel have less influence on the error rate. Tasks such as selecting a menu item from a cascaded menu or selecting a line of a sentence in a text editing software requires the mouse cursor to be moved inside a tunnel. Failure to do so causes the task to fail. Our finding suggests that if the index of difficulty (*A*/*W*) is high, a strong guiding force is preferred. Generally, larger forces are better, but this is not the case for movement time.

Unlike general tunneling tasks, tracing a trajectory (e.g. learn to write strokes of Chinese words) requires hand movement to stay within a tunnel defined by a tolerance region while keeping the deviation as small as possible. The results showed that reducing the index of difficulty (*A*/*W*) does not necessarily make the task easier. Therefore, tracing a long path with haptic guidance may not be more difficult than tracing a short path.

We demonstrated that the proposed model can be used to predict movement time of a force-enhanced steering task. The model suggests that for passive constraint guiding, increasing the guiding force has the same effect as increasing the width of the tunnel. Therefore, if for some reasons the width of a tunnel needs to be reduced, e.g. as a result of increasing the number of menu items, the guiding force should be increased to achieve similar performance levels. Conversely, if the width of a tunnel is increased, the guiding force can be reduced.

9 Conclusion and Future Work

We presented and empirically validated a model for force-enhanced steering tasks. We derived the model from the well-known steering law [1] with the addition of the guiding force. In two experiments, we demonstrated that the proposed model can be used effectively to predict movement time for steering tasks when force-guidance is provided. The model suggests that providing guiding force orthogonal to hand movement has the same effect as enlarging the width of the tunnel. We explored user performance with respect to the intensity of the guiding force. Results show that the effects of force-guidance on movement time reduced as the intensity of guiding force increased. This suggests that the maximum force of a force-feedback device is not necessarily be the most suitable for achieving the best performance.

Future work will focus on producing design guidelines for force-enhanced interaction for specific tasks. We are also interested in finding a model for the tasks where force-guidance is provided in an active manner, i.e. when a user's hand is haptically led by the guiding force through an ideal trajectory.

References

1. Accot, J., Zhai, S.: Beyond Fitts Law: Models for trajectory-based HCI tasks. In: ACM CHI, pp. 295–302 (1997)
2. Accot, J., Zhai, S.: More than dotting the i's foundations for crossing-based interfaces. In: ACM CHI, pp. 73–80 (2002)
3. Accot, J., Zhai, S.: Refining Fitts Law models for bivariate pointing. In: ACM CHI, pp. 193–200 (2003)
4. Ahlström, D.: Modeling and improving selection in cascading pull-down menus using Fitts' law, the steering law and force fields. In: ACM CHI, pp. 61–70 (2005)
5. Campbell, C., Zhai, S., May, K., Maglio, P.: What You Feel Must Be What You See: Adding Tactile Feedback to the Trackpoint. In: INTERACT, pp. 383–390 (1999)
6. Dennerlein, J., Martin, D., Hasser, C.: Force-feedback improves performance for steering and combined steering-targeting tasks. In: ACM CHI, pp. 423–429 (2000)
7. Forsyth, B., MacLean, K.: Predictive Haptic Guidance: Intelligent User Assistance for the Control of Dynamic Tasks. IEEE Transactions on Visualization and Computer Graphics 12(1), 102–113 (2006)

8. Kattinakere, R., Grossman, T., Subramanian, S.: Modeling steering within above-the-surface interaction layers. In: ACM CHI, pp. 317–326 (2006)
9. MacKenzie, S.: Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction 7(1), 91–139 (1992)
10. MacKenzie, I., Buxton, W.: Extending Fitts' law to two-dimensional tasks. In: ACM CHI, pp. 219–226 (1992)
11. Mensvoort, K., Hermes, D., Montfort, M.: Usability of optically simulated haptic feedback. International Journal of Human-Computer Studies 66(6), 438–451 (2008)
12. Pastel, R.: Measuring the difficulty of steering through corners. In: ACM CHI, pp. 1087–1096 (2006)
13. Rodgers, M., Mandryk, R., Inkpen, K.: Smart sticky widgets: Pseudo-haptic enhancements for multi-monitor displays. Smart Graphics, 194–205 (2006)
14. Sensable Technologies (1993), `http://www.sensable.com`
15. Teo, C.L., Burdet, E., Lim, H.P.: A robotic teacher of Chinese handwriting. In: HAPTICS, pp. 335–341 (2002)
16. Webster, R.W., Zimmerman, D.I., Mohler, B.J., Melkonian, M.G., Haluck, R.S.: A Prototype Haptic Suturing Simulator. Medicine Meets Virtual Reality, 567—569 (2001)
17. Yang, X.D., Bischof, W.F., Boulanger, P.: Validating the performance of haptic motor skill learning. In: HAPTICS, pp. 129–135 (2008)
18. Yang, X.D., Bischof, W.F., Boulanger, P.: Perception of haptic force during hand movements. In: IEEE International Conference on Robotics and Automation, pp. 129–135 (2008)
19. Zhai, S., Accot, J., Woltjer, R.: Human action laws in electronic virtual worlds: an empirical study of path steering performance in VR. Presence: Teleoperators and Virtual Environments 13(2), 113–127 (2004)

Designing Laser Gesture Interface for Robot Control

Kentaro Ishii[1], Shengdong Zhao[2,1], Masahiko Inami[3,1],
Takeo Igarashi[4,1], and Michita Imai[5]

[1] Japan Science and Technology Agency, ERATO, IGARASHI Design Interface Project,
Frontier Koishikawa Bldg. 7F, Koishikawa 1-28-1, Bunkyo-ku, Tokyo, Japan
[2] Department of Computer Science, National University of Singapore
[3] Graguate School of Media Design, Keio University
[4] Graduate School of Information Science and Technology, The University of Tokyo
[5] Faculty of Science and Technology, Keio University
kenta@designinterface.jp, zhaosd@comp.nus.edu.sg,
inami@designinterface.jp, takeo@acm.org,
michita@ayu.ics.keio.ac.jp

Abstract. A laser pointer can be a powerful tool for robot control. However, in the past, their use in the field of robotics has been limited to simple target designation, without exploring their potential as versatile input devices. This paper proposes to create a laser pointer-based user interface for giving various instructions to a robot by applying stroke gesture recognition to the laser's trajectory. Through this interface, the user can draw stroke gestures using a laser pointer to specify target objects and commands for the robot to execute accordingly. This system, which includes lasso and dwelling gestures for object selection, stroke gestures for robot operation, and push-button commands for movement cancellation, has been refined from its prototype form through several user-study evaluations. Our results suggest that laser pointers can be effective not only for target designation but also for specifying command and target location for a robot to perform.

Keywords: laser pointer, stroke gesture, human-robot interaction, user interface.

1 Introduction

Robot technology has now advanced to the extent that many companies produce robots that not only can carry out tasks in industrial factories, but also can support everyday human activities. Today's commercial robots can already perform some of the tasks required in the home, such as vacuuming rooms [1].

These robots, however, are still limited to pre-programmed functionalities, and many important issues must be addressed to introduce more capable robots into home environments. One such issue is the question of how to give unique instructions to specify the types of robot operation, target objects, and target location. For example, when a user wants a robot to carry laundry to the veranda, the user has to give an instruction that the operation is delivery, the target object is laundry, and the target location is veranda.

A laser pointer provides a powerful alternative as a user interface for robot control. It is cheap, lightweight, points accurately, and is reliably recognized. Kemp *et al.* and

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 479–492, 2009.

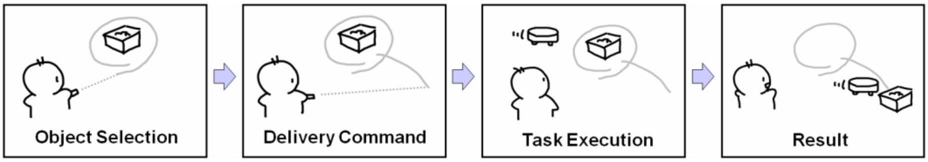


Fig. 1. Robot Control using Laser Gestures

Suzuki *et al.* showed that a laser pointer is useful for providing target designation for a picking-up robot [2] or a button-pushing robot [3]. However, target designation on its own is far from sufficient for a human-robot interface. The user also needs to be able to specify what to do and where to do with the target (e.g., *deliver* the laundry *to veranda*, *collect* the toys *in corner of the room*).

We extended the idea presented by Kemp *et al.* [2] and Suzuki *et al.* [3] to use a laser pointer not only for target designation but also for specifying command and target location. To achieve this, we applied gesture recognition to the laser trajectory. Fig. 1 illustrates our approach. After the user makes a gesture using the laser pointer to indicate the target object and desired command, the robot performs the task specified by the gesture. With our user interface, the user can command a robot to deliver, collect, and trash objects. The user can also instruct the robot to cancel, pause, undo, and redo a task.

This paper reports on our experience in designing an effective laser gesture interface. Using a vision-based tracking system by ceiling-mounted cameras and a remote-controlled iRobot Create, we developed a prototype robot system. We iteratively refined laser gesture sets by repeatedly testing them with first-time users. The final design of our system includes lasso and dwelling gestures for object selection, stroke gestures for robot operation, and push-button commands for movement cancellation.

The remainder of this paper is organized as follows. Section 2 describes previous work on robot control, and the use of laser pointers, and stroke gesture interfaces. Section 3 introduces the prototype system we built to test the laser gesture interface. Section 4 describes the initial design of our laser gesture interface. Section 5 describes the details of our system modification process that involves four rounds of user testing and design revision. Section 6 discusses possibility and limitations of our user interface, and Section 7 concludes the paper with a brief summary.

2 Related Work

2.1 Robot Control

Previous work on interfaces for robot control can be roughly divided into two approaches: (1) direct manual control and (2) as-autonomous-as-possible. A typical example of the former is joystick control. The user continuously gives low-level instructions to the robot, such as to move forward and turn left. This approach might be necessary for performing complicated tasks in a remote location (e.g., extraterrestrial robots). The other approach is the as-autonomous-as-possible method. Typically, these methods include the use of speech- and hand-gesture recognition. These methods require only a very abstract instruction from the user (e.g., "clean here") and the system automatically performs the task. While this might sound ideal, these high-level instructions are inherently ambiguous and often lack the robustness required for practical use.

There have been attempts to combine the two methods described above, creating systems that are more abstract than a direct form of control, but more concrete and specific than using speech and hand gestures. Perzanowski *et al.* provided multimodal interface in which the user can specify the target location by tapping a map on a PDA screen [4]. Lundberg *et al.* also proposed a PDA interface for a robot, where the user can specify the area for the robot to explore [5]. In this study, we also follow the approach to combine user control and robot autonomy.

2.2 Laser Pointer

Laser pointers can be used in several situations. Their most common use is as a pointing device to aid communication between a speaker and an audience during a presentation. Since a laser pointer can precisely point to a target location on a screen from a distance, the audience knows at which feature to pay attention. Taking advantage of this characteristic of a laser pointer, Yamazaki *et al.* developed a laser-pointing device that could be operated remotely called a GestureLaser [6]. Using a mirror with an actuator and a camera, a remote operator can give instructions while checking the actual pointing position on a remote camera image.

A laser pointer can also be used as a computer input device. For ordinary computer application, the user can use a laser pointer instead of a mouse to perform actions such as point-and-click or drag-and-drop [7], [8]. For robotic application, Kemp *et al.* used a laser pointer to tell a robot which object to pick up [2]. Using a laser pointer and a mobile robot, the user can pick up an object that is some distance away because the laser points over long distances. Their prototype system enables the user to "fetch" an object while remaining seated. Suzuki *et al.* also used a laser pointer to tell a robot which button to push [3]. With a flashlight, the user can specify an approximate direction for the robot to turn its face. After that, the user can designate the target button using the laser pointer, and the robot moves toward the button accordingly. Contrary to these two works, our interface uses a laser pointer not only for target designation but also for specifying command and target location.

2.3 Stroke Gesture Interface

Stroke gestures are used to input commands for various computer applications. The most significant advantage of using stroke gestures to input commands is that the user can specify several kinds of commands using just a simple pointing device. The idea of marking menus with a pointing device [9] is now widely used as a common method to input commands in applications such as web browsers.

Sakamoto *et al.* applied stroke gesture recognition to robot instructions [10]. They used four cameras capturing top-down images of a room, which the user can monitor during instruction. Skubic *et al.* developed robotic interface by free hand stroke [11]. This interface displays sensory information and robot status such as driving or not.

Igarashi *et al.* also proposed a similar technique to navigate a virtual avatar [12]. The user can navigate the avatar by drawing its path to walk through.

Our interface supports command specification based on stroke gesture recognition as well as marking menus [9] and Sketch and Run [10]. However, it is done by using a laser pointer.

3 Prototype System

In our prototype system, the user draws gestures on the floor using a laser pointer and the system executes the specified robotic task accordingly. This implementation supports three tasks: delivery, collecting, and trashing. All tasks are performed by pushing objects. The delivery task simply moves the selected objects from one place to another. The collecting task clusters the selected objects into a specified location. The trashing task moves the selected objects to a predefined trash box location. Fig. 2 shows the working environment for the prototype system.

Fig. 2. Prototype Environment

Fig. 3. Laser Pointer

Fig. 4. iRobot Create

3.1 Hardware Configuration

The frontend system consists of a laser pointer held by the user and a robot on the floor that execute the tasks. The laser pointer is an off-the-shelf product with two extra buttons for moving forward/backward through a slideshow presentation (Fig. 3). The task execution robot is a developer's version of the popular home vacuuming robot (Roomba), called iRobot Create [1] (Fig. 4). It comes with a serial interface for receiving control commands. We use a Bluetooth-serial adapter to control the robot wirelessly.

The backend system consists of three cameras, a speaker, a projector, and two computers (Fig. 5). The three cameras are fixed on the ceiling and cover 4m x 3m viewing area. The speaker and projector are used to provide feedback to the user. The two computers are connected via a TCP/IP network. One computer is used for robot and object tracking and the other computer is used to recognize laser gestures, drive the robot, and generate audio-visual feedback to the user.

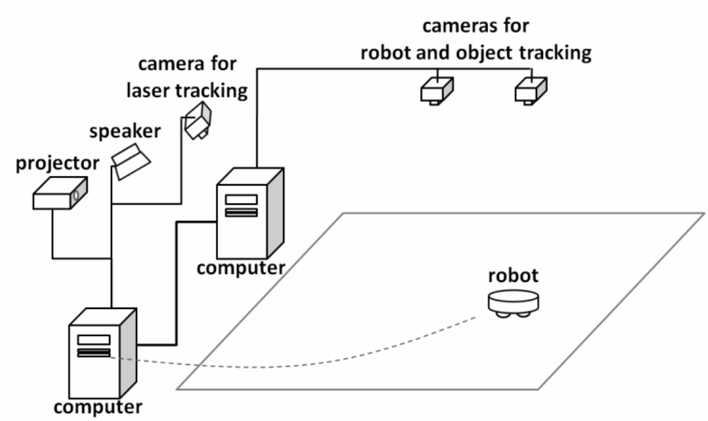


Fig. 5. Hardware Configuration

3.2 Laser Tracking

We use a simple web camera with a narrow band-pass filter for laser tracking. The narrow band-pass filter passes only light with a wavelength of the laser, shutting out all other wavelengths. Since the narrow band-pass filter passes only laser light, laser tracking is simply realized by tracking the most significant pixel on the image provided by the camera.

3.3 Robot and Object Tracking

Our system uses vision-based tag identification for robot and object tracking. 3 x 3 (9bit) 2D markers are used for visual tag tracking. An example of this design is illustrated in Fig. 6 left. In order to make the tags small, we use a proprietary tag tracking system because we are only concerned about the 2D robot motion on the floor, contrary to existing marker-based technology that can capture 3D postures, such as CyberCode [13], ARTag [14], and ARToolKit [15]. Fig. 6 right shows the results of our robot and object tracking. Each of the two cameras calculates the ID, position, and orientation of the existing tags. The coordinate system of each camera is manually calibrated with each other.

Fig. 6. Vision-based ID Tag and Tracking Results

It is important to note that visual tags may not be ideal for practical deployment because it is tedious initially to put visual tags on all objects in an environment. We envision that this will be replaced in the future by another registration-free technique such as image-based segmentation using overview camera, or probing the object by sensors on the robot.

3.4 Robot Control

Using the information collected by the tag tracking system, the robot control module controls the robot based on the results from the gesture recognition module. Since all tasks are comprised of pushing operations, the robot control module creates a path procedure and sends low-level control commands such as “move forward” and “spin left.” To complete an atomic pushing operation, the robot first moves behind the target object, adjusts its orientation toward the target location, and then moves to the target location. The system also instructs the robot to avoid other objects during its movement, tracking the robot and object positions during operation. In case the object slips off from the robot during pushing, the system orders the robot to retry from moving behind the object.

3.5 Audio-Visual Feedback to the User

A sound speaker is placed at the corner of the prototype environment for the user to receive sound feedback. Sound feedback works to understand gesture recognition results. In addition, we use poles to place a projector near the ceiling, projecting toward the floor (Fig. 7). The projector gives visual feedback on the laser pointer trajectory while the user is drawing a stroke. Visual feedback enables the user to understand laser tracking result (Fig. 8).

Fig. 7. Projector on Ceiling

Fig. 8. Visual Feedback on Laser Trajectory

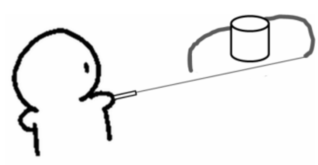


Fig. 9. Object Selection

4 Initial Laser Gesture Design

4.1 Instruction Procedure

The following steps summarize the robotic instruction procedure:

1. Select Object
2. Input Gesture Command

Since all tasks require target objects, the user needs to specify the target objects before inputting a command.

Object selection in initial design is performed by a lasso gesture (Fig. 9). We choose lasso gesture because the user can select single or multiple objects in the same manner.

4.2 Command Gestures

We designed command methods as single-stroke gestures. Fig. 10 shows stroke gestures representing delivery, collecting, and trashing, as well as the gesture to cancel a selection. We consulted the existing works on stroke gesture interfaces [16], [17], [18] to design our initial set of stroke gestures.

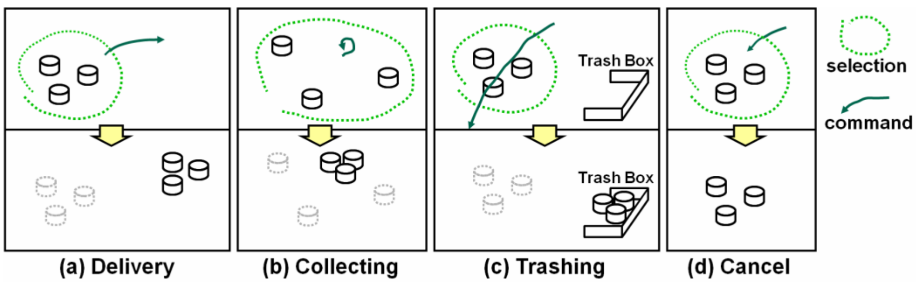


Fig. 10. Command Gestures

The delivery command tells the robot to move the selected objects to a target location. When the user draws a line from inside the selection lasso to another area, the system interprets it as a delivery command and the robot moves the objects to the end of the line (Fig. 10(a)). The collecting command tells the robot to collect the selected objects at the target location inside the selection lasso. When the user draws a circle inside the selection lasso, the system interprets it as a collecting command and the robot collects the objects in the center of the circle (Fig. 10(b)). The trashing command tells the robot to throw selected objects into a predefined trash box. When the user draws a stroke across the selection lasso, the system interprets it as a trashing command, and the robot moves the objects to the trash box (Fig. 10(c)). We also provide a cancel command. When the user draws a line across the selection lasso from outside to inside, the system interprets it as a cancellation of the selection and returns to the initial state (Fig. 10(d)).

4.3 Task Control Gestures

We also designed task control gestures so that users could adjust the execution of tasks after having given some commands. Using task control gestures, users can cancel, pause, undo, and redo a task (Fig. 11).

Task cancellation uses the same gesture as selection cancellation (Fig. 11(a)). When a task is cancelled, the robot stops performing the task immediately and leaves the objects as they are. Pause is used to suspend a robot's activity temporarily (e.g., to

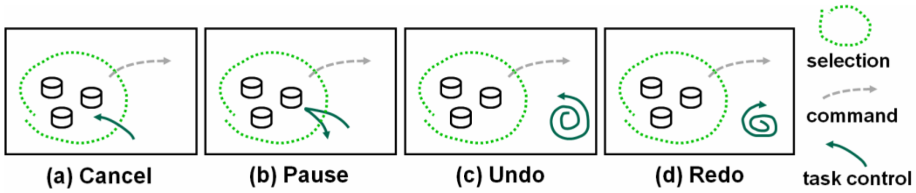


Fig. 11. Task Control Gestures

stop robot noise when answering a phone call). If the stroke enters the selection lasso and returns to the outside, the system pauses in the current task (Fig. 11(b)). After the pause, the robot resumes its task when the same gesture is recognized. Since our application only changes the position of real-world objects, the system realizes undo by restoring the object position from task history. If the stroke forms a double counter-clockwise circle, the system starts to undo the previous task (Fig. 11(c)). If no task history is stored, the system does not recognize the undo command. Redo is an inverse function of undo. The user can re-order the robot to redo the task that it has been instructed to undo. If the stroke forms a double clockwise circle, the system starts to redo the previously undone task (Fig. 11(d)). If there is no undone task, the system does not recognize the redo command.

4.4 Design of Audio-Visual Feedback

To ensure that the user understands how the system recognizes each gesture, it is important that users receive appropriate feedback. In addition to drawing the laser trajectory as described in section 3.5, the system emits sound and changes the color of the trajectory when the system recognizes a stroke gesture. The system emits a different sound for selecting objects, and for each command and task-control gestures. The trajectory color is white before recognition, and yellow after recognition.

5 Iterative Refinement Based on User Study

5.1 Overview

The interface introduced in the previous section was tested and refined by first-time users. We invited 10 participants (9 females, 1 male; 3 single, 7 married, 2 with kids; age range: 27 to 39 years old, mean: 31) who identified as having a non-science educational background, as we assumed that home users would be the target audience of our system. We had 3 participants in the first round, 3 participants in the second round, 2 participants in third round, and 2 participants in the fourth round. Although we had a small number of participants in each round, we intended to evaluate the interface qualitatively rather than quantitatively.

In each round, participants could only refer to a user's manual for instructions on how to use the system, and could not ask questions to the experimenter. We took this policy because in the real world, the home user can only refer to a user's manual of a product for information on the system.

Each participant was given the following tasks to accomplish in ten minutes:

- Move a trash box to the sofa where the participant is sitting.
- Cluster four toy boxes, which are placed around the room.
- Move a cardboard to trashing area.

It should be noted that we did not specify the destination for the cluster in the second task. This may have caused various completion times among the study participants.

The time restraint was set on the belief that real-world users tend to give up on a product that cannot be understood in less than ten minutes. Since it took about two minutes for experts to finish these tasks, the participants could have the remaining eight minute to understand system behavior.

The following describes the course of each user study:

1. The experimenter explains the aims of this study and user tasks to the participant.
2. The experimenter gives the user's manual to the participant.
3. After the participant has read the user's manual, the experimenter starts 10 minute trial session. If the participant finishes all tasks before 10 minutes, the experimenter stops the trial session. During the trial session, the experimenter records the session by videotape.
4. After the trial session, the participant fills out a questionnaire and the experimenter interviews the participant.

5.2 The First Round

The first round used the initial design of laser gesture interface described in Section 4. Specifically, all instructions were given as stroke gestures. All visual feedback was given as lines that indicate the trajectory of laser movement, which color changes from white to yellow when recognized.

It turned out that this initial design caused a log of failures. None of the three participants in this round reported to understand the system behavior. We observed they failed to complete any of the tasks, being confused and frustrated. The interview revealed the most fundamental problem with the initial design was that the participants failed to cancel the robot actions. We often observed the participants panicked when the robot started to move against their intention and they were unable to stop it. The interview also revealed the visual feedback was problematic. The initial design did not distinguish the color of the first stroke (object selection) and the second stroke (command). The participants often got confused if the system was in the first or second step.

5.3 The Second Round

Based on the results in the first round, the following revisions were made to the system:

- We changed cancellation command from a gesture to a secondary button (originally designed for returning to a previous presentation slide) on the laser pointer.
- We changed the color of the strokes in the visual feedback. The object selection stroke is shown in yellow and command stroke is shown in blue in the revised design.

The result of the second design proved to be more successful than the first. Of the three participants, two reported that they managed to understand the system behavior, with one of the two successfully completing all the three tasks in 8m33s. We observed that the easy cancellation greatly contributed to this improvement. The participants repeatedly cancelled their mistakes and tried again, which helped steady learning. The third user, who was unable to finish the tasks, did not understand the system, and tried to input command gestures before object selection.

In the interview, the participants reported a problem with the second design. They found it difficult to make the “delivery command”, which required the user to start a stroke inside of the first lasso, because the laser did not appear until they pressed the laser emitting button. From this, we learned that in general, stroke gestures using a laser pointer are ineffective at “starting” at a specific position because the pointer is invisible in the beginning.

5.4 The Third Round

Based on the results in the second round, the following revisions were made to the system:

- Since it was difficult for the user to start a command stroke inside of the object selection lasso, we decided not to take the starting point of a stroke into account. A closed stroke is recognized as object collection regardless of the position (Fig. 12(b)). An open stroke that crosses the selected objects is recognized as trashing (Fig. 12(c)), and other open strokes are recognizes as object delivery (Fig. 12(a)). In addition, pause command of task control gestures was replaced by using the third button (originally designed for forwarding a presentation slide) on the laser pointer.
- We revised the visual feedback of the object selection. The previous design showed the trajectory of the object selection lasso drawn by the laser pointer. The revised version simply shows a spot light on the selected target objects (Fig. 12). This helps the user to distinguish object selection and command specification phases.

Both participants successfully understood the system behavior and completed all the three tasks in 6m10s and 9m30s respectively. As with the second round study, we observed the participants frequently used cancellation button on the laser pointer to cancel selection and stop unexpected executions. The participants in the third round achieved the study goal that the user gets accustomed the system behavior in under ten minutes.

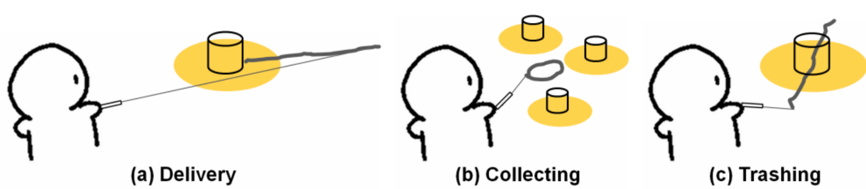


Fig. 12. Command Gestures after Object Selection (Revised for Third Round)

Although this study was successful, there were still areas of improvement for our system. One participant, who finished the tasks earlier, commented that lassoing for object selection was somewhat bothering. The participant suggested using dwelling behavior of laser to select an object, and commented that direct dwelling gesture can be more intuitive selection method than lasso gesture.

5.5 The Fourth Round

Based on the results in the third round, the following revisions were made to the system:

- We added a direct dwelling gesture on the object for object selection in addition to lasso selection. After making one second dwelling gesture, the object under the laser is selected. The user can select a single object directly and quickly. After lasso selection, the user also can add an object using another dwelling selection. In addition, we added a function to remove an object from the selected objects by using a dwelling gesture. In other words, the user can toggle the selection status of an object.

Both two participants successfully understood the system behavior and completed all the three tasks in less than ten minutes. Both participants finished their tasks earlier than the earliest participant in the third round, with 3m54s and 5m22s respectively.

5.6 The Final Design and Execution Sequence

In our final design, object selection is realized by lasso and dwelling gestures. If the user draws a lasso in the initial state, the system selects objects in the lasso. The user can also use dwelling gesture to select and remove an object. The system provides three commands for object control and distinguishes them by stroke gesture

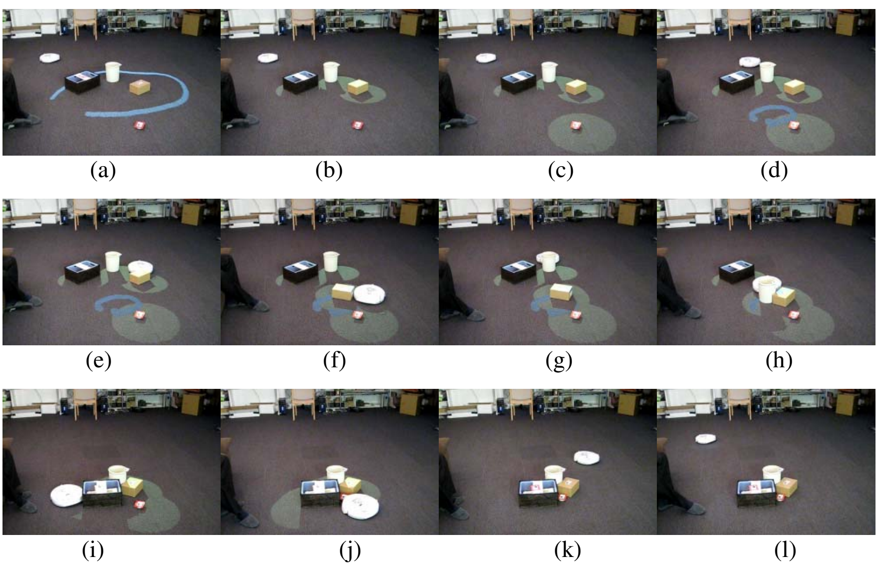


Fig. 13. Execution Sequence

recognition. Closed strokes are recognized as object collection. Open strokes that cross the selected objects are recognized as trashing and other open strokes are recognized as object delivery. The user can cancel or pause by pushing the extra buttons on the laser pointer.

Fig. 13 shows an execution sequence of the system. The user first selects the upper three objects using lasso selection (Fig. 13(b)). After that, the user adds the lower object using dwelling selection (Fig. 13(c)). When the user inputs the collecting command, the robot starts the task to collect the objects that have been selected (Fig. 13(d)). If the robot finishes the task, the visual feedback turns off (Fig. 13(k)) and the robot returns to its initial position (Fig. 13(l)).

6 Discussion

Through our user study, we observed the extent that a laser pointer can be used as a robot-controlling interface, and concluded that the laser pointer is good at staying at a position or jumping to another position, but not good at starting a stroke at a specific position. Starting a stroke at a specific position was one of the main difficulties in making command and task-control gestures in our initial design. However, we observed that it is still possible to use gestures to distinguish commands if the system is only concerned with the middle or end position of the path, as verified in the successful completion of tasks in the third and fourth round.

The study also showed that visual feedback is essential, especially for the first-time users. If it is not well designed, the user cannot understand the behavior of the system. Although visual feedback is essential, the light provided by the projector potentially has an occlusion problem. Designing another occlusion-free feedback is one of the future modifications in this system.

Easy cancellation methods are important for robot applications. Since the robot affects the objects in the real world, it is necessary to have an emergency stop button. In the initial design, we did not provide this feature. In the initial and second design, the users often made mistakes in specifying commands due to the difficulty in starting at a specific position. In such cases, easy cancellation especially works. Easy cancellation also helps first-time users learn the system behavior faster. In the final design, we do not use the stroke gesture interface for cancellation; however, this interface still works for specifying other commands as well as designating targets.

7 Conclusion

This paper explored the possibility of using a laser gesture interface to give instructions to a mobile robot. The user makes a gesture using a laser pointer, specifying target objects, commands, and target locations. The system then recognizes the gesture and executes the specified task by driving the robot. We implemented a prototype system by combining a robot moving on the floor and cameras on the ceiling. Using the system, we iteratively improved the interface design by conducting user studies. The results revealed the importance of easy and reliable cancellation methods, the difficulty of starting a gesture at a specific location, and the importance of understandable visual feedback. Based on the results of iterative studies, we added the dwelling gesture for object selection, eliminated considering the starting point of strokes, changed the laser gesture cancellation function to a push button function, and revised the visual feedback from stroke-based to status-based.

The main limitations of the current implementation are in using vision-based tags for object tracking and visual feedback occlusion problem. To overcome these limitations, our future work includes object tracking by other registration-free techniques, such as image-based segmentation using an overview camera, or probing objects by sensors on the robot; and designing occlusion-free feedback.

References

1. iRobot Corporation: Home Page, `http://www.irobot.com`
2. Kemp, C.C., Anderson, C.D., Nguyen, H., Trevor, A.J., Xu, Z.: A Point-and-Click Interface for the Real World: Laser Designation of Objects for Mobile Manipulation. In: 3rd ACM/IEEE International Conference on Human-Robot Interaction, pp. 241–248 (2008)
3. Suzuki, T., Ohya, A., Yuta, S.: Operation Direction to a Mobile Robot by Projection Lights. In: IEEE Workshop on Advanced Robotics and its Social Impacts, TPO-2-2 (2005)
4. Perzanowski, D., Schultz, A.C., Adams, W., Marsh, E., Bugajska, M.: Building a Multimodal Human-Robot Interface. IEEE Intelligent Systems 16(1), 16–21 (2001)
5. Lundberg, C., Barck-Holst, C., Folkeson, J., Christensen, H.I.: PDA interface for a field robot. In: 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 3, pp. 2882–2888 (2003)
6. Yamazaki, K., Yamazaki, A., Kuzuoka, H., Oyama, S., Kato, H., Suzuki, H., Miki, H.: GestureLaser and GestureLaser Car: Development of an Embodied Space to Support Remote Instruction. In: Sixth European Conference on Computer Supported Cooperative Work, pp. 239–258 (1999)
7. Olsen, D.R., Nielsen, T.: Laser pointer interaction. In: SIGCHI Conference on Human Factors in Computing Systems, pp. 17–22 (2001)
8. Kirstein, C., Muller, H.: Interaction with a Projection Screen using a Camera-Tracked Laser Pointer. In: International Conference on MultiMedia Modelling, pp. 191–192 (1998)
9. Kurtenbach, G., Buxton, W.: User Learning and Performance with Marking Menus. In: SIGCHI Conference on Human Factors in Computing Systems, pp. 258–264 (1994)
10. Sakamoto, D., Honda, K., Inami, M., Igarashi, T.: Sketch and Run: A Stroke-based Interface for Home Robots. In: 27th International Conference on Human Factors in Computing Systems, pp. 197–200 (2009)
11. Skubic, M., Anderson, D., Blisard, S., Perzanowski, D., Schultz, A.: Using a hand-drawn sketch to control a team of robots. Autonomous Robots 22(4), 399–410 (2007)
12. Igarashi, T., Kadobayashi, R., Mase, K., Tanaka, H.: Path Drawing for 3D Walkthrough. In: 11th Annual ACM Symposium on User Interface Software and Technology, pp. 173–174 (1998)
13. Rekimoto, J., Ayatsuka, Y.: CyberCode: Designing Augmented Reality Environments with Visual Tags. In: DARE 2000 on Designing Augmented Reality Environments, pp. 1–10 (2000)

14. Fiala, M.: ARTag, a Fiducial Marker System Using Digital Techniques. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 590–596 (2005)
15. ARToolKit Home Page, http://www.hitl.washington.edu/artoolkit/
16. Baudel, T.: A Mark-Based Interaction Paradigm for Free-Hand Drawing. In: 7th Annual ACM Symposium on User Interface Software and Technology, pp. 185–192 (1994)
17. Landay, J.A., Myers, B.A.: Interactive Sketching for the Early Stages of User Interface Design. In: SIGCHI Conference on Human Factors in Computing Systems, pp. 43–50 (1995)
18. Igarashi, T., Matsuoka, S., Tanaka, H.: Teddy: A Sketching Interface for 3D Freeform Design. In: 26th Annual Conference on Computer Graphics and Interactive Techniques, pp. 409–416 (1999)

A Haptic-Enhanced System for Molecular Sensing

Sara Comai and Davide Mazza

Politecnico di Milano
Department of Electronics and Information (DEI)
Piazza L. da Vinci 32
I-20133 Milan
{comai,mazza}@elet.polimi.it

Abstract. The science of haptics has received an enormous attention in the last decade. One of the major application trends of haptics technology is data visualization and training. In this paper, we present a haptically-enhanced system for manipulation and tactile exploration of molecules.The geometrical models of molecules is extracted either from theoretical or empirical data using file formats widely adopted in chemical and biological fields. The addition of information computed with computational chemistry tools, allows users to *feel* the interaction forces between an explored molecule and a charge associated to the haptic device, and to *visualize* a huge amount of numerical data in a more comprehensible way. The developed tool can be used either for teaching or research purposes due to its high reliance on both theoretical and experimental data.

1 Introduction

Molecular analysis has acquired a great importance in recent years. The study of interactions among molecules and the ways in which they can be aggregated are at the basis of drug design. A deep understanding of the *inter-molecular* forces that govern this kind of processes is fundamental. Interactions among molecules are usually described by *huge sequences of data*, which describe the attraction/repulsion forces and the positions of interesting sites for binding over molecular surfaces in a very detailed way, but awkward to interpret *as-is* even by scientists and chemists of the field. Haptic interaction could greatly help here, because the sense of involved forces could be used to better understand the intensity of interactions that humans are not able to feel in everyday life, due to their existence at nanoscale or atomic levels. In this way, scientists could reach a better awareness on the location and the nature of these interactions, that could be exploited in further experimental studies.

At this aim we developed a virtual environment, where the molecule to be analyzed is shown to the user and a *probing* charge is associated with a haptic device. The interaction between the molecule and the electric charge is felt via the haptic device and allows one to explore the electrostatic surface of the molecule. This paper is organized as follows: Section 2 provides an overview of the system and shows how information and results are rendered; Section 3 reports feedbacks obtained by users; Section 4 presents related works, and, finally, Section 5 draws our conclusions and future plans.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 493–496, 2009.

2 Description of the Haptic-Enhanced System

The developed system takes in input the geometrical representation of the molecule in terms of theoretical model (.PDB file format) or experimental data (.CIF file format, obtained from crystallography). For a reference on PDB and CIF file format see [5, 3]. Such data are then elaborated by a computational chemistry tool [4], which provides in output the electric field data needed to model the force interaction. Figure 1 sketches the architecture of the developed system.

Both input and output data are huge in size and it is therefore difficult to catch atoms position and electronic fields' interactions directly from such data. Figure 2 shows some examples of fragments of input and output data. Just to give an idea of the amount of information to deal with, the original files shown in Figure 2, referred to the ammonia molecule, are composed of 1508 and 22933 lines of numbers!

The developed system provides a virtual environment showing a 3D representation of the molecule. Different visualizations are possible, like, for example, space-filling and ball-and-stick. Beside providing the graphical representation, the system allows to *feel* the electronic fields' interactions by means of a haptic device (in our case, the Sensable PHANTOM® Omni [6]): the graphical environment shows the current position of the haptic proxy, to indicate which part of the space the user is interacting with. Auxiliary key information are also shown in an intuitive way, such as a plot of the electrostatic field along the direction connecting the proxy position and the center

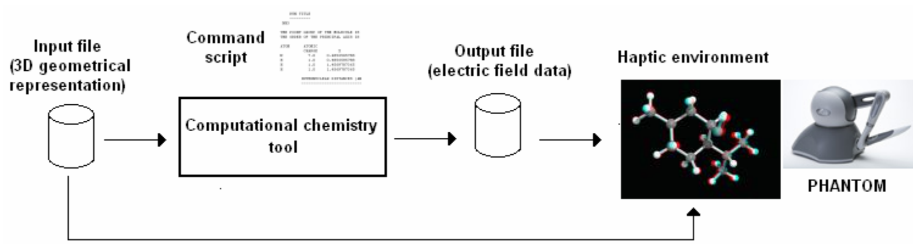
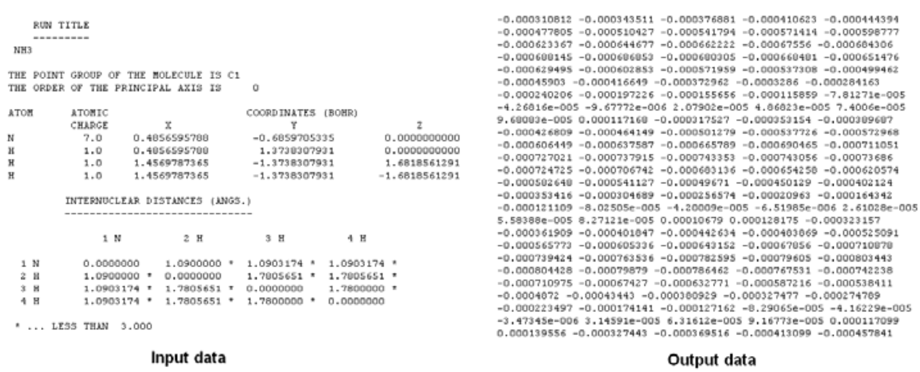


Fig. 1. The architecture of the designed haptic system



```
    RUN TITLE
    ---------
 NH3

THE POINT GROUP OF THE MOLECULE IS C1
THE ORDER OF THE PRINCIPAL AXIS IS     0

ATOM      ATOMIC                      COORDINATES (BOHR)
          CHARGE         X                   Y                   Z
N           7.0     0.4856595788      -0.6859705335        0.0000000000
H           1.0     0.4856595788       1.3738307931        0.0000000000
H           1.0     1.4569787365      -1.3738307931        1.6818561291
H           1.0     1.4569787365      -1.3738307931       -1.6818561291

       INTERNUCLEAR DISTANCES (ANGS.)
       ------------------------------

            1 N          2 H          3 H          4 H

  1 N    0.0000000    1.0900000 *  1.0903174 *  1.0903174 *
  2 H    1.0900000 *  0.0000000    1.7805651 *  1.7805651 *
  3 H    1.0903174 *  1.7805651 *  0.0000000    1.7800000 *
  4 H    1.0903174 *  1.7805651 *  1.7800000 *  0.0000000

 * ... LESS THAN  3.000
```

```
-0.000310812 -0.000343511 -0.000376881 -0.000410623 -0.000444394
-0.000477805 -0.000510427 -0.000541794 -0.000571414 -0.000598777
-0.000623367 -0.000644677 -0.000662222 -0.00067556 -0.000684306
-0.000688145 -0.000686853 -0.000680305 -0.000668481 -0.000651476
-0.000629495 -0.000602853 -0.000571959 -0.000537308 -0.000499462
-0.00045903 -0.000416649 -0.000372962 -0.0003286 -0.000284163
-0.000240206 -0.000197226 -0.000155656 -0.000115859 -7.81271e-005
-4.26816e-005 -9.67772e-006 2.07902e-005 4.86823e-005 7.4006e-005
9.68083e-005 0.000117168 -0.000317527 -0.000353154 -0.000389687
-0.000426809 -0.000464149 -0.000501279 -0.000537726 -0.000572968
-0.000606449 -0.000637587 -0.000665789 -0.000690465 -0.000711051
-0.000727021 -0.000737915 -0.000743353 -0.000743056 -0.00073686
-0.000724725 -0.000706742 -0.000683136 -0.000654258 -0.000620574
-0.000582648 -0.000541127 -0.00049671 -0.000450129 -0.000402124
-0.000353416 -0.000304689 -0.000256574 -0.00020963 -0.000164342
-0.000121109 -8.02505e-005 -4.20009e-005 -6.51985e-006 2.61028e-005
5.58388e-005 8.27121e-005 0.00010679 0.000128175 -0.000323157
-0.000361909 -0.000401847 -0.000442634 -0.000483869 -0.000525091
-0.000565773 -0.000605336 -0.000643152 -0.00067856 -0.000710878
-0.000739424 -0.000763536 -0.000782595 -0.00079605 -0.000803443
-0.000804428 -0.00079879 -0.000786462 -0.000767531 -0.000742238
-0.000710975 -0.00067427 -0.000632771 -0.000587216 -0.000538411
-0.0004872 -0.00043443 -0.000380929 -0.000327477 -0.000274789
-0.000223497 -0.000174141 -0.000127162 -8.29065e-005 -4.16229e-005
-3.47345e-006 3.14591e-005 6.31612e-005 9.16773e-005 0.000117099
0.000139556 -0.000327443 -0.000369516 -0.000413099 -0.000457841
```


Fig. 2. Abstracts from sample input and output data ¯les in case of ammonia (NH3) molecule

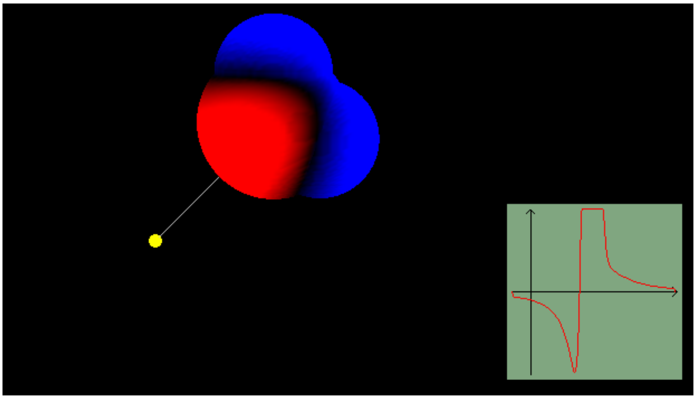


Fig. 3. A screenshot of the tool in case of water molecule. The proxy position (yellow circle) and the plot of the electrostatic field are visible. The plot is computed using the value of the field along the thin white line connecting the proxy with the center of the molecule.

of the molecule, and the force field on the molecule surface (represented by means of different colors). Figure 3 shows a snapshot of the tool in action. The 3D molecule representation and the proxy pointer are clearly visible. The plot of the electrostatic field is shown in the bottom-right corner.

The electrostatic field is represented by a quantized grid surrounding the molecule till a certain distance. The force felt by the user is the one associated with the *voxel* of the grid within which proxy position falls in. The forces are meant as opposed to probe movement done by the user and consists of attraction/repulsion forces according to the situation. Their intensity are naturally very weak (nanoscale). A human operator cannot feel them *as is*. After some usage tests, an amplification has been determined to scale the force appropriately without losing the proportions among real forces.

3 Users' Feedbacks

The tool has been tested by different kinds of users on different molecules (e.g., water, ammonia, methane). The students of a chemistry course tried the *experience* of the interaction as a didactic tool; chemical technicians appreciated the improved awareness of the phenomena they are used to work with, which are otherwise not so intuitive; researchers tried to test how the tool is able to render well-known interactions. We have tested the tool on molecules with known electrostatic behaviour and interesting binding sites, and checked with chemical experts the obtained behavior to confirm that the simulation is correct. In general, the most important benefit is represented by the possibility of combining the visualization of data with the rendering of the *feeling* of nano-scale/atomic interactions that can improve the understanding of real phenomena.

4 Related Works

[7] provides a survey on haptic rendering techniques: most of the haptic algorithms and applications focus on the rendering of surface tactile feedback, rather than *forces*

distributed in the whole space as this does. In the field of chemical visualization and haptics, our tool is a border-line one and stands along the boundary of old computational chemistry (which privileges awkward numerical outputs, disregarding user-friendliness) and the brand-new bioinformatics (which introduces better and friendly UIs to show results). Most of the developed tools, whenever presenting haptic interaction, exhibit it as a side feature, usually because it has been added later. In our tool it is the *core* interaction. Examples of similar tools include VMD, PyMol, and others (see [2] for a survey), which offer a wide range of representations, but interpretation of data is not so straightforward for non-technician of the field. [1] presents a similar tool, but it is strongly designed on theoretical models and, compared to our systems, it does not allow to use or explore empirical data directly obtained in laboratories.

5 Conclusions and Future Work

This work is meant as the first step towards the development of a framework for the simulation of the interaction between complex chemical compounds, such as a protein and another -whatever complex- molecule and is carried on in collaboration with the Chemical Department of our University. In particular the focus will be on molecules exhibiting *halogen bonding*, a kind of ligand known in literature since 60s but arisen again in recent years due to its importance in drug design. Further improvements include the implementation of interactions of pairs of halogenated compounds, either from a statical or dynamical point of view, to better understand the behavior of such bonding and to allow the usage of the tool in scientific/research frameworks.

Acknowledgements. We would like to thank Pierangelo Metrangolo (chemistry), Antonino Famulari (chemistry), Marzio Ghezzi (labs and device) and Marco Faverio (development) for their precious contributions.

References

1. Chemical force feedback, http://cff.itn.liu.se/public/
2. Chemistry visualization software survey, http://personal.cscs.ch/mvalle/ChemViz/tools.html
3. The cif file form, http://www.wrcad.com/manual/xicmanual/node493.html
4. The general atomic and molecular electronic structure system (gamess), http://www.msg.ameslab.gov/GAMESS/
5. Pdb file format specifications, http://deposit.rcsb.org/adit/docs/pdbatomformat.html
6. Sensable Phantom® Omni, http://www.sensable.com/haptic-phantom-omni.htm
7. Laycock, S., Day, A.: A survey of haptic rendering techniques. Computer Graphics forum 26(1), 50–65 (2007)

Designing with Only Four People in Mind? – A Case Study of Using Personas to Redesign a Work-Integrated Learning Support System

Amir Dotan, Neil Maiden, Valentina Lichtner, and Lola Germanovich

Centre for Human-Computer Interaction Design, City University London
Northampton Square, London, EC1V 0HB

Abstract. In this paper we describe and reflect on the use of personas to redesign the 3rd prototype of APOSDLE – a system to support informal learning and knowledge transfer in the workplace. Based on the results of a formative evaluation of the 2nd prototype we used personas to explore how users' goals, behaviour and preferences could be communicated to project members during a two-day design workshop, in order to ensure useful and usable design solutions. We actively involved stakeholders representing the target market throughout the process as they helped to create, validate and interpret the four personas we used. By doing so we aimed to address methodological weaknesses and practical limitations of using personas, primarily those relating to the validity of the personas used and the way they are interpreted. We reflected on how effective the personas were by referring to data generated during the workshop and discussion transcripts. As reported by others, and as we have experienced ourselves, using personas can be quite challenging as rich narrative descriptions are expected to produce insight and design solutions. In light of this challenge, we contribute a case study illustrating how personas were implemented in a real world situation to engage project members with user information and drive the design process. We specifically discussed the strengths and weaknesses of actively involving stakeholders in creating and using personas.

Keywords: Personas, User-Centred Design, Work-Integrated Systems, Interaction Design, Usability.

1 Introduction

APOSDLE (Advanced Process-Oriented Self-Directed Learning Environment) is a 48 month research and development project involving twelve European organisations from five countries. It aims to support work-integrated learning by providing people working in knowledge-intensive industries the infrastructure needed to acquire, understand and communicate knowledge. The system detects and analyses the user's computer activity in order to infer a task the user is working on and suggest relevant resources that might prove helpful. These resources are tailored to the user's 'work context' and profile information, which includes data such as inferred level of knowledge based on previous interactions with the system among other parameters. Following a formative evaluation of the 2nd system prototype, which highlighted

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 497–509, 2009.

various usability problems, it was decided to introduce personas in order to explore how user information could be integrated more effectively to inform the design of the 3rd and final prototype. Personas are a vivid 'day in the life' description of fictional individuals representing key user groups used to communicate their goals, behaviour and preferences in an engaging and practical way [9, 12]. Establishing a shared understanding of key user groups was especially important ahead of a two-day design workshop involving 21 project members from different professional backgrounds. While the results of the formative evaluation served as the main catalyst for the redesign and provided valuable insight into users' experience using APOSDLE, they did not communicate the goals and work environment of those users in a memorable and actionable manner. Personas were regarded as a potential way to communicate such information to participants so it could be effectively referred to during the discussions.

In this paper we describe how we used personas and reflect on how effective it was to design with only a handful of specific individuals in mind. By doing so we provide a real life account that might inform others wishing to use the tool. While literature focuses on analysing field data to identify and establish personas, we chose to complement such data by collaborating with stakeholders in order to create and use the four personas. We discuss the strengths and weaknesses of this collaborative approach by presenting data gathered and observed during a two-day design workshop. We begin with an overview of the main research work carried out as part of the APOSDLE project before the personas were introduced. This is followed by a review of relevant literature. We then describe the methodology we implemented to create the personas. The results section includes an analysis of how the personas were used by referring to comments made by workshop participants and discussion transcripts. In the discussion section we address the strengths and weaknesses of our approach and personas in general as well as lesson learned from the experience.

2 Previous Prototype Development Work

The development of the APOSDLE project included a variety of research activities (e.g. creativity workshops, focus groups) executed in part to elicit requirements and evaluate the system. A functional prototype has been evaluated every year as part of an iterative development approach. Four of the twelve organisations involved in the project are companies seen as the target market of the APOSDLE system. Project members working in these four organisations (referred to in this paper as *application partners*) are familiar with and represent the future end users of the system (e.g. engineers, consultants). Their involvement proved important to ensure the project was aligned with the needs and work environment of its intended users. The need to consider additional design tools, which led to the use of personas described in this paper, arose following an extensive four-week formative evaluation of the 2nd prototype [8]. Primarily work-based, it highlighted various usability problems ranging from ambiguous terminology to unclear system logic. A considerable number of users could not fully grasp how APOSDLE might benefit them at work [1]. We concluded the user interface had to be redesigned so it supported users' goals and knowledge needs more effectively. In order to achieve this objective we had communicate and remind 21 project members of those goals so they could be referred to during design discussions. The target audience of APOSDLE had been mapped from the outset; however the descriptions used lacked contextual data consisted of such abstract terms such as *Knowledge Seeker* and *Knowledgeable Person*. Such segmentation provided little if any useful data about the people the project was targeting and their work environment, therefore we decided to use personas instead.

3 Previous Research into Personas

Personas have become a popular tool since they were first introduced by Alan Cooper [3] as means of incorporating user information into the design process. By referring to carefully constructed rich descriptions of fictitious individuals, a design team aims to engage with 'real people' in order to assess how the product may or may not suit their goals. Similar to other techniques such as *Cultural Probes* [5], personas are meant to engage teams with useful information about end users, so that it can be referred to during the design process. This is often done in conjunction with other user-centred methods such as user testing as personas are not expected to provide all the answers regarding product usability and usefulness. Pruitt and Adlin [10] note the idea of designing for a very small set of concretely defined user archetypes is still fairly new and considered radical by many since the conventional mindset is to ignore the preferences of the individual user in order to satisfy those of a wider audience. However, Cooper [3] argues that contrary to common logic, it is better to design with a single person in mind rather than try to accommodate most people, assuming the single person being referred to is constructed correctly and is seen as a valid representation. Grudin and Pruit [6] list the following benefits of using personas: 1) They create a strong focus on users and work contexts 2) They utilise people's powerful ability to extrapolate from partial knowledge to create coherent wholes and project them into new settings 3) The act of creating personas makes assumptions about the target audience explicit 4) They provide a shared basis for communication 5) They focus attention on a specific target audience.

Using personas during the design process effectively creates numerous challenges, as they need to be memorable and engaging in order to have an effect [9, 11]. People need to buy-into the personas and get to know them if they are expected to champion their goals. The personas also need to be used on a regular basis in order to influence the design process. Grudin and Pruit [11] reported on the successful use of personas at Microsoft; however they ended up developing different tools in order to communicate the personas and ensure they are believable representations. Similarly, Nieters et al. [9] also reported how using the tool required some degree of improvisation to make the personas memorable and engaging (the personas were eventually represented by plastic action figures). Such accounts demonstrate that in practice, personas are often used in different and at times creative ways, and are seen as flexible rather than a rigid tool. Chapman and Milham [2] present a sceptical view of personas and describe methodological weaknesses and practical limitations, which they argue undermine the validity of the tool. From a methodological perspective, they assert personas are difficult or impossible to verify and represent only a very small portion of the intended user population. In terms of practical limitation they point to the fact that it is not always clear how the personas are reconciled with other data and who is responsible for interpreting them. Existing literature about personas indicates the tool is debatable with various strengths and weaknesses. As we describe in later sections, we acknowledged methodological and practical weaknesses and made attempts to address them by collaborating with stakeholders to create the personas and involving them in the design process to support the interpretation of those personas.

4 Methodology – Creating the Personas

In this section we describe how the APOSDLE personas were created before being used during a two-day design workshop. We developed the personas based on empirical data gathered from observations and interviews as well as previous research carried out as part of the APOSDLE project (e.g. [6]). Given the four *application partners'* wealth of internal organisational knowledge, we decided to augment the standard persona creation process by getting them actively involved in creating and validating the personas. They were asked to produce a detailed description of employee types using a template we provided and based on their input, as well as additional data, we synthesised a set of four personas. By doing so we sought to address methodological and practical weaknesses expressed in the literature with regards to persona validity and the way the personas are interpreted and utilised. Apart from validating the personas the *application partners* were frequently consulted with during the design workshop when a persona was referred to. As we later describe, while this helped resolve ambiguity and disagreements it also contributed in our view to the personas being underused overall.

4.1 Data Collection

We provided the *application partners* with a template that included basic demographic details and sections frequently mentioned in literature about personas such as 'typical day at work', 'goals', 'dislikes' and 'attitude towards technology'. We received eight descriptions in total from three of the four *application partners*. A consultant and administrative assistant emerged as the main employee types in two consultancies. An engineering company provided four descriptions, which included a senior engineer and an intern. Based on these descriptions and empirical data we already had, we identified the following dimensions as the main distinctions between the different employee groups:

- Nature of work – Scientific and rigid (Engineering) versus creative and flexible (Consultancy). While the former relied heavily on a structured approach, calculations and precision, the latter required being able to adapt to clients.
- Level of experience - An experienced, possibly senior, employee versus a new employee, probably an intern.

While these dimensions were not new to us, and such user segmentation was identified early in the project, our goal was ensuring this data was made visible and engaging to all 21 participants during the workshop so it could help establish empathy with end users needs and circumstances.

4.2 Constructing the Personas

We ended up constructing four personas based to a large extent on the ones provided by the *application partners*. The personas represented senior and junior employees in scientific and creative industries. The two primary personas represented senior employees in an engineering company (Pierre) and a senior consultant working in an innovation consultancy (Eva). One secondary persona represented an intern in an engineering company (Paul) and the other an administrative assistant in a consultancy (Lisa). Only one *application partner* listed an intern as a possible persona, therefore we considered this type of employees to be secondary compared to senior and experienced employees. We did not expect APOSDLE would equally satisfy all the personas given their different goals; hence the distinction between the primary and secondary personas i.e. the goals of a primary persona had priority over those of a secondary one. As can be seen from the following example, we produced a rich persona description that captured daily work routine as well as behavioural patterns that could inform the persona's goals, preferences and attitudes:

> Eva – 37 years old, married + 1 and lives in Graz, Austria. She has PhD in Industrial Economics. She worked for two years as a junior consultant, mentored by a senior consultant before taking over projects. Her work requires her to be creative and constantly adapt to new situations created by working with new clients. She is often away from the office due to meetings or travels. She prefers as much face-to-face communication with her clients. She is self-organised, self-directed and sometimes find herself overcommitted. She uses phone and email for contact colleagues, many meetings with other team leaders and heads of department. Eva uses mainly MS Office, two different web content management systems, two address databases, accounting system, Lotus Notes and internal Wiki. She is technologically interested, curious and open-minded. She is very busy and critical and therefore prefers to have a quick look in Google instead of searching her notebook or the server. Typical day: many meetings and emails, multitasking; deciding, planning, coordinating controlling and monitoring activities. In her free time she likes sports, she is an active triathlete. She is very interested in school topics and the education of her children and is chairman of the parents association at her childrens' school.

Each description included 10-15 related visuals that were meant to help participants extrapolate additional information by visualising various aspects of the written narrative. Using visuals to augment a persona's textual description is common practice [e.g. 9, 11], and similar to Gaver et al.'s [5] *Cultural Probes* techniques, is meant to generate further insight by engaging people with visual or tangible objects. In Eva's case the visuals included a mindmap, flipchart, a handshake, a train and *Google*'s logo. They emphasised key attributes such as the creative and client-facing nature of her work, as well as the fact she often travels and is therefore out of the office. Before using the personas we asked the *application partners* to review and validate the personas we created by examining how well they represented employees in their organisation. We received positive feedback (see example below) and were satisfied the personas were a realistic representation and ready to be used.

> *I like very much the personas you have synthesized from the application partners' input. I really can recognize [our company's] two main groups of employees in Eva and Lisa.* (Personal email correspondence)

4.3 Introducing the Personas

The personas were introduced at the beginning of the design workshop during an hour long 'familiarisation session'. 21 participants representing each project partner (programmers, learning experts, knowledge management scientists, *application partners*) took part in the workshop, which was facilitated by two of the authors. Since the participants were not familiar with personas the challenge we faced was two fold. First, they had to buy into the tool and appreciate how it could benefit the design process. Second, they had to accept the personas we developed as realistic representations of the end users. The familiarisation session was designed to introduce the personas and allow the participants to eventually perceive them as real individuals and advocate their goals. During the session each persona was projected on a screen one at a time and a facilitator introduced it and asked key questions in order to generate a discussion. The process had four main objectives:

1. Get participants to view the personas as real people they can relate to and empathise with;
2. Explore general assumptions about the personas and talk about their work and learning needs (explicit and implicit);
3. Explore specific assumptions regarding each persona's attitude towards the current version of APOSDLE e.g. what aspects of the system the persona would likely appreciate and find useful;
4. Explore similarities and differences between the four personas;

Each discussion began with the following key questions: "What kind of information is the persona likely to look for at work?" and "How does the persona find answers to questions she has at work?" These questions set the scene for a discussion about the personas' likely attitude towards the existing version of APOSDLE. The main questions we asked with regards to the system were: 'What aspects of the current APOSDLE prototype the persona would not like?' and 'What aspects of the current APOSDLE prototype persona would like?' Having the application partners present was useful as they contributed examples and helped resolve disagreements by either supporting or rejecting assumptions made by participants about the personas.

5 Results – Utilising the Personas

In this section we report how the personas were utilised during the workshop and the impact they have had on the redesign process. We first analyse the data gathered during the familiarisation session and then present a sample of transcripts from the discussions showing how the personas were used to inform various design decisions. In order to evaluate the impact the personas may or may not have had on the workshop we reviewed video footage of the two days and observed how the personas were used. We isolated the events where the personas were referred to in order to assess: 1) How much impact they had on the discussions 2) What was their impact in terms of redesign and 3) how the way they were utilised might have been affected by the participation of the *application partners* who helped create them.

5.1 Familiarisation Session

The familiarisation session resulted in general comments made by participants about the personas and assumptions as to how they were likely to perceive the existing version of APOSDLE. This part of the workshop lasted 63 minutes and generated a

Table 1. Genereal comments and assumptions made about the personas

<table>
<tr><th></th><th>Rigid work process</th><th>Flexible work process</th></tr>
<tr><td>Senior employee</td><td>Pierre (Primary persona)
Regarded as expert
Responsible for training and supporting new employees (interns)
Set in his ways
Contacts internal and external experts
Knows who knows what in the company
Deals with processes, simulations and calculations
Would like to spread his knowledge and build a positive reputation</td><td>Eva (Primary persona)
Knowledge needs depend on clients
Frequently Out of the office
Relies on past projects and peers' experience
Deals with people (clients)
Process influenced by clients</td></tr>
<tr><td>Junior employee</td><td>Paul (Secondary persona)
Question driven
Internet savvy
Willing to explore new things
Explicit learning needs
Single task at a time
Requires internal knowledge (procedures)
Uses Google products
Does not know who to ask for help
Very social
Requires detailed and specific knowledge
He has time to learn and likes to explore new things</td><td>Lisa (Secondary persona)
Supports others work
Free to plan her work
Work is influenced by requests made by consultants
Acts as a 'hub' in the office</td></tr>
</table>

lively discussion, which actively involved more than half of the participants. The personas were introduced and discussed in the following order: Pierre (22 minutes), Eva (16 minutes), Paul (13 minutes) and Lisa (11 minutes). Table 1 summarises the general points raised during the discussion about each persona; from personal preferences to the type of knowledge needs encountered at work.

The participants' comments enriched the personas' description and added important information. The *application partners* were referred to on a number of occasions to support or dismiss assumptions participants made, and their input was very useful. As can be seen, Pierre and Paul generate the highest number of comments. We can speculate on what were the reasons why the personas were not equally engaging: 1) Participants found it easier to relate to Paul and Pierre because they were more familiar with their work domain; 2) Pierre and Paul were seen as more relevant to the project; 3) The description and accompanying visuals associated with Pierre and Paul were clearer and possibly more engaging. A particular reason why Lisa was not as engaging as the other personas could be due to the fact she was the last persona to be introduced after more than 50 minutes. It is reasonable to assume that at that point general fatigue set in and affected the degree of participation. In hindsight, it would have been better to discuss the personas in parallel and ensure they are each discussed for an equal amount of time. It is worth noting Lisa was included as a persona because two *application partners* described an assistant as a significant employee type. This did not automatically make her an end user of APOSDLE. However, we wished to explore ways APOSDLE could potentially benefit new user types. We believe the Lisa persona was also less engaging since it was not perceived as relevant to the project. We tend to exclude the possibility of gender-bias in this case since there was an equal number of men and women taking part in the discussions and overall in the workshop.

Table 2 summarises the comments expressed regarding each persona's likely attitude towards the 2nd prototype. These issues were discussed following the general comments section while the participants were still engaged with a particular persona. As can be seen by the comments, participants felt comfortable inferring how the personas would likely perceive APOSDLE. This was encouraging and showed the first part of familiarisation session was effective as it enabled participants to extrapolate new information based on the persona description. Ensuring the discussion revolved around two questions i.e. what the personas would like and not like about APOSDLE, helped focus the responses and was very effective. It also helped ensure the comments would be comparable and could be easily grouped to uncover patterns,

Table 2. Comments and assumptions made regarding the personas' attitude towards APOSDLE

	Rigid work process	Flexible work process
Senior employee	**Pierre (Primary)** • Will not use task selection • Will not use the Wiki to communicate • Might not use APOSDLE's generated Learning Events and jump straight to a PDF without reading any introductory text offered by the system • Would appreciate collaboration options and a place to discuss with others • Would appreciate detailed and advanced search options • Would appreciate a good domain model • Requires incentives to use the annotation tool and 'tag' portions of resources according to the domain	**Eva (Primary)** • Would consider APOSDLE as 'just another tool' • Would not use APOSDLE's process view as her work is less about 'ticking boxes' • Would be interested to get a quick overview of her clients' work domain (especially if she is not familiar with that work domain) • Would likely compare APOSDLE to Google and conclude it is not as good • Would appreciate different work domains to learn from about her clients • Would appreciate collaboration options within APOSDLE • Would appreciate being able to locate experts using APOSDLE • Would consider APOSDLE a creative concept
Junior employee	**Paul (Secondary)** • Would not like the fact APOSDLE does not interface with his favourite applications • Would perceive APOSDLE as too difficult to learn compared to web applications he uses • Would not like the fact APOSDLE does not run on Linux • Would not understand the modelling terminology used in APOSDLE • Would appreciate the fact that APOSDLE highlights relevant portions of resources instead of having to read through a 500 page long PDF • Would appreciate the fact that by effectively using APOSDLE he does not need to ask people questions all the time • He is not too concerned with privacy and whether APOSDLE records his personal data	**Lisa (Secondary)** • Would not appreciate the lack of interoperability with other collaboration tools she uses at work • Constantly changing her tasks could hinder the automatic context detection process • May not see APOSDLE as a tool that directly supports her work

similarities and differences between the personas. The following assumptions focused participants' attention on key issues and served as the basis for the design session we describe in the next section.

The patterns which emerged during this discussion referred to mainly to APOSDLE's emphasis on structured learning support and how well it suited different personas. While experienced employees such as Eva and Pierre would likely appreciate various unique features of APOSDLE like the use of semantic models to link resources to tasks and topics, its focus on learning would probably deter them and appeal more to an inexperienced employee such as Paul. This had significant impact on functionality and terminology. Such assumptions were derived from the data presented in Table 1 as it was understood Paul had explicit learning needs and was question driven compared to Pierre who was more problem driven. Out of the four personas, Paul was seen as the only one that was given explicit learning tasks by his supervisors and was most likely to be in a 'training mode'. Once it was established that experienced employees like Eva and Pierre were less likely to require structured learning support, we began exploring how they might benefit from APOSDLE's various features in their everyday work. One significant outcome of this realisation was that the *Search* function became more prominent in the user interface; acknowledging it would be frequently sought by someone familiar with the work processes and who is interested in quick and easy access to information rather than following a step-by-step approach.

5.2 Reviewing the 2nd Prototype

Following the familiarisation session, we presented participants the 2nd prototype and began focusing on specific system components from the point of view of each persona. We focus on two of several examples, which illustrate how the personas were used to evaluate the prototype and inform design decisions. In the first case, participants were asked to evaluate APOSDLE's *Learning Events* concept. A *Learning Event* was an automatically generated collection of annotated documents, which included an introductory paragraph and a form allowing users to reflect and assess how well they have learned the material. A list of available *Learning Events* was displayed once the user selected a *Learning Goal* (e.g. "How to plan a creativity workshop") from a list of learning goals relating to a specific work task (e.g. "Run a creativity workshop"). The formative evaluation of the 2nd prototype showed users struggled to make sense of *Learning Events* so assessing and redesigning this system component using the personas was important.

> *Facilitator A: Paul is probably the main recipient of Learning Events [according to the previously mentioned assumptions]. What Paul might want?*
> *Participant 1(Knowledge management scientist): He had a Learning Goal. Otherwise he wouldn't end up with a Learning Event.*
> *Facilitator A: Is Paul able to select his own Learning Goal? How did users react to Learning Goals in the formative evaluation?*
> *Participant 2 (Application Partner A): They didn't understand what they meant.*
> *Participant 3 (Application Partner B): Some of our users didn't understand the Learning Goal type (apply, understand etc.)*
> *Participant 4 (Learning expert): Perhaps we should phrase them in a more intuitive manner.*

> ***Participant 5 (Programmer): The Learning Events were created to facilitate learning. I think we should reflect on the four personas and see if they have need for this information. We are discussing low-level details and terminology before addressing the users' real goals.***
> *Facilitator B: There were two personas, Lisa and Eva, that we said need quick overview of material.*
> *Facilitator A: What would Paul want to learn from?*
> *Participant 4 (Learning Expert): he wants a definition*
> *Participant 6 (Learning Expert): A slider could be used to show different level of learning support depending on the user's learning needs.*

It is interesting to note the comment made by Participant 5. This was one of the programmers coding APOSDLE not involved with the learning theory aspects of the system. He noticed the discussion was focusing on low-level design details while overlooking the personas' high-level goals i.e. focusing on how *Learning Events* were presented rather than asking whether they were useful in the first place given the personas' goals. This was an encouraging sign that some participants appreciated the personas' contribution and referred to them without the facilitators' intervention. As a result of a subsequent discussion it was decided to modify *Learning Events* in APOSDLE so they are presented as optional learning support an inexperienced employee like Paul might utilise, while not deterring more experienced employees from using the system. The next transcript illustrates how the personas were used to evaluate APOSDLE's collaboration component. It is used to formulate a collaboration request and share with the contacted person relevant resources before actual communication takes place using one of several means (e.g. phone, *Skype*, email).

> *Facilitator A: What kind of collaboration would Paul want to do?*
> *Participant 7 (Programmer): If Paul has a visual basic problem, he won't hesitate to use the collaboration option. He will go there immediately and it will be very helpful if an expert with the relevant knowledge is suggested.*
> *Facilitator B: Is he more likely to collaborate with an expert or another intern?*
> *Participant 4 (Learning expert): He will turn to an intern first and then an expert, then a book and then an expert again*
> *Participant 3 (Application Partner B): Interns have to prepare their question. He will ask his question according to some resources. He will look for something in the resource repository and if he doesn't understand something then he might consider a collaboration using APOSDLE*
> *Participant 8 (Application Partner C): Paul will first talk to his advisor.*
> ***Participant 9 (Knowledge management scientist): Will Paul like this kind of thing? I wonder if people will go through this process. I can imagine Paul simply pressing 'ask' and bypassing the form***
> *Participant 10 (Programmer): What about a drag and drop? What if Paul finds a document but doesn't understand something he can drag it into the collaboration wizard and share it with the expert he is about to contact.*
> *Facilitator B: Would Paul use the collaboration wizard to contact another fellow intern given there is no power distance between them?*
> *Participant 3 (Application Partner B): [laugh] they would probably use Google Talk for that*

Again, the discussion involved a number of participants from different professional backgrounds and the *application partners* contributed relevant contextual data. Given Paul's status as an intern it became clear that he might appreciate a structured collaboration request when contacting an expert in the company. However, when contacting someone at his own organisational level, he would either use an alternative more informal communication tool such as *Google Talk* or would opt to submit an unstructured request via APOSDLE, which will not require him to state his request in a rigid manner. As the sample transcripts show, getting participants to think about the system from the perspective of the personas helped articulate issues that we believe would have otherwise probably been overlooked. Referring to the personas helped focus discussions on end users and reduced the impact of personal preferences on the decision making process.

6 Discussion

In this section we discuss the strengths, weaknesses and limitations of using personas as part of our redesign efforts and specifically address the approach we adopted. As demonstrated by the discussion transcripts, when designing the 3rd prototype of APOSDLE we focused solely on four fictitious individuals. We found that thinking about the system from the point of view of those people was by far more productive than referring to abstract terms such as *User* or *Knowledge Seeker*, which conveyed little if any meaningful data. We found personas a useful way of making information about end users, which tends to be scattered across several documents in different forms, visible so it could be easily referred to. It made such information tangible, meaningful and most importantly – referable by name. The personas were by no means seen as a perfect solution, however personifying key considerations such as the type of work users were involved in brought to the forefront essential consideration that probably would have been otherwise ignored or forgotten. Instructing participants to justify their suggestions by referring to a specific individual by name helped ensure the process was more user-centred and made people think explicitly about the individuals APOSDLE was meant to support.

In order for the personas to *work*, it was essential to ensure we were using the *right* personas i.e. realistic and valid representations of key user groups. Most of the accounts regarding persons we came across describe how personas were created based on methods such as interviews and observations. This raises questions concerning the validity of the end result and the way in which it is interpreted. Our collaborative approach sought to balance our own subjective interpretation of the data by involving stakeholders representing the intended end users in the creation and utilisation of the personas. Getting these stakeholders involved proved useful as they contributed data in the form of initial persona descriptions and validated the end result. Would we have been able to produce the personas without the *application partners*' contribution based solely on empirical data? We believe so, however we feel incorporating direct contributions from these stakeholders enriched the descriptions and made them more believable and ultimately more valid. This led to an effective buy-in process, which was essential in order for participants to accept and use the personas. We could have asked the *application partners* to merely review and comment on our personas but wanted them to feel a greater sense of ownership over the end result by getting involved.

The process of combining our own data with the *application partners*' contribution was not without its problems there are several limitations to the approach we have described in this paper. First, we relied on the *application partners* to provide vital data. At times this caused delays and in one case we did not receive the information we requested. The collaborative nature of the process meant our work developing the personas depended on the *application partners*, which was not ideal though ultimately beneficial in our view. Second, despite having provided a template, the style of the responses we received varied considerably. Consolidating the different styles in a consistent format across all four personas was challenging, but given the outcomes of the workshop we believe we achieved that goal successfully. In terms of utilising the personas in the workshop it was very helpful to have the *application partners* present. Because they were familiar and personally knew the people the personas represented, they helped interpret them. This eliminated ambiguity, uncertainty and disagreements among participants about the personas. However, as the design session progressed we observed that since the *application partners'* were seen as representatives of the personas, participants bypassed the personas by engaging with them directly. Not surprisingly, it was easier for them to ask a real person a question rather than attempt to view the system from someone else's perspective and make assumptions about a fictitious character. We view this as a weakness of our collaborative approach, which should be noted and addressed. Nevertheless, we do not think the *application partners* could have replaced the personas since each served a different purpose. The personas were first and foremost a communication tool, which personified user information and made it visible to project members thus enabling them to consider problems and solutions by thinking about the system from the point of view of different users. The stakeholders complemented and enriched the personas – they did not replace them. As facilitators, we felt at times that constantly reminding participants to refer to the personas to justify every design decision would have been potentially too annoying and counter-productive. This in our opinion is a significant weakness of the tool. Although we expected the personas to be 'present' at all time, we realise now this was not a very realistic expectation.

Despite the fact that during the workshop the personas were not referred to as rigorously as we had initially expected, they created a strong focus on APOSDLE's users and their goals, and as such we found them to be effective. Active participants appeared to be comfortable referring to Paul, Pierre, Eva and Lisa as they justified their ideas and assessed others' suggestions. For the most part, the persona-driven discussions resulted in actionable design decisions. We acknowledge that participants' familiarity with personas is essential to ensure they are naturally woven into the discussions. Overall, our experience using personas was a positive one and contributed significantly to our redesign efforts.

7 Conclusions

In this paper we described a case study of using personas to redesign the 3rd prototype of the APOSDLE system. We used personas in order to explore how users' goals, behaviour and context could be communicated effectively to 21 project members during a two-day design workshop. The notion of designing with only a handful of specific individuals in mind has been debatable and some might argue - controversial. There are obvious pros and cons to the persona approach; however in our case we conclude the strengths outweighed the weaknesses. Keeping the end users in mind when considering design solutions is highly challenging and personal preferences, rather than the users' goals, can quickly end up dictating design decisions. We found personas to be an effective way to encapsulate and communicate user information so it served as constant reminder during discussions. Creating and utilising the personas was challenging yet beneficial. By combining initial persona descriptions generated by stakeholders with our empirical data we feel we were able to address to some degree methodological weaknesses and practical limitation of the tool expressed in the literature. The contribution of the user representatives helped in our view to increase the personas' validity and credibility, which were essential to ensure they were utilised successfully. Despite the fact the four personas were not referred to as rigorously as we had intended, we feel their contribution was nonetheless significant. They confronted project members with key questions regarding the ways APOSDLE could support the goals and needs of users in different industries and career stages. As reported by advocates of personas, we also maintain that personifying user information proved a useful way to communicate essential data about the end users, which despite possible inherent weaknesses was still better than any alternative we were familiar with.

Acknowledgments. APOSDLE (www.aposdle.org) is a 48 months integrated project partially funded by the European Community under the Information Society Technologies (IST) priority of the 6th Framework Programme for R&D.

References

1. APOSDLE project internal deliverable - Formative Evaluation Report of 2nd Prototypes (2008)
2. Chapman, C.N., Milham, R.P.: The Persona's New Clothes: Methodological and Practical Arguments Against a Popular Method. In: Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting (2006)
3. Cooper, A.: The inmates are running the asylum. In: Why High-Tech Products drive us Crazy and How to Restore the Sanity, Macmillan, Indiana (1999)
4. Davis, F.D.: A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. Doctoral dissertation, MIT Sloan School of Management, Cambridge MA (1986)
5. Gaver, W., Dunne, T., Pacenti, E.: Cultural probes. Interactions 6(1), 21–29 (1999)
6. Grudin, J., Pruitt, J.: Personas, Participatory Design and Product Development: An Infrastructure for Engagement. In: 7th Participatory Design Conference (PDC), pp. 144–161 (2002)
7. Kooken, J., Ley, T., de Hoog, R.: How Do People Learn at the Workplace? Investigating Four Workplace Learning Assumptions. In: Duval, E., Klamma, R., Wolpers, M. (eds.) EC-TEL 2007. LNCS, vol. 4753, pp. 158–171. Springer, Heidelberg (2007)
8. Lichtner, V., Kounkou, A.P., Dotan, A., Kooken, J.P., Maiden, N.M.A.: An Online Forum as User Diary for Remote Workplace Evaluation of a Work-Integrated Learning System. In: 27th Computer-Human Interaction Conference (in press, 2009)
9. Nieters, J.E., Ivaturi, S., Ahmed, I.: Making Personas Memorable. In: Conference on Human Factors in Computing Systems (CHI), pp. 1817–1824 (2007)
10. Pruitt, J., Adlin, T.: The Persona lifecycle. In: Keeping People in Mind Throughout Product Design. Morgan Kaufmann Publisher, San Francisco (2006)
11. Pruitt, J., Grudin, J.: Personas: practice and theory. In: Conference on Designing for User Experiences DUX 2003, ACM press, San Francisco (2003)
12. Sharp, H., Rogers, Y., Preece, J.: Interaction Design: Beyond Human-Computer Interaction, 2nd edn. Wiley, New-Jersey (2007)

Play-Personas: Behaviours and Belief Systems in User-Centred Game Design

Alessandro Canossa[1] and Anders Drachen[2]

[1] Denmark Design School, Strandboulevarden 47 2100 Copenhagen Ø
IO Interactive, Kalvebod Brygge 4, 1354 Copenhagen K, Denmark
aca@dkds.dk, alessandroc@ioi.dk
[2] Center for Computer Games Research, IT University of Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen S, Denmark
anty@itu.dk

Abstract. Game designers attempt to ignite affective, emotional responses from players via engineering game designs to incite definite user experiences. Theories of emotion state that definite emotional responses are individual, and caused by the individual interaction sequence or history. Engendering desired emotions in the audience of traditional audiovisual media is a considerable challenge; however it is potentially even more difficult to achieve the same goal for the audience of interactive entertainment, because a substantial degree of control rests in the hand of the end user rather than the designer. This paper presents a possible solution to the challenge of integrating the user in the design of interactive entertainment such as computer games by employing the "persona" framework introduced by Alan Cooper. This approach is already in use in interaction design. The method can be improved by complementing the traditional narrative description of personas with quantitative, data-oriented models of predicted patterns of user behaviour for a specific computer game Additionally, persona constructs can be applied both as design-oriented metaphors during the development of games, and as analytical lenses to existing games, e.g. for evaluation of patterns of player behaviour.

Keywords: Play persona, emotion, game design, user centered design, user experience design.

1 Introduction: Users and Computer Games

End users of electronic interactive entertainment such as computer games are accustomed to expect a high degree of customization from the products during actual use. They expect sensitivity, awareness and responsiveness, with the standard being set by Web 2.0 [23] compliant applications and services such as Wikipedia, Google Maps and Facebook. The essence of the innovations posed by these applications is how they place the user at the center of the experience, and how they make the user feel: users think they matter, that they can make a difference and it gives them a way to express themselves. For example Nicole Lazzaro recognized the importance of involving players, at least conceptually, when designing games by stating: *"A game's value proposition is how it makes its players think and feel. Players don't buy games, they*

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 510–523, 2009.

buy experiences" [11]. A failure to incorporate means for mass customization [16] could risk alienating the large majority of people that are becoming more and more acquainted with the practice of expressing themselves at almost every occasion. Furthermore, non-trivial choices and ability to express oneself are seen as determining factors for critical acclaim and are required for adding player choice and broad appeal to computer games [21].

Implementing these innovations in computer games requires both substantial technological advancements and a philosophy of game design that directly incorporates the user needs from the early stages of the design and development process [13, 15, 10, 5, 22]. Technological advancements can include platforms for user generated content, adaptive AI, eventually even biometric feedback [17, 18, 14, 12]. We are only starting to scratch the surface of the many possibilities unleashed by user-aware technology. For example, Left4Dead's [1a] AI Director is a dynamic system that monitors the players' performance and orchestrates accordingly the distribution of enemies, items, visual effects, dynamic music and dialogues between characters. Little Big Planet's [2a] magic relies almost completely on the global exchange of its users' creativity, similarly a big portion of the appeal of games such as Spore [3a] or SingStar [4a] seems to be the possibility of sharing user generated content. All these technology-based innovations require considerable investment of resources. On the other hand, a paradigm shift in game design philosophy appears to be emerging, which can bring about a strong focus on the user and create a set of games that attempt to empower players, all of this requiring minimal technical efforts. Such a game design philosophy mirrors what has already happened in the field of Human Computer Interaction: the introduction of User-centered and Experience design [15, 10]. In order to accelerate this paradigm shift in the field of game design this paper advocates first the adaptation and then the adoption of tools and practices already in use in HCI for modeling users, such as the persona framework [2, 3].

2 Designing User Experience in Games

According to Frijda [7], emotions are complex dynamic processes and arise as states of action-readiness, including affect and arousal, when we interact with people and things in the world. Past experiences are evaluated and remembered and influence the appraisal of future occurrences to inform ensuing behaviour.

Damasio defines emotional processes as sets of rational, bodily, and behavioural responses to the perception (or memory) of an experience [4].

It is possible to see how, as Sanders correctly points out [19], all efforts aimed at engendering defined experiences and emotions are doomed to failure since experiences and emotional responses alike are too individual, subjective and rooted in people's past to be able to scientifically aim at re-producing them.

It is necessary then to utilize a design philosophy that takes into account players and allows them a certain leeway for expressing themselves, but without assuming universal emotional responses to experiences. Such a frame of mind should be accountable for different player's motivations, goals, behaviours and belief-systems.

In the domain of user centred design, such a mental tool exists already. Alan Cooper developed a method called Goal-Directed design [2]. This method makes use of personas - “archetypes that represent distinct groupings of behaviours, attitudes, aptitudes, goals, and motivations” [3] - to help developers understand the end user and to foresee its way of interacting with the product.

2.1 Origin and History of Personas

Persona is a Latin word and it indicates the mask that actors put on before becoming their characters, it’s a socially agreed convention used to represent certain *types*. Currently it refers mostly to “social masks” or roles that all humans have to play on the stage of life [8]. Goffman uses the term “fronts” to address the different masks that we have to wear according to the different contexts we are presented with. We must act differently in different settings, as the world is a stage. It is in this sense that Jung listed it as one of the archetypes populating the human unconscious.

Personas or fictional identity-constructs have been recognized as fundamental in many creative practices. In literary theory, Iser [9] introduced the term “implied reader” to address the certain “reader that a given literary work requires”. Within the frame and the context imposed by the text, this implied reader makes assumptions, has expectations, defines meanings that are left unstated and adds details to characters and settings through a "wandering viewpoint". For example, by Joyce’s own admission, Finnegan’s Wake should be read by “that ideal reader suffering from an ideal insomnia”. Eco expanded on the concept introducing the “model reader” [6] as “the author’s foreshadowing of a reader competent enough to provide the best interpretation of a text”. The author tries to prefigure a model reader by imagining what could be the actualization of the text. The author, consciously or not, is concerned with how the text / type becomes interpretation / token.

In social sciences, Max Weber introduced the concept of Idealtyp as: *“Formed by one-sided accentuation of one or more points of view and by the synthesis of a great many diffuse, discrete, more or less present and occasionally absent concrete individual phenomena, which are arranged according to those one-sidedly emphasized viewpoints into a unified analytical construct”* [20]. The ideal type is a pure mental construct used to assess the behaviour of social groups. It is totally theoretic, almost fictitious and generally not empirically found anywhere in reality, it is not backed by statistical data nor a model personality profile, it’s more used as some sort of unit of measure, standards much like "meter", "second" or "kilogram" not really found in nature, but useful to measure it.

In similar ways game designers could benefit greatly by making assumptions on the nature of players using personas to map the extreme boundaries of the field of possibilities afforded by their game.

2.2 Traditional Persona Modeling

Alan Cooper’s [3] goal directed design process starts with the research phase, in which behaviours patterns and modes of use of products are identified. These patterns suggest goals and motivations and in turn these inform the creation of personas. Personas are detailed, composite user archetypes and they serve as main characters in narrative, scenario-based descriptions that iteratively inform the design of a product, so that features emerge directly from the goals.

Typically a persona is a description of behaviour patterns, goals, skills, attitudes, and environment, with a few fictional personal details to make it a realistic character. For each product there should be set of 3 to 12 personas, it's not necessary to design for all of them, but an extensive cast helps articulate the user population, the primary focus for the design will be a limited subset of maximum three personas. Persona description should be precise including as many details as possible, but not necessarily accurate, it does not need to represent a real person. Name, physical appearance, education, and idiosyncrasies should be included.

The main benefits of personas for product development purposes are:
A) It is easier to relate to a personal human face and name instead of abstract customer data.
B) It is possible to infer user needs not openly stated by drawing on personal people-experience.
C) Personas provide a shared, fast and effective language for communication between engineers and designers.
D) Personas states what a user needs and wants so that no stakeholder can reshape the user to their convenience.
E) Personas avoids self-referential designs, where designers might unconsciously project own mental models.
F) Personas also work as reality-checks, helping designers keeping the focus on the limited subset of users that have been deemed "primary".
G) Proposed designs, features and solutions can be evaluated against the needs of individual persona models.

Personas have also been criticized mostly because if they are fictional, they have no clear relationship to real customer data and therefore any data gathered cannot be considered scientific [1]. In any case, in order to apply this design method to computer games some changes are necessary.

3 From Persona to Play-Persona

Play-personas are further defined as clusters of preferential interaction (*what*) and navigation (*where*) attitudes, temporally expressed (*when*), that coalesce around different kinds of inscribed affordances in the artefacts provided by game designers. This means that personas can no longer just be limited to narrative descriptions of motivations, needs and desires distilled in ethnographic interviews.

The persona hypotheses emerge as a relation of parameters from the set of interaction and navigation possibilities that the game-rules and game spaces can afford [22]. Personas can be augmented and strengthened by a quantifiable, parametric, data-driven perspective. Furthermore, if directly coupled with instrumentation data in the form of gameplay metrics gathered from game engine software during play sessions, play-personas can provide a powerful evaluation tool to confirm whether a certain hypothesis also turns out to represent a sizeable slice of players. That is why play-personas are both theoretical models of ideal users (*metaphor*) and data-driven representations of player behaviours (*lens*).

3.1 Gameplay Metrics

Gameplay metrics are instrumentation data extracted from computer game engines during play. The collection of quantitative data about user-product interaction is an established method within Human-Computer Interaction [24, 25], but has only recently been adapted to user testing in computer game development [10, 22]

The type of game instrumentation data relevant to persona modeling is gameplay metrics, which are measures of game-player interaction. Gameplay metrics can be recorded for any type of user-initiated behavior where interaction takes place in or with the virtual environment; as well as behaviors initiated by agents and systems operating in the virtual environment but which are not controlled by the player, e.g. autonomous agents. Tracked and logged actions can vary from low-level data such as button presses to in-game behavior. For example, tracking the location of a player through the virtual environment, the use of specific character skills, weapons, hit probability etc. Gameplay metrics analysis provide the ability to generate highly detailed, objective and quantitative analyses about player (end user) behavior.

Metrics analysis can be used to evaluate user behavior, and the data type forms a supplement to existing methods for user-oriented research and –testing in the game industry, such as usability testing, playtesting, focus groups and similar qualitative/semi-quantitative approaches. Different methods have different strengths and weaknesses. For example, usability testing focuses on measuring the ease of operation of a game. Playability testing explores if users have a good playing experience. Gameplay metrics analysis offers however detailed insights into how the users are actually playing the games being tested. Furthermore, gameplay metrics can be used to develop data-driven persona models.

3.2 Play-Persona as Metaphor

A metaphor is a rhetorical device that allows describing something unknown by transferring attributes from a known entity. Metaphors are utilized before the accumulation of experience, in a similar way personas allow designers to “imply” unknown player behaviour in the process of creating digital games, i.e. by pre-defining the ideal play-patterns possible in the game in question and design to accommodate these. It is the case in *Tomb Raider Underworld* [18], where players can choose between different identically optimal strategies to progress and express preferences for some modes of interaction and navigation instead of others. Play-personas as design tools represent an expectance of how players would like to craft their experience. As metaphors, play-personas are hypotheses that emerge as relations of parameters from the set of possibilities that the game can afford.

Designers can use personas as categories of behaviours prior to a playable version of the game in order to plan coherent navigation and interaction modes. They are also precious as guides to select which variables are interesting enough to monitor as game metrics, which will lead to the creation of play-personas as lens.

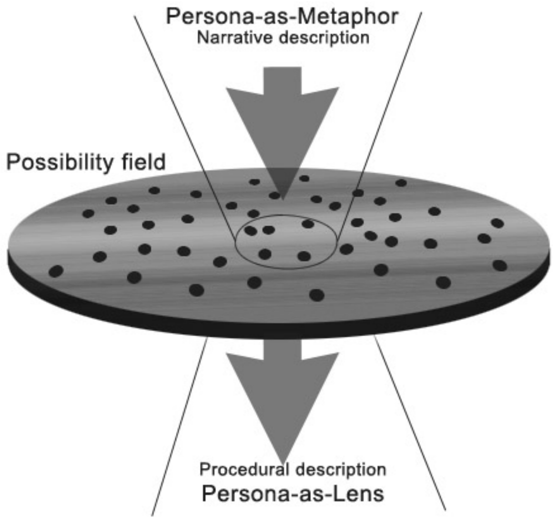


Fig. 1. Two sides of the same coin: persona as metaphor and lens. The black dots on the "possibility-field" plane represent game mechanics; thanks to the a-priori description of the persona-as-metaphor a certain subset is individuated. This persona hypothesis can be checked against metric data gathered from players and inform the creation of persona-as-lens.

3.3 Play-Persona as Lens

Lens is here intended as the choice of a context (*Ümwelt*) from which to sense, categorize, measure or codify experience. Lenses are utilized to examine the accumulated experience. Gameplay metrics data can form the basis of defining data-driven personas during game testing. As lenses, play-personas are derived from game metrics gathered from players, after they have been interpreted as clusters of similar behaviours. Personas can be used as tools when evaluating games by comparing the goals set by the designers with those of the players. By comparing designers' and players' goals it is possible to check whether the game design actually supports and facilitates the planned experience in practice, and if any new personas emerge from the user-interaction with the game software. Analyzing game metrical data with multivariate statistical tools can provide a way of discovering patterns in the usage of game elements and features, thus enabling the building of personas of how players interact with the game, and whether the game design facilitates the specific play patterns of the personas assumed as hypotheses. It's a sense-making perspective, a code that allows extracting meaning form an otherwise unclear list of numbers.

Playstyles (or patterns of play) are possible ways in which certain subsets of the rules and mechanics provided by the game can be combined. Player that maintains consistent choices of styles eventually identify with a play-persona.

Personas are aggregate description of possible player behaviour both in theory, as an expectation of the designer – "a priori" metaphor - and in practice, as a description of what actual, real players do during a play session – "a posteriori" lens (Fig. 1).

4 Tomb Raider Underworld: A Case Study

Taking as a concrete example the game *Tomb Raider Underworld* [5a] it becomes easier to see the two sides of the play-persona. The spaces, rules and mechanics of the game indicate few important skills necessary to identify personas:

- Shooting: Indicated by high or low number of deaths inflicted to enemies & animals (interaction with NPC)

- Jumping: Indicated by high or low number of deaths caused by falling, drowning, being crushed (navigation)

- Puzzle solving: Indicated by high or low requests for help to solve puzzles (interaction with the world)

Player can have "high" or "low" skills for each of the parameters individuated. Keeping these values down to two, although loosing granularity and resolution, simplifies immensely the possibility space.

	Expert	Rookie	Grunt	Athlete	Chess-player	Hybrid personas		
Shooting	+	-	+	-	-	+	-	+
Jumping	+	-	-	+	-	+	+	-
Puzzles	+	-	-	-	+	-	+	+

Fig. 2. Table showing how the three parameters unfolded in the two values (+ and -) map the possibility space of the game Tomb Raider Underworld

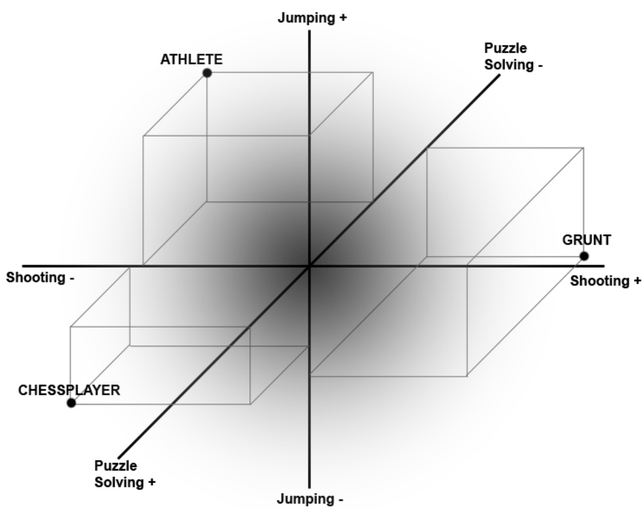


Fig. 3. Possibility space and the three play-personas: Grunt, Athlete and Chess-player. The axes can be defined abstractly such as "+" and "-" prior to a playable version of the game, when personas are used to imply user behavior patterns in the design. However, the axes can also be defined based directly on collected gameplay metrics. For example, the percentage of total puzzles solved. This provides a means for defining detailed quantitative components of the persona models, and even post-launch data-driven analysis of whether these personas emerge in the player behavior.

Mapping the possibility space consists of exploding all the combinations of parameters underlying the skills listed before: shooting, jumping and puzzle solving.

Each of the parameters individuates an axe of an n-dimensional space, where n equals the total number of variables worked with in the given situation. In this specific case there are only two possible values that each parameter can assume: plus or minus. Hence the total amount of combinations is the number of values that each parameter can assume, at the power of the number of all the parameters found: Two at the power of three for a total amount of eight combinations. This initial unfolding of gameplay parameters maps the a-priori possibility space within which personas will be found (Fig. 2).

At this point it is possible to select which personas-as-metaphors will guide the design process. Play-personas have been defined as extreme cases, one-sided accentuations that delimit the field, statistically most of the players will be identified by one of the *hybrid personas*, but these exaggerations are helpful both during design, to frame the kind of experience targeted by the game, and during post production, to evaluate how players actually relate to the game. Exaggerating one-sidedly the three parameters individuated lead to three personas: The Athlete, the Grunt and the Chess-player (Fig. 3).

4.1 Grunt Persona

<u>Narrative description:</u> this persona obviously excels at shooting. Grunts particularly enjoy fights and are quite good at them, they are interested in twitch stimulation, taking pride on physical domination, at the same time they might be not so precise with movements, hence the risk of failing at jumps. They might find slightly annoying tight jumping puzzles. Grunts are not interested in spending time figuring out solutions to puzzles, they are more prone to bypass cerebral strain if possible.
<u>Procedural description:</u>
Shooting: +
Jumping: -
Problem solving: -

4.2 Athlete Persona

<u>Narrative description:</u> athletes enjoy exploration of the environment; they will try to avoid fights if at all possible. Considering the amount of skill involved in performing precise jumps we can assume that they are players with experience and are comfortable with navigation controls, this points towards a high proficiency. Taking for granted proficient navigation skills, it is possible to assume a good sense of direction, athletes will rarely loose direction and will display relatively fast completion times, this also reflects on relatively few requests for help for spatial puzzles
<u>Procedural description:</u>
Shooting: -
Jumping: +
Problem solving: -

4.3 Chess-Player Persona

<u>Narrative description:</u> Chess-players get their biggest kicks out of solving puzzles, they will log almost no requests for help, there will be very few deaths caused by

mistakes in solving puzzles, they also possesses a good sense of direction that helps a lot in navigating the environments leading to few deaths caused by the environment. At the same time they might not necessarily be good at precise jumps hence they could die several times by falling. They are not expected to show particular skills with a gun. They could resort to asking for hints to solve certain puzzles, but never to complete answers.

Procedural description:

Shooting: -

Jumping: -

Problem solving: +

5 Play-Personas as Relations between Game Parameters

Now that all the possible combinations have been manifested it is possible to plot them as relations between parameters (Fig. 4).

The personas as metaphors, defined as one-sided accentuations, form the hypotheses around which game and level designers plan, produce and test their work. Because of the historical legacy of the *Tomb Raider* franchise, it could be expected that there will be a bias of players expecting physical and mental challenges in the form of jumping and puzzle solving experiences. Even though the shooting component has always been a part of the *Tomb Raider* series of games, it has never been the main focus, and players looking for a full fledged shooting experience would probably pick a different title, nevertheless the developers consciously decided to strengthen the combat elements and include a varied array of enemy NPCs for players to express shooting tendencies.

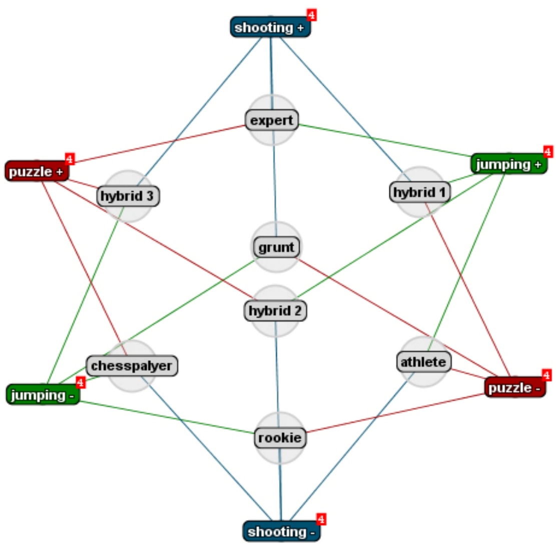


Fig. 4. Persona hypotheses emerge as relations between parameters that have been derived from gameplay mechanics

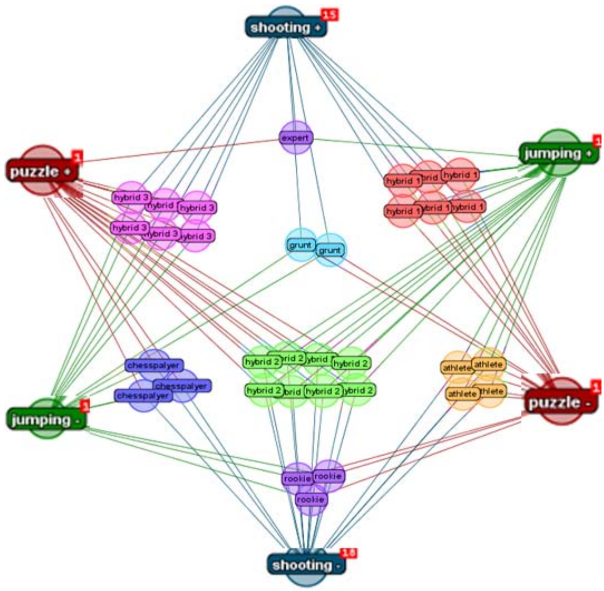


Fig. 5. Personas as lenses show the statistical layering of the population of players condensed into clusters of similar behavior profiles

The next step consists of individuating which game mechanics and variables translate as high or low puzzle solving, shooting and jumping skills, and how these can be tracked via gameplay metrics.

For example, players have the possibility to request hints and clues on how to solve puzzles in the game or they can ask directly for the complete answer via the Help On Demand-system. Thanks to the metrics tracking system developed internally at EIDOS [22], it is possible to collect data regarding players' behaviour. The decisions of hundreds of thousands players around the world to ask for help solving puzzles can therefore be collected, analyzed and form the basis for detailed persona models that relate directly to game mechanics. This information helps designers answering straight forward questions such as *"where do players get stuck more often?", however, gameplay metrics datasets – depending on the variables tracked – also provide statistical evidence to the subdivision of players in the various persona-defined categories* (Fig. 5).

For Tomb Raider: Underworld, data on player progression, level completion times, locations of player death, requests through the Help-on-Demand system, causes of death and a variety of other variables are tracked. Via the application of multivariate statistical methods such as cluster analysis, factor analysis and population statistical methods such as ordination/correspondence analysis, patterns of player behavior in the data can be sought out. Additionally, neural networks, decision three analysis and similar techniques can be used to locate data-driven patterns. This permits game developers to check how and if the pre-defined persona models (as metaphor) actually emerge in the way the end users are playing the game, i.e. as data-driven personas-as-lens [22].

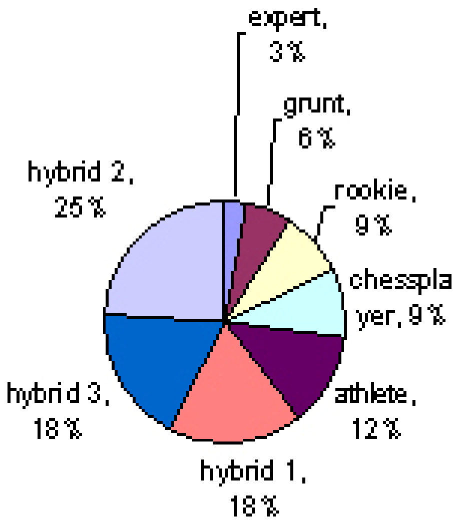


Fig. 6. Distribution of players according Tomb Raider Underworld persona profiles

In the case example for Tomb Raider: Underworld, it is apparent that the large majority of players do not fit under the umbrella of the three personas used as hypotheses, but rather is to be found in the profiles of usage defined by the hybrid personas (Fig. 6).

It is in this phase that persona hypotheses are checked against the numbers: As expected the Grunt is not very well represented, while the Athlete is the largest "pure" persona. This kind of information generates knowledge that has both short and long term effects: On a concrete level, the information provides valuable feedback to level designers, who can use the input to modify a level to achieve better balance and facilitate the different persona models; to obtaining a greater insight on the landscape of possible player types and eventually make games that can cater for a broader audience.

6 General Applicability of the Method in Extreme Cases

The suggested framework could potentially come short of being in any measure exhaustive when applied either to puzzle games with one solution or extremely linear games with very few areas designed for the player to make non-trivial choices. For example, in games such as Braid [6a], where there is only one way to solve the puzzles, it makes no sense to even talk about play-styles let alone personas: the player's sole duty is to second guess the designer's mind and therefore push forward the story. Similarly, in classic point-and-click adventure games, players had to advance the story by collecting and putting together a narrative that the designers exploded and spread around the game world, with no influence on what to do nor how to do it.

Even modern AAA-level games sometimes fail to provide non trivial choices, collapsing all the potential personas into only one profile that players have to conform to in order to proceed. This type of games would not particularly benefit from the application of the method, but it's important to notice how they do not represent the majority of titles on the market, often populating the casual end.

Beside these considerations, the play-persona framework is directly useful in situations where players have more than one way to accomplish a task, any time that the combination of different game-rules in different ways could give rise to something similar to "play style", every time that players are allowed to express a preference towards a course of action instead of another, even if this only means choosing between killing a NPC with a silenced gun or with poison, as is the case in the Hitman series developed by IO Interactive.

7 Conclusions

A key design challenge in an increasing number of game development projects is to ensure optimal experience for a variety of player preferences, in order to reach as broad a target audience as possible [11, 15]. In the above, a tool for addressing this requirement has been presented in the form of play-persona frameworks. These function both as models of preliminary hypothesis of in-game behaviour, and a means for categorizing and analyzing character-bound gameplay metrics variables. Play-personas have been shown to be a tool for informing the process of developing and testing a diversity of play behaviours in a computer game.

The play-persona framework aggregates data in a way that binds ludic and narrative aspects of the game. In this paper a case study has been presented that shows how play-personas allow designers to expand their role from mere dealers of rewards and punishments, disseminating challenges and skills for a general undistinguished audience, to become crafters, orchestrators of experiences and weavers of playing modes.

Acknowledgements. The authors would like to extend their gratitude to colleagues at IO Interactive, Crystal Dynamics, the Danish Design School and the IT University of Copenhagen.

References

1. Chapman, C.N., Milham, R.: The personas' new clothes: methodological and practical arguments against a popular method. In: Proceedings of Human Factors and Ergonomics Society 50th Annual Meeting, pp. 634–636 (2006)
2. Cooper, A.: The Inmates Are Running the Asylum. SAMS Publishing, Indianapolis (2004)
3. Cooper, A., Reimann, R., Cronin, D.: About Face 3: The Essentials of Interaction Design. Wiley Publishing, Indianapolis (2007)
4. Damasio, A.: Descartes' Error: Emotion, Reason, and the Human Brain. Penguin, London (2005)

5. Davis, J., Steury, K., Pagulayan, R.: A survey method for assessing perceptions of a game: The consumer playtest in game design. Game Studies: The International Journal of Computer Game Research 5 (2005)
6. Eco, U.: The Role of The Reader. Indiana university Press, Bloomington (1984)
7. Frijda, N.H.: The Laws of Emotion. Lawrence Erlbaum Associates Publishers, Mahwah (2007)
8. Goffman, E.: The Presentation of Self in Everyday Life. Penguin, London (1990)
9. Iser, W.: The Implied Reader. Johns Hopkins Paperback Editions, London (1978)
10. Kim, J.H., Gunn, D.V., Schuh, E., Phillips, B.C., Pagulayan, R.J., Wixon, D.: Tracking Real-Time User Experience (TRUE): A comprehensive instrumentation solution for complex systems. In: Proceedings of CHI, pp. 443–451 (2008)
11. Lazzaro, N.: Why We Play Games: Four Keys to More Emotion Without Story (2004), http://www.xeodesign.com/xeodesign_whyweplaygames.pdf
12. Mandryk, R.L., Inkpen, K.M., Calvert, T.W.: Using Psychophysiological Techniques to Measure User Experience with Entertainment Technologies. Behaviour & Information Technology 25(2), 141–158 (2006)
13. Medlock, M.C., Wixon, D., Terrano, M., Romero, R.L., Fulton, B.: Using the RITE method to improve products: A definition and a case study. In: Proceedings of UPA (2002)
14. Nacke, L., Lindley, C.A.: Flow and Immersion in First-Person Shooters: Measuring the player's gameplay experience. In: Proceedings of the 2008 Conference on Future Play: Research, Play, Share, pp. 81–88 (2008)
15. Pagulayan, R.J., Keeker, K., Wixon, D., Romero, R.L., Fuller, T.: User-centered design in games. In: The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications, pp. 883–906. Lawrence Erlbaum Associates, Philadelphia (2003)
16. Pine II, J.: Mass Customization: The New Frontier in Business Competition. Harvard Business Press (1992)
17. Ravaja, N., Saari, T., Laarni, J., Kallinen, K., Salminen, M., Holopainen, J., Järvinen, A.: The Psychophysiology of Video Gaming: Phasic Emotional Responses to Game Events. In: DiGRA conference Changing views: Worlds in play. Digital Games Research Association Press (2005)
18. Ravaja, N., Saari, T., Turpeinen, M., Laarni, J., Salminen, M., Kivikangas, M.: Spatial presence and emotions during video game playing: Does it matter with whom you play? Presence: Teleoperaters and Virtual Environments 15(4), 381–392 (2005)
19. Sanders, E.B.-N.: Scaffolds For Building Everyday Creativity, in Design for Effective Communications: Creating Contexts. Allworth Press, New York (2006)
20. Shils, E.A., Finch, E.A.: The methodology of the social sciences. Free Press, Glencoe (1997)
21. Swain, C.: Master Metrics: The Science Behind the Art of Game Design. Presentation at NLGD Conference, Utrecht, Holland (2008)
22. Tychsen, A., Canossa, A.: Defining Personas in Games Using Metrics. In: Proceedings of Future Play 2008, pp. 73–80 (2008)
23. Ó, Reilly, T.: What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software, September 30 (2005), http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html
24. Goetz, P.: Too Many Clicks! Unit-Based Interfaces Considered Harmful. Gamasutra, August 23 (2006)
25. Hilbert, D.M., Redish, J.C.: A practical guide to usability testing. Intellect books, Bristol (1999)

Ludology

1a. *Left4Dead*, Valve, Electronic Arts, (2008)
2a. *Little Big Planet*, Media Molecule, Sony (2008)
3a. *Spore*, Maxis, Electronic Arts (2008)
4a. *SingStar*, SCE London, Sony (2008)
5a. *Tomb Raider Underworld*, Crystal Dynamics, EIDOS, (2008)
6a. *Braid*, Number None Inc. (2008)

Developing and Validating Personas in e-Commerce: A Heuristic Approach

Volker Thoma[1] and Bryn Williams[2]

[1]University of East London, Romford Road, London E15 4LZ
[2]Betfair Limited, Winslow Road, London W6 9HP

Abstract. A multi-method persona development process in a large e-commerce business is described. Personas are fictional representations of customers that describe typical user attributes to facilitate a user-centered approach in interaction design. In the current project persona attributes were derived from various data sources, such as stakeholder interviews, user tests and interviews, data mining, customer surveys, and ethnographic (direct observation, diary studies) research. The heuristic approach of using these data sources conjointly allowed for an early validation of relevant persona dimensions.

1 Introduction

One of the greatest challenges in user centered design is the elicitation and communication of user requirements. Results from traditional methods such as questionnaires and interviews are often too complex or too cumbersome to digest [1]. Furthermore, design specifications can fairly quickly become out-of-date and may simply not be adhered to by the development team. To resolve these common problems, Cooper [2] proposed the use of "personas" to achieve a more effective tool to elicit and convey user requirements. Personas are fictitious, yet specific and concrete representations of target users. The description of a persona usually consists of a name, basic demographics, motivations, product usage patterns, personal likes, etc. The aim of personas is to inform design and facilitate communication among stakeholders [1]. The persona method has gained a substantial following among developers [3] and is now a standard component of user experience projects in many companies. Advocates claim that personas engage teams to think about users, are easy to use for designers and developers, and help overcome the typical problems associated with more complex or statistical representations of the user base.

The general approach to persona development is to gather information about users' needs, behaviors, and preferences, which are collated and used to construct vivid descriptions about admittedly fictional individuals [4, 5]. The exact method of how to derive and construct personas is a topic of debate [1]. Grudin and Pruitt [4] advocate a multiple-methods approach, with the use of a range of qualitative and quantitative data to achieve greater accuracy and utility for persona creation. Unfortunately, because of lack of time or resources some practitioners select either purely qualitative or purely quantitative methods in the development of personas. Such limitations have their own problems, such as running the risk of being dismissed by stakeholders as

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 524–527, 2009.

purely creative exercises [6] or not having the depth of understanding of the users behaviours to generate fully-rounded personas [4]. Finally, it is very difficult to verify that personas are accurate [3]: how can we test whether the combination of attributes actually represents any users? In summary, any persona development project faces three major questions: Which and how many methods to use? Which procedure to follow? And can one validate personas? The main aim of the paper is to illustrate – using a case study – how data gathering methods can be prioritized, adapted and used conjointly for persona development projects. The issue of cost is beyond the scope of this paper, but our results help to focus on core methods for persona projects. A particular focus is on different ways to validate personas early in the research process.

2 Persona Research, Validation and Creation

A large exchange betting company requested a set of company wide personas which would guide design processes as well as product development decisions. Betting exchanges offer online betting in a "many-to-many" form, where users are betting against each other using the betting exchange platform. Additionally, the betting exchange website offers other online products (e.g., poker, casino games).

The process for developing personas consisted of four phases which partly overlapped. The first phase comprised of gathering qualitative and quantitative data and included a review of in-house knowledge. Qualitative data such as observations and interviews were analyzed using a simplified thematic analysis approach [7]. Quantitative data were analyzed using cluster analysis techniques and principal components analysis [8]. The data gathering methods used are summarized in table 1. Phase 2 involved the organization and coding of datasets to allow for the extraction of high-level patterns and themes while preserving the details of the data for later analysis. In phase 3 these themes and high-level patterns were used to define the major dimensions representing our users. Finally, in phase 4 (not described here), the personas were developed in detail drawing on the data and themes extracted earlier.

Stakeholder interviews (semi-structured) were conducted early on in the project. A common theme from stakeholders was that high volume users should be differentiated according to which products they use and their sophistication. In addition to interviews, workshops were held at various stages throughout the persona development process. From the results of these activities and other previous requirements engineering projects (customer surveys, market segmentation) we compiled a long list of categories and parameters that would potentially be used to describe personas, such as interests, betting behaviour, psychographics, motivation, media profile, and many more. Over 200 categories and subcategories were recorded and were entered in a dynamic "master" spreadsheet and regularly updated.

We then carried out user interviews (n=24) and also directly observed eight users in their natural context to elicit typical motivations, needs, context of use, etc. Home visits were arranged in advance and users were instructed to keep a week-long "diary" of their betting related experiences. These methods provided direct context-driven information that would be difficult to otherwise uncover (e.g. sharing of accounts, social aspects of betting, etc.). The extracted themes were around the expertise with betting and emotional involvement with a particular sport or product. In addition, a clear distinction was found between sport-only bettors and those also using the casino/games products. A further method to elicit persona information was user testing. Although usually centered on prototype development and testing, we designed the user tests to probe clients for persona relevant information. Emerging themes here were the reliance on old navigation habits, the lack of awareness of other products, and betting related features on the site that made betting more complex. User tests also validated themes that had previously emerged. For example, the inclusion of minimized games within the sports betting page was tested on users with different product usage profiles. Five out of six casino users spotted the minimized game, whereas only one out of six non-casino users found the game, thus validating the dimension of product-awareness and usage. A further way to validate previous themes was through data mining. An analysis of navigation paths showed that the way to navigate to customer's favorite markets depended on how experienced they were and how often they betted, validating themes extracted from user tests. Finally, a customer panel with over 1000 users provided feedback to electronic and phone surveys. This data base was linked to the data warehouse which allowed for the analysis of usage patterns across clients. Quantitative patterns (clusters) emerged on volume of product use (e.g., sports betting, games, casino, etc.) and channels (online, mobile).

After data were collated from all sources, emerging themes and patterns were extracted [7] and summarized as various persona "dimensions" (see [1,5] for structuring persona attributes). Initially 10 dimensions were extracted and then reduced to four, which were mapped onto a "persona space" (Figure 1). A persona could vary in: 1. betting volume, 2. passion or emotional involvement in a product (e.g. horse betting), 3. degree of betting sophistication (from punting to complex and automated trading), and 4. type of product usage (e.g. sports betting vs. Casino).

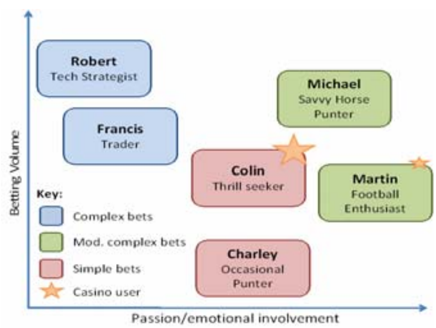


Fig. 1. Persona space

Table 1. Methods used in persona research

Data (phase)	Recommend for
Stakeholder Interview (early)	Themes
User tests/interviews (throughout)	Qual. & quant. data, validation
User observations User diaries (early)	Qualitative data, themes, quotes
Stakeholder workshop (early/late)	Qualitative data, validation
Data mining (throughout)	Quant. data, validation
Customer survey (throughout, late)	Quant. data, validation

To summatively validate these dimension, 24 variables from the user panel response (n=408) were subjected to an exploratory factor analysis (oblique rotation). The variables represented quantitative information on usage patterns and Likert-scaled answers on betting related questions. The first of the four main factors extracted loaded highly (partial correlations > .48) on variables such as: Joy from betting, strategic betting, betting as income, information based betting, betting on specialty areas. The second factor loaded (> .20) on: Revenue (volume), profit & loss, mobile betting volume, in-play betting volume, Casino betting volume; The third factor loaded on (> .16): Age, Internet experience, technical know-how, betting experience, frequency of Internet use, betting habits, product loyalty. The fourth factor loaded (>.17) on: Perceived lack of know-how, betting as entertainment, amount of online gambling; and games volume. This result confirmed the previously derived themes from qualitative sources (the variables loading on the fourth factor reflect the various motivations of using different products such as games and casino). The dimensions and the "locations" of different personas were further validated and fine-tuned in stakeholder workshops. The persona space greatly facilitated the validation and communication of persona dimensions to stakeholders. In addition, it allowed the team to quickly produce and test early persona sketches.

3 Conclusions

We began this article by describing the three main problems of persona creation: the choice of methods, the procedure, and validation. We propose a heuristic approach that includes a limited number of both qualitative as well as quantitative methods, including user observations, user tests, in-depth interviews, data mining, and surveys. Persona development must be based on both qualitative and quantitative data if a fully rounded and believable set of personas is to be developed [1, 4]. A second important point is rather than using a particular method in isolation, we recommend using qualitative and quantitative data gathering methods in parallel to extract relevant themes. Further, we propose validating the extracted persona dimensions early on in the process: through triangulation (comparing with themes from different methods) as well as by using quantitative data reduction (here: factor analysis). The use of a persona space to represent relevant dimensions is a pragmatic solution to validate that personas reflect true user groups, while at the same time providing a framework for how to conduct research and integrate the findings across methods [4]. We contend that although we may not have validated personas in a fully scientific sense [3] or exhausted other (qualitative) approaches of validation, the approach described here is a practicable approximation to ensure the creation of relevant and cogent personas.

References

1. Pruitt, J., Adlin, T.: The persona lifecycle. Morgan Kaufmann, San Francisco (2006)
2. Cooper, A.: The inmates are running the asylum. Macmillan, New York (1999)
3. Chapman, C.N., Milham, R.: The personas new clothes. Human Factors and Ergonomics Society (HFES) 2006, San Francisco, CA (2006)
4. Grudin, J., Pruitt, J.: Personas, participatory design, and product development: An infrastructure for engagement. In: Proc. PDC 2002, pp. 144–161 (2002)
5. Pruitt, J., Grudin, J.: Personas, Practice and Theory. In: Proc. DUX (2003)
6. Mulder, S.: The User Is Always Right: A Practical Guide to Creating and Using personas for the Web. New Riders, Berkeley, CA (2007)
7. Braun, V., Clarke, V.: Using thematic analysis in Psychology. Qualit. Res. Psy. 3, 77–101 (2006)
8. Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C.: Multivariate data analysis. Prentice Hall, New Jersey (1998)

Picking Up Artifacts: Storyboarding as a Gateway to Reuse

Shahtab Wahid[1], Stacy M. Branham[1], Lauren Cairco[2], D. Scott McCrickard[1], and Steve Harrison[1]

[1] Center for Human-Computer Interaction, Virginia Tech, Blacksburg, VA 24060
{swahid,sbranham,mccricks,srh}@cs.vt.edu
[2] Department of Computer Science and Quantitative Methods, Winthrop University, Rock Hill, SC 29733
caircoL2@winthrop.edu

Abstract. Storyboarding offers designers the opportunity to illustrate a visual narrative of use. Because designers often refer to past ideas, we argue storyboards can be constructed by reusing shared artifacts. We present a study in which we explore how designers reuse artifacts consisting of images and rationale during storyboard construction. We find images can aid in accessing rationale and that connections among features aid in deciding what to reuse, creating new artifacts, and constructing. Based on requirements derived from our findings, we present a storyboarding tool, PIC-UP, to facilitate artifact sharing and reuse and evaluate its use in an exploratory study. We conclude with remarks on facilitating reuse and future work.

Keywords: design reuse, storyboarding, claims.

1 Introduction

Storyboards are visual stories of actors engaging in a series of actions for a certain purpose. Typically, they consist of multiple panels made of pictures and an accompanying narrative that illustrates the progression. Used by those involved in the creation of movies, cartoons, and commercials, they are powerful tools in highlighting the most important aspects of a narrative [7],[9]. In Human-Computer Interaction (HCI), storyboards have been used within the design process to illustrate how users may interact with a system [4],[18]. By leveraging graphical descriptions of artifacts, they prove to be powerful tools in pitching and communicating ideas for new technologies.

But what are the artifacts that will be included in the storyboard? Choosing artifacts is key as it is a reflection of the vital functionality that a proposed technology might support. Designers strive to derive new artifacts through creative efforts to improve the design of their products. This involves determining the nature of the artifacts, how they connect to other artifacts, and the visual elements that will represent them.

It is also true that practitioners often informally continue to turn to features that have been used before because they have been proven to stand the test of time. Everyday tools such as remote controls and microwave ovens have maintained the same

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 528–541, 2009.

basic design principles for many years. While new iterations continue to be produced, designers naturally reuse techniques of the past—providing impetus behind an often ignored need to gain and even build on what has been used before [24]. Improvements based on prior work can enable designers to potentially continue to maximize successes while avoiding or mitigating identified pitfalls and lowering costs [2]. In fact, the benefits of reuse on usability have been demonstrated [23], justifying efforts within HCI that have investigated ways of facilitating the reuse of various components often in the form of design knowledge [3],[11],[12],[20]. Although this might seem fruitful, designers exposed to collections of reusable artifacts are often burdened by the large amount and restricted by tools that enforce inflexible design processes, hindering reuse [10].

We propose storyboarding as a fitting activity to provide easier access to design knowledge and aid designers in determining what and how artifacts can be reused. This offers us the chance to combine reusable claims, rationale encapsulating tradeoffs of design features [5], with inspiring images. While pictures can ease access to reusable knowledge, claims serve to familiarize designers with features. Throughout the process of creating the storyboard, identifying and understanding the relationships and connections that exist between components can aid designers in choosing and creating components used in the construction of the final storyboards.

Our goal is to develop a tool that supports storyboarding through reuse. In this effort, we present a study to discover the nature of what the tool must embody. We study groups of designers asked to construct storyboards for a given system by reusing cards consisting of pictures and claims. Results of the study demonstrate that the cards and the relationships identified by designers were useful in jumpstarting efforts to build a prototype through reuse. We then present a tool developed on the results acquired from the study and a preliminary study of its use. We conclude with remarks about lessons we have learned and future efforts.

2 Related Work

Storyboarding is the process of describing a user's interaction with a system over time through a series of graphical depictions, often sketches, and units of textual narrative. Storyboards have been used to help understand the flow of the story, to eliminate costly elements of a design, and even to decide how to pitch ideas to others [4],[17]. Thus, they can be seen as early prototypes in the design process. Key aspects of a storyboard are the portrayal of time, the inclusion of people and emotions, the inclusion of text, and the level of detail [21]. Tools such as SILK [13], DENIM [15], and DEMAIS [1] facilitate storyboarding to create prototypes early in design. However, such tools do not support sharing and reusing artifacts.

Design reuse has been researched by many to lower development time and costs [6]. Its value has been seen in industry in different situations. For example, the IDEO Tech Box serves to collect and store objects that can inspire and improve the design of products [4]. Patterns, reusable knowledge structures that incorporate contexts of use, conflicting forces, and potential solutions, have been adopted by those in HCI as a form of design rationale [3] and by those in software engineering [8]. The Yahoo! Patterns Library is an example of a repository that stores reusable components for web design [25]. However, such tools lack an overarching design activity, such as storyboarding, that leads to a prototype.

Claims are a way of recording knowledge initially proposed by Carroll and Kellogg [5]. Delivered in informal natural language, claims address a variety of situational and interface aspects that affect the compatibility of the design and user models, such as user satisfaction and feeling of reward, color and object layout, and strength of affordances. They were later proposed as design knowledge units that could be stored in repositories for reuse by designers [20]. A claims library was also created to support their reuse [16]. More recently, they were even included as a part of the design pattern structure by Hughes [11]. However, because claims focus on design rationale, creating tools can be difficult since they often face challenges in capturing, using, and organizing rationale [10].

3 Uncovering Storyboarding through Reuse

To investigate how we may provide novice designers access to reusable design knowledge in new ways, we conducted a study in which designers took part in a storyboarding design session. We were particularly interested in how the participants balanced the use of pictures and rationale, connections established among the components, creation of new artifacts, approaches taken to completing the design tasks, and the structures of the storyboards.

3.1 Participants

Twenty-one graduate students actively engaged in conducting HCI research or enrolled in a graduate HCI course participated in the study. Their familiarity with claims and storyboarding varied widely, giving us a broad look at the potential ways in which the participants could be impacted by the artifacts.

3.2 Materials

A video camera was mounted over a table on which thirty cards describing design features were scattered (see Fig. 1). The front of each card had a picture representing the feature along with a label. The back of the card had a claim for the feature. Blank pieces of paper and pens were also provided, as well as an instruction sheet that explained the task, definitions for concepts such as claims and storyboards, and a prepared design problem.

Each group was given a unique design problem that centered around the need for a notification system. Notification systems are designed to be used in dual-task situations in which a user engaged in a primary task needs to be interrupted for a reason [14]. Some of the design problems we assigned involved notifying nuclear plant operators of changing core temperatures, passengers in airports of flight status changes, commuters of empty parking lot spots while driving, theme park visitors of ride wait times, and students of empty spots for classes they wish to register for.

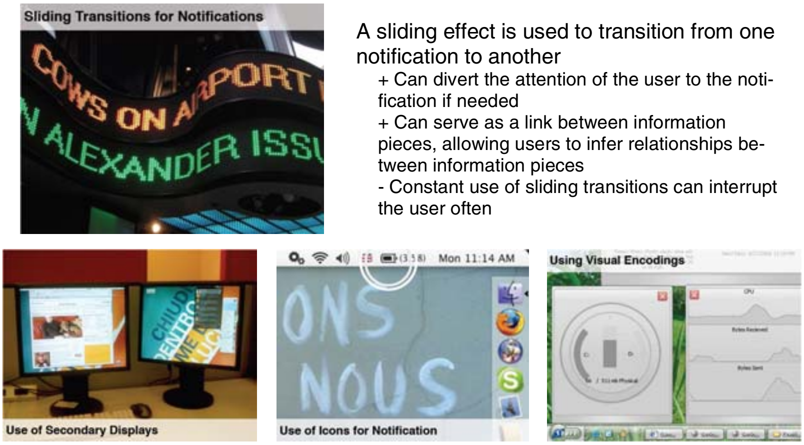


Fig. 1. The front of the cards have pictures illustrating the design feature along with labels (top left). The back describes the consequences of using the feature in the form of a claim (top right). Additional examples of other cards are shown in the bottom row.

3.3 Procedure

The participants were randomly divided into 7 groups with 3 in each group. Each group was asked to create a storyboard with 4-7 panels representing a system that would solve the given design problem. They were given the option to create their own cards if they felt they needed a new feature. However, they were told that the new cards must have pictures and claims. If a new card was created, the card was kept in the pile for subsequent groups to see and use. Upon completion of the storyboard, they were asked to write a narrative for the storyboard describing a usage scenario. Each group was told they had 40 minutes, but we did not stop groups that went over the time limit. Once the storyboard was completed, the participants were interviewed as a group for approximately 30 minutes. Interview questions revolved around explaining the constructed storyboard, usage of the cards, and creating cards.

3.4 Analysis

We took a grounded theory approach and adopted the open coding technique to analyze the data [19]. Our categories of analysis included analyzing pictures and claims, connections between cards, decision-making, creating cards, task approaches, and storyboard construction. Each group video was analyzed by two coders who watched the complete videos and identified critical points of interest.

3.5 Results and Discussion

In this section we present the results focusing on the qualitative data that was acquired from the study. The participants displayed interesting behavior regarding the use of the cards, the connections established between cards, approaches to carrying out the task, and structure of the storyboards. We describe these activities, but expand on only some of them due to space limitations.

Use of Imagery. For the designers, a critical part of carrying out the activity was familiarizing themselves with the artifacts. Therefore, the participants spent time looking at the cards, sharing them with each other, and referring to them as they talked about possible design ideas. On average, the groups explored cards 86 times. Group 4 explored the cards the most, 120 times, while group 5 investigated cards the least, 56 times. For the majority of the time all the groups kept the cards with the pictures facing up at all times.

The use of pictures to summarize the idea of the reusable artifacts was critical to how the participants explored the cards. The images provided them with ways of recognizing the cards. This was important because of the number of cards that were scattered across the table during the study sessions. One participant from group 1 summarized her thoughts by mentioning the picture, "*…was the main thing. I first looked at the photo and then read through.*" A participant in group 7 emphasized the utility of the pictures: "*I guess the picture gives us a very…makes it easier to identify the object so it stands out from the other objects. The pictures are quite different so maybe we employ less effort to recognize the objects instead of going through the text.*" In response to this comment, a participant in the same group said, "*If you had pieces of paper with text on them it would be harder to sort through them.*" These comments tell us that the use of imagery can potentially make it easier to find and access rationale that would have otherwise been harder to sift through.

The pictures also served to facilitate creative thinking for the design task at hand. For example, one card about showing trends in data had a picture of a graph showing changes in values over time. Group 1 was told to create a system that would allow commuters to find empty parking lot spots. A participant from this group looked at the trends card and said, "*you see [some spot] is empty so you click that one and it just opens up…[shows] when it got available and also this graph…you know…gives you an idea about what time of the day it gets full and what time of the day it gets [empty] and stuff like that. That might be a cool thing I think.*" In this case we notice that feature was directly taken as a possible component for the system being created. There were other cases where an idea was extracted from a picture, but was unrelated to the feature that the picture was representing—perhaps a more creative act. For example, a card about relating preexisting user knowledge to a notification had a picture of a chat window with a chat history. A participant from group 4 picked up the card and said "*if you want a timestamp…then [use] something like this.*" The timestamps happened to be a part of the picture, but were not necessarily there to illustrate the idea of the card. This demonstrates creativity may play a vital role in reuse, potentially making a positive impact on the product. Linking pictures to the claims provided designers with a way of quickly recognizing artifacts, giving them quick access to reusable components that might have otherwise been lost and lowering the burdens of dealing with large amounts of rationale. However, perhaps the more promising utility lies in their ability to inspire, establishing a balance between creative thought and the rationale associated with the images.

Use of Claims. We observed many situations where the participants felt a need to flip the cards over to read the claim for the artifact. On average the groups turned the cards over 20.7 times. Group 4 flipped cards the most, 33 times, and group 1 flipped cards just 10 times.

The claims served as a way for the participants to familiarize themselves with the details of the feature and the possible consequences of using the feature. Speaking of the claims, a group 4 participant said, "*I guess it clarifies what the creator of these [cards] intended to mean by each of these.*" It served to clarify thoughts if designers were confused about the cards because they misinterpreted or were unsure of what a picture meant. In one instance we saw two participants in group 7 who were looking at a card about personalizing notifications were unsure about what the feature was about. They flipped it over and one participant said, "*this seems to be more about the mode of notifications rather than…it's more about the mode rather than the content [of notifications].*"

Previous work identified a set of relationships that exist among claims during the design process [22]. This study served to demonstrate that designers, without knowledge of the idea of claim relationships, identify and leverage relationships during their design work once they are familiar with the claim. For example, one common relationship was the idea of generalization vs. specification. Several groups identified a card about notification systems as having this relationship with many of the other existing cards. One person in group 2 mentioned, *"[the] notification systems [card] seems too general for this task. It seems it should be the entire overarching second half of our storyboard.*" In this case, because the group understood the relationship, they were able to decide how relevant the card would be.

Another common relationship that was critical to decision-making involved comparing alternatives. In group 1, a member found two cards on sporadic notification and continuous notification. Having identified through the claims that they are alternatives to each other, the person said, "*sporadic vs. continuous…this is something we need to think about. Will it be continuously displayed on your handheld?*" A participant in group 3 identified the relationship between voice interaction and touch interaction and then asked, "*can we compare the upsides and downsides of the two features to see which one is better?*" Comparing alternatives is an important task. Realizing this relationship allows the designers to think about the larger impact reusing the feature could have on the system that is being created.

One advantage of using claims is that it makes designers aware of potential downsides that might need to be mitigated. In such cases, there might be another claim that can serve to reduce or eliminate the negative effects of a claim. This relationship was also realized by a member of group 3. While reading the claim on a card she said, "*may miss notification and wait…but we can compensate it with this one. If they miss it they will get a continuous notification like an icon on the corner [of the screen].*" Once again, being aware of the tradeoffs and then immediately identifying potential solutions through the connection established paved the way to reusing the card in question.

Finally, two other connections that were observed, but not always explicitly talked about, involved combining cards together and linking information to interaction elements. Often cards were combined together such that both cards would be used in conjunction to represent a new idea. For example, a new notification method could be created by combining a card about a blinking light with an audio notification card. The participants also linked cards appraising the state of an interface to cards about interacting with the elements in the interface. As the participant in group 1 who wanted to use the trends in data card mentioned, the use of the card was also being tied to how a user would interact with the feature if it were to be included. These two relationships were used more often in the construction of the storyboards.

As demonstrated through these examples, claims play an important role in the design process. Just as pictures contributed much to the activity, the rationale presented through these claims is often critical to how the participants learn and express ideas to others. When needed, the claims can be used to argue for or against a line of reasoning. The identified connections force designers to think about choosing between claims and how the claims will eventually come together in the storyboard.

Creating New Cards. During the design session we also gave the participants the option to create their own cards if they felt it was needed due to the nature of their design. Because we knew that designers often resist formally contributing knowledge, we wanted to explore how cards were created when needed.

There were only three design sessions, groups 1, 4, and 7, in which cards were created. In all three cases, the groups had concluded that the new card was essential to the design and therefore needed to be created. To illustrate the process we explain how one group created such a card. Group 1 created a card called *using geospatial representation of information* so that they could use it for their commuter parking lot system. The conversation that took place during the creation of this card turned out to be quite interesting. First, the person who was writing the claim said, "*What was the graphical [card] one? How was that phrased? I lost it…so using geospatial representation of information? I almost feel like…they're related or something.*" As he proposed a name for the new card, he realized that this card could potentially be similar to a card about graphical information. His need to refer to the other card demonstrated that he wanted to maintain the same level of scope, making it generic and trying not to over-specify the card so that its potential reuse would not be restricted. Thus, his authoring of the claim was influenced by the claims that were already around him. While referring to the graphical information card, he then discussed the possible tradeoffs for the new card: "*Yeah, I think the last two [tradeoffs] are particularly relevant. I guess you can specify them for geospatial. If the person doesn't know how the parking lot is laid out…if they see it from overhead…they might understand it the first time they see the map…that sounds like a pretty similar downside…*" Finally, once the card was created the participant reflected, "*I feel like we're specializing this one...cause it's a kind of graphical information, but it's a special kind. It brings its own pros and cons to an extent in terms of understanding the information. It's about understanding the place, not just the representation of it…*" His reflections serve to demonstrate that other cards can influence how new cards are created and that new cards can also be a result of a noted connection. In this case, identifying that relationship with the graphical information card served the group well, allowing them to create a unique and reusable card. This card was eventually reused by group 7 in their own storyboard.

Creating new artifacts is important to reuse because it provides a fresh source of knowledge. While the burdens of creating a new card in terms of content might have been lowered by introducing the picture and the simple structure of claims, we noticed that other factors such as the designer's own knowledge and trust influence whether a card is created. This might be overcome, however, if the designers were exposed to more cards that look like they were contributed by other designers.

Building Storyboards. While we observed rich interactions as the complete storyboards were constructed, we focus on reporting on the cards in the structures of the storyboards and the strategies used to assemble them. On average the groups had 4.8 panels in their storyboards (where each panel was one or more cards marked by a single section of the written narrative). Group 2 had the most with 6 panels while group 6 had just 3 panels. An average of 2.4 panels contained multiple cards within them. All of the group 5 panels had multiple cards in them. Group 1 and 4 did not have any such panels.

The structures in the storyboards continued to represent the reliance on establishing connections between cards. A common connection was the connection between displaying an artifact and interacting with the artifact, as acknowledged by group 3 participants when they sequenced their cards accordingly. As we mentioned earlier, cards were often combined together to represent a form of fusion. Group 4 was the only group to structure their storyboard such that alternate paths were represented. At one point, their storyboard split into three possible routes, representing different notification methods based on preference, and then converged again. We find it is not only necessary to demonstrate what is being reused, but also how it is being reused—an aspect reflected through the structure of the representation.

We spent time investigating the strategies used throughout the design activity. In general we found the groups typically first familiarized themselves with the cards, made choices, and then constructed the storyboard. Group 3 was the only group that decided to try to group all the cards into categories such as displays, notification methods, and input methods while familiarizing themselves with the cards.

In determining how the group should approach the task, we noticed a clear tension between determining a scenario to inform the design and letting the features determine a possible scenario. Group 1 was the only group that clearly determined a scenario first and then proceeded to create the storyboard. The rest of the groups allowed the combinations of features to determine a possible scenario, but groups 5 and 6 did engage in some initial discussion of a possible task flow. While many groups engaged in a debate about how to approach it, a participant in group 3 put it this way: "*So how would you want to start this? Do you want to put individual pieces together and build a storyboard…and then build a story? There are two ways of going about it right? Either you take the abstract chunks and place it then form a storyline or decide [on] a story and then place the chunks.*" In group 1, we noticed that focusing on the scenario first led to a new card being created because the story was determined independently of what was being offered to them. One should consider whether a scenario-first approach can lead to increased contributions, but we do not have further data to explore this route. While an open and flexible design process can be beneficial, this important tension must be investigated further if we are to work toward proceduralizing storyboarding through reuse as it might have an impact of the quality of the storyboards.

4 PIC-UP

Our primary motivation for creating a tool for storyboarding through reuse was to facilitate the collection and sharing of design knowledge. We conducted the study to understand how those engaged in storyboarding through reuse would leverage components—giving us a chance to develop a tool that is grounded in actual experiences. The results allowed us to distill a set of goals for a tool that could support this activity:

Supporting recognition and ideation. To avoid having designers expose themselves to large amounts of textual information, images should be used to allow for quick identification. They will also serve as springboards for ideation because they can inspire and provoke thought that can be further solidified with access to rationale.

Discovery through relationships. Because artifacts do not exist in isolation, relationships among them can allow designers to explore other artifacts that might also be relevant. Using relationship types allows designers to find artifacts that are meant to fit the design in a certain way.

Encouraging reusable contributions. While the tool must enable quick and easy creation of artifacts, designers must also be encouraged to make their artifacts as reusable as possible so that others can take advantage of them. Through exposure to other artifacts and guidance, designers can be encouraged to create appropriately scoped and complete artifacts.

Guidance. While an open and flexible process should be adopted, a sense of the important tasks of exploring, deciding, and constructing storyboards should be fostered.

4.1 PIC-UP Description

We created a digital storyboarding environment called PIC-UP that builds on the activities carried out for our study and takes advantage of the additional storing, organizing, and guidance that technology enables. Although we focus on the experience on a traditional display, our application was designed to support nontraditional displays such as multiple display environments and tablet PCs.

While PIC-UP can certainly allow designers to create full storyboards from scratch, we focus on supporting creation through reuse. Thus, we provide access to an online library of claims in the form of cards to enable to construction of storyboards by leveraging a table metaphor. Each card appears with a claim title and image and, when necessary, can be flipped to show the complete claim on the back. The cards can be moved easily within the space, allowing users to sift through artifacts as they please.

While our goal was to guide novice designers in their efforts to build storyboards, we did not want to impose a strict process so that designers could be free to carry out the process in their own way. For this reason, we adopted a lightweight method that would illustrate the main types of activities that would need to be carried out. We divided the environment into three resizable spaces: browse, workspace, and storyboard (see Fig. 2). As cards are moved from one space to another, the designer is made aware of the possible tasks that may need to be carried out by referring to a menu on the left of each space.

Browse. The first space in which browsing activities are carried out addresses the need for quick identification of and brainstorming about cards that may be of value to the designer. While initially confronted with a scattered set of cards, users can narrow down and search through the cards in a number of ways with constant focus on images. They can choose to see the top 10 most used cards, a random set of cards, or

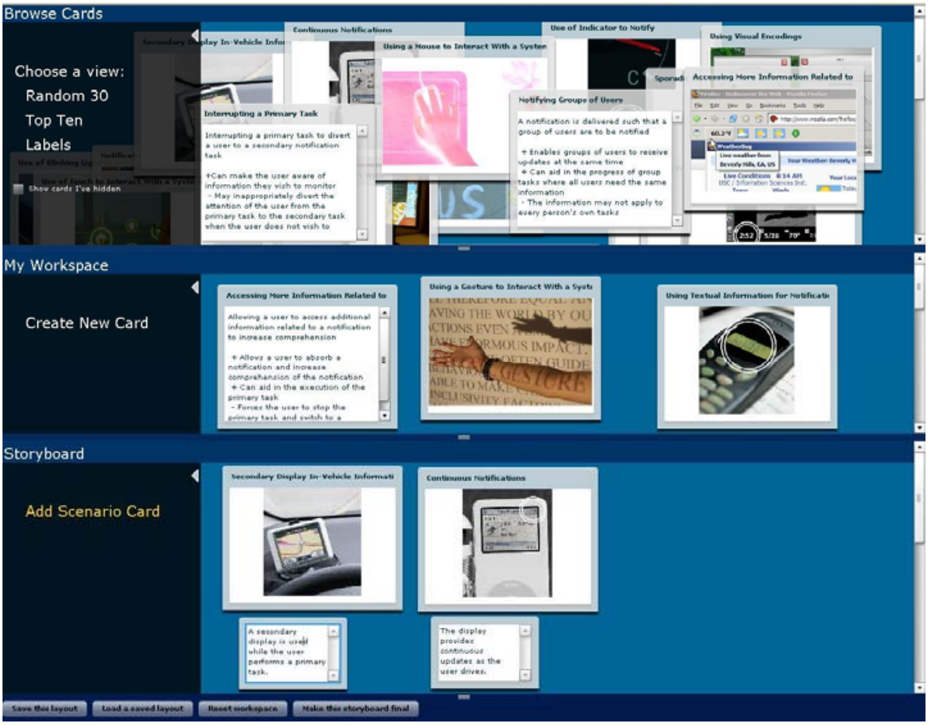


Fig. 2. A screenshot of PIC-UP showing three spaces: browse, workspace, and storyboard. Cards are moved from one space to another as users progress through the design activity.

cards that belong to a certain category through labels assigned to them. Labels such as notification methods, displays, and input methods are used to categorize the cards. Functionality to hide and restore cards is also provided.

A detailed view of the cards allows for further investigation. As we found in the study, the relationships established among the cards can play an important role during the reuse process. This enlarged version of the selected card includes thumbnails of related cards and allows for navigation to them. Relationships such as generalize, specialize, mitigate, fuse, execute, and evaluate are used to describe the connections we observed in our study.

Workspace. Once a card is chosen by a designer it is moved to the workspace. This space is intended to give users a space to collect and decide on cards to use in their storyboard. Users may continue to see the detailed view of cards and read the claims to familiarize themselves with the artifacts.

Because this space is used to collect components, it also gives designers the opportunity to further determine how suitable an artifact might be. In some cases the artifacts might be only somewhat appropriate. An artifact can be edited so that it better fits the context of the system being created. When an artifact is not available, the user may create a new card. Creating a new card involves adding a title, image, and full claim. We deliberately wanted to make the process of creating a new card quick and simple to reduce the burden on the designer. When the designer's work is complete, the edited or new cards can be stored for others to view and reuse. Because this creates a source for contributions to the repository, we try to encourage designers to make their contribution as reusable as possible. We believe that exposure to the cards in the browse space can inform designers as to how their own cards should be authored to make them reusable.

Storyboard. The cards that will make it into the final storyboard are moved from the workspace to the storyboard space. The space is used to sequence the cards to represent an appropriate scenario of usage. Our choice of using the table metaphor for our tool proves useful in this space because it continues to allow the designers to represent different paths and relationships in the arrangement of the cards. This section also allows the user to add scenario cards, blank text boxes that the user can type their narrative into and place underneath corresponding artifacts. Like the regular cards, these can be moved around in the same way, although they can only exist within this specific space.

4.2 Exploratory Study

To gather feedback on the tool we developed, we conducted an exploratory study with six undergraduate students. The participants previously participated in other sessions of the same storyboarding activity for a different study. Although we realize this gives them an advantage, our primary goal was not to observe their design habits, but to collect initial reactions to aid in further tool refinement to conduct future studies.

Each participant was asked to create a storyboard for a system that would notify a traveler in an airport when their flight was departing. We asked that the storyboard use at least two existing cards, at least one new card, contain three to five panels, and include a narrative describing the usage scenario to ensure they used all of the available functionality. Once the task was complete, participants filled out a survey responding to Likert-scale questions about finding cards and creating cards. They also provided comments on what they liked the most, what they found most frustrating, and any changes or additions they would make to PIC-UP.

We found all participants used between five and eleven cards to construct their storyboards. None chose to edit an existing card for use in their storyboard. We were surprised to find only two of the participants included a picture in the cards they created even though they were free to find an image online or draw an image themselves and import it. This points out that inserting images might be an impediment that could not be overcome easily or was perceived to be time-consuming. While we know there are benefits to having imagery, potential barriers to inserting imagery needs further study as it might be detrimental in the long run.

Bailey et al. found paper and pencil techniques were preferred to communicate design structure by juxtaposing artifacts in a physical space [1]. In our questionnaire, five of six participants reported they would prefer using the digital method for storyboarding over paper-based methods because of the drag-and-drop interface. Perhaps the physical nature of PIC-UP allowed them to juxtapose artifacts in ways that mimicked the findings of Bailey et al. The one participant who preferred paper-based storyboarding mentioned the physical cards would provide more freedom and "responsiveness" and might explain why images were not included in some cards.

Our questionnaire asked the participants how easy it was to find cards and create new cards on a scale of 1 (very easy) to 5 (very difficult). Participants ranked creating their own cards as easier (mean of 1.7) than finding and reusing preexisting cards (mean of 2.7). No participant ranked the difficulty of either task above 3, indicating that participants considered both tasks to not be too difficult.

The participants gave us the impression that they generally enjoyed using the tool. When asked what they liked best about the tool, we received positive comments about the emphasis on the visuals, the drag and drop interaction, and the different spaces organizing the work for them. However, they did offer some suggestions for improvements. While viewing images initially helped them find the artifacts they wanted to use, they felt that an arranged positioning of cards (instead of the current scattered placement) would have aided them in finding cards more easily. Others request larger labels for the cards. When creating cards most participants wrote the point and click approach to entering the parts of the card made the task simple, but some participants encountered problems in finding out how to modify their cards.

5 Conclusions

Our study on storyboarding demonstrated we can create scenarios of use by reusing artifacts. We found the use of imagery aids ideation, quick recognition, and access to claims that might have otherwise been harder to identify—creating a balance between creativity and design rationale. Key to the building of the storyboard are the relationships that naturally emerge and aid in identifying, deciding on, and sequencing artifacts. Grounded in these observations, we developed a tool to support storyboarding, but also facilitate the artifact creation and sharing needed for reuse. Our research efforts have lead us to learn some important lessons regarding storyboarding and reuse.

Encapsulate reuse within a prototyping activity. Providing reusable artifacts and expecting designers to reuse them is not enough. There is a need to support a design activity that is widely accepted by practitioners. Storyboarding itself is a highly creative and fun activity early in design that can be beneficial to the designers trying to express their ideas and to other stakeholders who wish to understand how a system will function. Since storyboarding naturally has a componentized nature due to the panel structure, it fits well as the basis for a flexible reuse-based design process.

Encourage learning while making contributions. Encapsulating reuse within larger prototyping activities means that contributions are not just being made for the benefit of others, but that designers have a personal reason for creating new artifacts. However, for their contributions to ultimately be useful to others they must be authored in a way that is not too restrictive. Exposure to other artifacts of a similar nature, either through mechanisms such as categories or relationships, can not only familiarize the designers with the domain, but also the ways in which they are created.

Weigh the tradeoffs of structural changes in artifacts. Since we knew of the challenges designers face when trying to reuse, one of our intentions was to explore the benefits of attaching inspiring imagery to claims. While they provide easier access to rationale, it also means that we must explore how to encourage people to also include their own imagery. Although further iterations on the storyboards and support for sketches might provide incentives to designers, this teaches us an important lesson in understanding the impact on reuse when trying to shift the balance of the challenges.

Our work provided us with valuable insight into the design concerns that impact the building of storyboards through reuse. Other aspects of the activity, such as the collaboration, can also impact what and how artifacts are reused. We realize there is still more work to be done. We intend to refine PIC-UP to fix identified problems. Additional support for assigning relationships can also be included so that new contributions are not left in isolation in the repository, but connected to the larger network of artifacts. Because our storyboarding is based on generic images as opposed to specific sketches, this places greater emphasis on the narrative when trying to understand the task flow. Further design iterations to specify the storyboard may be required by designers, but we believe this approach can aid novice designers in learning about components needed for their system. Thus, we plan on conducting a summative evaluation on the use of the tool with practitioners to better define how well the storyboards are able to capture the information most desired.

Acknowledgments. This work was partially funded by the National Science Foundation (IIS-0552732).

References

1. Bailey, B., Konstan, J.A.: Are Informal Tools Better?: Comparing DEMAIS, Pencil and Paper, and Authorware for Early Multimedia Design. In: Conference on Human Factors in Computing Systems (CHI), pp. 313–320 (2003)
2. Bias, R.G., Mayhew, D.J.: Cost-Justifying Usability. Morgan Kaufmann, San Francisco (2005)
3. Borchers, J.O.: A Pattern Approach to Interaction design. In: Conference on Designing Interactive Systems (DIS), pp. 369–378 (2000)
4. Buxton, B.: Sketching User Experiences: Getting the Design Right and the Right Design. Morgan Kaufmann, San Francisco (2007)
5. Carroll, J.M., Kellogg, W.A.: Artifact as theory-nexus: Hermeneutics meets theory-based design. In: Conference on Human Factors in Computing Systems (CHI), pp. 7–14 (1989)
6. Dusink, L., van Katwijk, J.: Reuse Dimensions. In: Symposium on Software Reusability, pp. 137–149 (1995)
7. Finch, C.: The Art of Walt Disney: From Mickey Mouse to the Magic Kingdom. Harry Abrams, New York (1973)
8. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns: elements of reusable object-oriented software. Addison-Wesley Longman, Boston (1995)
9. Hart, J.: The Art of the Storyboard: Storyboarding for Film, TV, and Animation. Focal Press (1998)
10. Horner, J., Atwood, M.E.: Design rationale: rationale and the barriers. In: Nordic Conference on Human-Computer Interaction (NordiCHI), pp. 341–350 (2006)
11. Hughes, M.: A pattern language for user assistance. Interactions 14(1), 27–29 (2007)
12. Landay, J.A., Borriello, G.: Design Patterns for Ubiquitous Computing. Computer 36(8), 93–95 (2003)
13. Landay, J.A., Myers, B.A.: Interactive sketching for the early stages of user interface design. In: Conference on Human Factors in Computing Systems (CHI), pp. 43–50 (1995)

14. McCrickard, D.S., Chewar, C.M., Somervell, J.P., Ndiwalana, A.: A Model for Notification Systems Evaluation–Assessing User Goals for Multitasking Activity. Transactions on Computer-Human Interaction (TOCHI) 10(4), 312–338 (2003)
15. Newman, M., Lin, J., Hong, J., Landay, J.A.: DENIM: an informal web site design tool inspired by observations of practice. Human-Computer Interaction 18(3), 259–324 (2003)
16. Payne, C., Allgood, C.F., Chewar, C.M., Holbrook, C., McCrickard, D.S.: Generalizing Interface Design Knowledge: Lessons Learned from Developing a Claims Library. In: International Conference on Information Reuse and Integration (IRI), pp. 362–369 (2003)
17. Rosson, M.B., Carroll, J.M.: Usability Engineering: Scenario-Based Development of Human Computer Interaction. Morgan Kaufmann Publishers, San Francisco (2002)
18. Sharp, H., Rogers, Y., Preece, J.: Interaction Design: Beyond Human-Computer Interaction. John Wiley & Sons Ltd., West Sussex (2007)
19. Strauss, A., Corbin, J.: Basics of qualitative research: Grounded theory procedures and techniques. Sage Publications, Inc., Newbury Park (1990)
20. Sutcliffe, A.G., Carroll, J.M.: Designing Claims for Reuse in Interactive Systems Design. International Journal of Human-Computer Studies 50(3), 213–241 (2000)
21. Truong, K.N., Hayes, G.R., Abowd, G.D.: Storyboarding: an empirical determination of best practices and effective guidelines. In: Conference on Designing Interactive Systems (DIS), pp. 12–21 (2006)
22. Wahid, S., Allgood, C.F., Chewar, C.M., McCrickard, D.S.: Entering the Heart of Design: Relationships for Tracing Claims Evolution. In: Conference on Software Engineering and Knowledge Engineering (SEKE), pp. 167–172 (2004)
23. Wania, C.: Examining the Impact of an Information Retrieval Pattern Language on the Design of Information Retrieval Interfaces. Ph.D. Dissertation, Drexel University (2008)
24. Whittaker, S., Terveen, L., Nardi, B.A.: Let's stop pushing the envelope and start addressing it: A reference task agenda for HCI. Human-Computer Interaction 15, 75–106 (2000)
25. Yahoo! Patterns Library, `http://developer.yahoo.com/ypattern`

Are User Interface Pattern Languages Usable? A Report from the Trenches

Regina Bernhaupt, Marco Winckler, and Florence Pontico

Institute of Research in Informatics of Toulouse (IRIT), University Paul Sabatier
118, route de Narbonne, 31062 Toulouse cedex 9, France
{bernhaupt,winckler,pontico}@irit.fr

Abstract. Patterns languages pattern languages for interaction design have been an active research field in the area of Human-Computer Interaction. However, only few researchers have explored the impact of pattern languages in real application domains. This work reports a field study on acceptance and adoption of pattern languages by development teams in industry. Our results show that pattern languages should take into account the idiosyncrasies of the application domain and results show that pattern languages can be a great leverage to improve usability culture in industry.

Keywords: pattern languages, field usability study, acceptance.

1 Introduction

Design patterns were first introduced in the field of Architecture as a way to describe design solutions in an understandable and generalized way. In recent years design patterns have received considerable attention in the field of Software Engineering and in Human-Computer Interaction (HCI) [1] for their potential for recording and communicating design options and for supporting knowledge reuse during the design process. Contrary to other strategies such as guidelines and heuristics, design patterns focus on concrete solutions of problems rather than on general recommendations. A design pattern can be defined as 'an *invariant solution* to a *recurrent problem* within a *context*'. Despite the fact that patterns can be described single standing, they quite often appear in a set of interlinked patterns for a specific domain. Such a collection of patterns is called a pattern language. Although many patterns and pattern languages have been published by the software engineering and the HCI community, little has been reported about their actual use in industry [3]. In order to understand how design patterns are actually used by development teams in industry we realized an empirical evaluation of the usage of a pattern language. This study was carried at SmalS-MvM[1], a non-profit organization employing more than 1.000 persons (mainly administrative and information technology experts) devoted to the design, development, deployment and handling of public e-Government applications in Belgium. Based on the very general nature of the web-service development in that area, we assume that lessons from this study can be generalized to other web-oriented domains.

[1] http://www.smals.be/

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 542–545, 2009.

2 The Usability Study: Participants and Procedure

SmalS-MvM started developing a pattern language in 2006 as a means to support the discussion of solutions with stakeholders (many and with different background) and to finally introduce a standardization of the user interface for all applications (which are produced by different teams in the organization). The User Interface (UI) patterns have been identified by browsing existing applications designed by SmalS-MvM. They were organized in three different levels of granularity covering: screen flow (e.g. sequence of pages), web page (e.g. layout), and basic components (e.g. form fields).

The description of the UI patterns is rather classical (advice on implementation and rationale around a given UI design problem) but it also includes wireframes for supporting low-level fidelity UI prototyping using design patterns. According to the level of granularity, wireframes feature a schematic representation of the layout and the disposition of an UI element (e.g. a page or a form). Navigation diagrams are depicted using a formal description technique named StateWebCharts (SWC). Individual patterns also offers pointers to user interfaces components (e.g. MS Visio templates, style sheets) supporting the implementation phase. Design patterns are available in the form of an intranet web site.

To investigate the usage of this pattern language in a real application domain, a pre-questionnaire was used to recruit participants among 34 employees that are (expected to be) working with the pattern language. Only eight participants agreed to take part in the expert interview (6 M, 2 F), aged 24 to 31. Five participants indicated to work mainly as analysts; two indicated to be mainly leading projects, and one was mainly working as a system architect. Three participants were working more than three years at SmalS-MvM, two participants up to two years and three participants were at SmalS-MvM for 3, 6 and 9 months respectively. The pattern language was used in projects by six participants; two participants did not use the pattern language within their projects. One participant reported to never have looked at the pattern language before, six mentioned looking up solutions for specific problems, one added comments to improve the pattern language, four used personalized wireframes and five used the pattern language to argue for solutions in their projects.

The study consisted of three tasks participants conducted at their workplace, followed by a final interview about possible improvements for the pattern language. Users were given the following problem: "*Your goal is to set-up a webpage for the lunch distribution of a school with a personal identification for parents and the possibility to select the weekday lunch taken by the kids. Parents should have the possibility to cancel lunch three days in advance and the school staff to look up the number of ordered lunches*". Task 1 was to find a general solution for the lunch web page using the pattern language by selecting a wireframe page (T1/a) and producing an html page incorporating the solution (T1/b). Task 2 was to compare the solution of the participant with the proposed pattern language solutions and describe how to modify the pattern language if necessary. Task 3 was to develop a solution for paying the lunch fee by credit card. The pattern language does not provide a complete solution to this problem so it should evoke the need for updating the pattern language.

3 Results

The evaluation showed various aspects on how to improve the usage of patterns in a user-centered development process. Due to space constraints, Table 1 summarizes the main findings from the expert interviews.

Table 1. Results for the three working hypothesis

Task	Investigate Topic				
T1/a: Designing with Wireframe	pattern language usage +	integration in MS Visio ++	user satisfaction +++	adequacy of pattern language +++	method take-up ++++
T1/b: Designing with SWC	knowledge of SWC +++	hypertext modeling +++	usage of SWC +	user satisfaction with SWC +++	
T2: Discussing alternatives	readability +++	usability culture +++	adaptability of patterns +++		
T3: Updating	(because) pattern is missing +	resources used -	agility ++	gratitude ++++	motivation +++

Legend: ++++ very frequently, +++ frequently, ++ sometimes, + rarely, - never

H1: The pattern language is usable and helps to find solutions within the conceptual phase. The usability of the pattern language is acceptable but leaves room for improvements. User reported to be satisfied with the pattern language in general but they rarely looked up the pattern language for performing task 1a and 1b. Some difficulties were observed especially for subjects that have little training with MS Visio and SWC; This finding suggests that the users background may affect the general understanding of the solutions provided by the pattern language.

H2: The wireframe and the hypertext structures are helpful to reach consensus for the user interface solutions in the project. The pattern language itself is well perceived in terms of usefulness, it is valued in terms of readability and that it can be adapted and modified if necessary. The general usability culture of the enterprise shows that the patterns are used to reach consensus in discussion and during project development.

H3: The pattern language helps to support the reuse of solutions made during the conception of user interfaces. The pattern language is helping to reuse solutions but tight project schedules limit their usage as a leverage to create a usability culture. Problematic in the use of the pattern language is the missing knowledge of what patterns/solutions are already available and what patterns are missing. The success of a pattern language in a real application domain will depend on the ability to keep the language a living document, as pointed out by participant 4: "*I feel concerned by the evolution of the pattern language, because if it does not evolve it will become unusable soon*". Time constraints stemming from the various projects may affect the evolution of the pattern language as stated by participant 6: "*I understand the importance of this issue [about the evolution of the pattern language], however project deadline is our priority*". From the perspective of a general usability culture, the pattern language is accepted by the analysts.

4 Discussion and Conclusion

The contributions of this work are two-fold: based on this qualitative study we can compare theoretical claims on the use of patterns and their actual use in industry; in addition, we report many of the difficulties one can encounter when evaluating pattern languages which stem from the absence of sound quality models for describing pattern languages. In fact, despite abundant literature on HCI pattern languages, very few patterns authors' clearly state rules for ensuring a certain level of quality for their pattern languages [6]. Few reports of empirical validation of pattern languages exists, however they are limited by artificial testing conditions with users that have never experienced pattern language in their real projects [2]. Moreover, quite often the participants are students and/or researchers which are not representative for a real user group from industry [4]. An exception for this rule is the work of Malone et al. of the Yahoo! pattern language validation by an open (web) community [5].

The originality of the present study is to assess the use of design patterns in an organization that has formally adopted pattern languages as a solution for improving its usability culture. As we learned, the formal adoption by the company does not mean that employees easily integrate the pattern language in their daily work.

As we have seen at SmalS-MvM, pattern languages are a good solution to leverage the usability culture into organizations but, as living documents, pattern languages should evolve to meet organization needs (in particular stakeholders' needs for their projects). Otherwise pattern languages do have limited credibility and lose their added value to support the communication and to generate ready-made solutions. Due to several constraints, this study focused on only one category of stakeholders (i.e. the analysts) and their activity in a specific design phase of the development process (i.e. design of prototypes). We suggest that further investigation should be performed to understand if the pattern language available fulfills also other stakeholders' needs. Moreover, it would be interesting to observe different stakeholders at different stage of the development process.

Acknowledgments. Our gratitude goes to Quentin Limbourg, SmalS-MvM's employees who participated in this study and to the Action COST294-MAUSE (http://www.cost294.org/).

References

1. Borchers, J.: A Pattern Approach to Interaction Design, p. 268. John Wiley and Sons, Chichester (2001)
2. Chung, E.S., Hong, J.I., Lin, J., Prabaker, M.K., Landay, J.A., Liu, A.: Development and Evaluating the Emerging Design Patterns for Ubiquitous Computing. In: Proc. of ACM DIS 2004, Cambridge, USA, August 1-4, 2004, pp. 233–241 (2004)
3. Dearden, A., Finlay, J.: Pattern Languages in HCI: A Critical Review. Human-Computer Interaction 21(1), 49–102 (2006)
4. Golden, E., John, B.: L. Bass. The Value of a Usability-Supporting Architectural Pattern in Software Architecture Design: A Controlled Experiment. In: Proc. of ICSE 2005 (2005)
5. Malone, E., Leacock, M., Wheeler, C.: Implementing a Pattern Library in the Real World: A Yahoo! Case Study (April 29, 2005), `http://leacock.com/patterns/`
6. Todd, E., Kemp, E., Phillips, C.: What makes a good User Interface pattern language? In: Proceedings of the 5th Australasian User Interface Conf. (AUIC 2004), Australian Computer Society, pp. 91–100. Australian Computer Society, Inc. (2004)

Get Your Requirements Straight: Storyboarding Revisited

Mieke Haesen, Kris Luyten, and Karin Coninx

Hasselt University - tUL - IBBT,
Expertise Centre for Digital Media,
Wetenschapspark 2, B-3590 Diepenbeek, Belgium
{mieke.haesen,kris.luyten,karin.coninx}@uhasselt.be

Abstract. Current user-centred software engineering (UCSE) approaches provide many techniques to combine know-how available in multidisciplinary teams. Although the involvement of various disciplines is beneficial for the user experience of the future application, the transition from a user needs analysis to a structured interaction analysis and UI design is not always straightforward. We propose storyboards, enriched by metadata, to specify functional and non-functional requirements. Accompanying tool support should facilitate the creation and use of storyboards. We used a meta-storyboard for the verification of storyboarding approaches.

1 Introduction

We expect the quality of a software product to improve on various aspects when people with complementary know-how are involved in user-centred software engineering (UCSE) approaches (e.g. software engineers, graphic designers, usability engineers, stakeholders, end users). However, ambiguity of (both functional and non-functional) requirements, because of different disciplines involved in the early stages of the development process, leads to undesired behavior of the final application [1].

An earlier literature study of user-centred design (UCD) techniques and tools showed that few tools actually include support for the cooperation of several disciplines [2]. It also revealed a lack of notations and tools for early stages of UCSE.

In order to verify whether our findings correspond with common practice, we interviewed companies that are active in UCD. These companies have between five and twenty years experience, employ five to forty professionals with various backgrounds, and accomplish projects in a wide range of application domains. Eight employees, mainly usability engineers and designers, participated in the interviews.

The interviewees often collaborate with external partners for the actual development of an application. With no exception, they indicated there is no sufficient support to translate their artefacts into a notation appropriate for software engineers or developers. Shortcomings of existing notations require close collaboration between all team members to avoid misconception.

Because of the independence of UCD relative to requirements engineering as known in software engineering [1], we investigate storyboarding for non-technical as

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 546–549, 2009.

well as technical team members. Two of the interviewed companies use storyboards, but only for specific applications, such as location-based systems or applications with a complex workflow. Nevertheless, they confirmed that storyboards are usable by all team members and contribute to later stages when suitable tool support should be available, which is confirmed in literature [3,4]. We propose a storyboarding tool for the creation and use of storyboards in multidisciplinary teams.

The storyboarding technique and our accompanying tool are described in the remainder of the paper. Furthermore, we propose the use of a meta-storyboard, depicting storyboarding in common practice, to evaluate and verify storyboarding in multidisciplinary teams.

2 Storyboards as a Common Language in UCSE

A suitable notation to support the early stages of UCSE needs to be (1) *understandable for all team members*, and have the possibility to (2) *present both functional and non-functional requirements*. Recent studies [3] show that the combination of stories and sketches is helpful to reason about a future application, errors, temporal information and contextual information. Consequently, scenarios, which are frequently used in UCSE and describe the future use of a system in narrative stories, become a much more powerful tool if visual information is added.

A storyboard consists of *scenes* that are related to each other and visualise sequences in a narrative scenario of use. This visualisation makes the narrative story more explicit and invokes empathy by depicting end users that interact with the future system [4]. Starting from a storyboard composed of scenes, we propose that each team member can add metadata according to their expertise. *Metadata* is similar to the free annotations designers usually add to storyboards [5], and concerns persona information resulting from a user needs analysis, device specifications or other annotations.

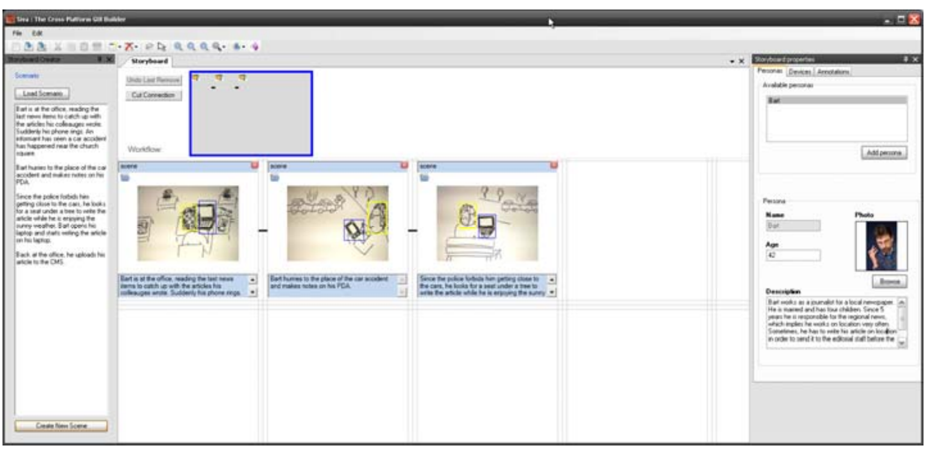


Fig. 1. A screenshot of our storyboarding tool, showing the narrative scenario (left panel), the storyboard with highlighted metadata (center) and the metadata details (right panel)

The combination of storyboards and metadata can contribute to later stages of the UCSE process: device information provides input and assistance for the UI design; persona information supports roles in the UI design and the evaluation of designs. As such, the storyboard is used as a common language for the requirements elaboration phase. Integrated in our research tool, that supports several types of artefacts that are created within a multidisciplinary team, storyboards enriched by metadata can increase the visibility and traceability of a UCSE project. Furthermore, understandability is maintained for all team members.

Our storyboarding tool (Fig. 1) features the possibility to load a narrative scenario. Storyboards can be composed by adding scenes that are aligned with a particular sequence of the scenario. For each scene, a sketch or photo depicting end users interacting with the future application can be loaded. Metadata can be highlighted in each scene, while the metadata information can be entered in a form that is related to this highlighted part of the scene. Storyboards enriched by metadata avoid ambiguities in the scenario and provide connections to later stages such as the UI design.

3 Storyboard Usage Exploration

To verify if the proposed storyboarding approaches fit in common practice, we created a storyboard describing the usage of storyboards in a multi-disciplinary team. This *meta-storyboard* is based on the aforementioned interviews with practitioners and describes how storyboards are used in all stages of UCSE. The accompanying personas personify the basic disciplines involved in UCSE projects.

In order to verify the concept of storyboarding in an entire UCSE process, we evaluated the meta-storyboard in an informal test consisting of two phases: an evaluation of the storyboarding tool and a walkthrough to validate the meta-storyboard. The latter additionally enabled us to verify the suitability of storyboards and the tool to discuss the use of a future application. Four female and three male subjects participated. Their backgrounds were very diverse, including interaction design, graphic design, computer science and media studies. The subjects' experience in multidisciplinary project teams ranged from a few months to more than five years.

After the subjects got a first impression of storyboarding and the transition to UI design using the tool, they participated in the meta-storyboard walkthrough. Each subject was asked to keep in mind a given persona description while discussing the meta-storyboard scene by scene using the tool. The discussions were recorded using a voice recorder and after the walkthrough, the subjects answered a questionnaire concerning the persona description and the use of storyboards.

The visual representation of the meta-storyboard stimulated discussion and the tool facilitated the understanding: clicking the metadata available in a scene enabled the exploration of personas and devices. After the walkthrough, all subjects declared it was easy to understand the approach depicted in the meta-storyboard and they all confirmed that the persona information contributed to this understanding.

All subjects agreed with the general approach presented by the storyboard. They accepted the use of storyboards in a multidisciplinary team. Most subjects prefer various disciplines to be involved in the creation of the storyboard to avoid misunderstandings at later stages. The technical subjects confirmed their work would benefit from storyboards because storyboards and provide more explicit information concerning activities carried out by users. Four subjects noticed that the tool should adapt its view and metadata according to the background of the team member working with the tool or the purpose of the storyboard at a particular stage.

4 Conclusion

The survey, we held among companies involved in UCD, inspired us to introduce storyboarding in multidisciplinary teams to specify functional and non-functional requirements. Storyboarding and tool support for storyboarding, were evaluated using a meta-storyboard that visualises the use of storyboards in a multidisciplinary team.

Our meta-storyboard was a valuable artefact for discussion and proved to be very understandable for team members with different backgrounds. The metadata, highlighted in the tool, contributed to the understanding of the meta-storyboard. The meta-storyboard walkthrough revealed that the tool should provide various visualizations according to the team members' backgrounds and the purpose of the storyboard. These and other results will be incorporated in the storyboarding tool. In the future we will frequently consult practitioners to investigate how metadata of storyboards contributes to UCSE projects and to adapt the storyboarding tool to common practice.

Acknowledgments. Part of the research at EDM is funded by the ERDF (European Regional Development Fund) and the Flemish Government. The research described in this paper is based on our experiences in IWT project AMASS++ (IWT 060051).

References

[1] Lindgaard, G., Dillon, R., Trbovich, P., White, R., Fernandes, G., Lundahl, S., Pinnamaneni, A.: User needs analysis and requirements engineering: Theory and practice. Interact. Comput. 18(1), 47–70 (2006)

[2] Haesen, M., Coninx, K., Van den Bergh, J., Luyten, K.: MuiCSer: A Process Framework for Multi-Disciplinary User-Centered Software Engineering processes. In: Proc. of Human-Centred Software Engineering, September 2008, pp. 150–165 (2008)

[3] Brown, J., Lindgaard, G., Biddle, R.: Stories, Sketches, and Lists: Developers and Interaction Designers Interacting Through Artefacts. In: Proc. of Agile 2008, pp. 39–50 (2008)

[4] Truong, K.N., Hayes, G.R., Abowd, G.D.: Storyboarding: an empirical determination of best practices and effective guidelines. In: DIS 2006: Proc. of the 6th conference on Designing Interactive systems, pp. 12–21. ACM Press, New York (2006)

[5] Myers, B.A., Park, S.Y., Nakano, Y., Mueller, G., Ko, A.: How designers design and program interactive behaviors. In: VL/HCC, pp. 177–184 (2008)

Hello World! – Experiencing Usability Methods without Usability Expertise

Elina Eriksson, Åsa Cajander, and Jan Gulliksen

Uppsala University, Dept. of IT/HCI, PO Box 337, SE-75105 Uppsala, Sweden
{Elina.Eriksson,Asa.Cajander,Jan.Gulliksen}@it.uu.se

Abstract. How do you do usability work when no usability expertise is available? What happens in an organization when system developers, with no previous HCI knowledge, after a 3-day course, start applying usability methods, and particularly field studies? In order to answer these questions qualitative data were gathered through participatory observations, a feedback survey, field study documentation and interviews from 47 system developers from a public authority. Our results suggest that field studies enhance the developer's understanding of the user perspective, and provide a more holistic overview of the use situation, but that some developers were unable to interpret their observations and see solutions to the users' problems. The field study method was very much appreciated and has now become standard operating procedure within the organization. However, although field studies may be useful, it does not replace the need for usability professionals, as their knowledge is essential for more complex observations, analysis and for keeping the focus on usability.

Keywords: Field studies, ethnography, usability, user centered systems design, case study, public authority, sense making.

1 Introduction

Despite major research efforts within HCI and extensive practical adoption of HCI-methods, poor usability is still a major problem in applications resulting from in-house systems development, and causes frustration and stress in computer supported work [1]. There is a great need of usability activities, but fully incorporating usability aspects in the system development processes or in the organization is still difficult [2, 3]. HCI research stresses the importance of addressing usability early in the development process. However, design decisions that contribute to the usability of the end result happen throughout the development process. System developers[1] have great impact on the end-result through their needs to make design decisions throughout the system development process. Previous research on system developers in a public authority [4] examined how design decisions affecting usability were made, and on what grounds, showing that system developers regard their work as problem solving. The result of their work was judged by technical quality attributes rather than its

[1] In this paper system developer refers to a person mainly programming systems, although the work of the system developer may include modeling or design as well.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 550–565, 2009.

contribution to the end-user's work situation. The design of the user interface simply emerged without anybody consciously making the design decisions leading to the results. The results of these late design decisions often affect the usability of the system in a negative way, as they are not informed by the end-users' situation. However, system developers are not evil or ignorant in making these decisions. Rather, they do not know enough about the work of the end user. To get a better understanding of users and system usage, system developers need to apply usability methods. One such method is field studies.

2 Purpose and Justification

Our overall research objective is to make usability a key concern in practical systems development, and to impact the development processes to address usability issues. Field studies are widely accepted within Human Computer Interaction (HCI), and the general view is that the quality of the observations and documentation depend on the experience and sense making of the person conducting the field study. Experienced usability experts are thus likely to produce richer data and make observations of a different nature when conducting field studies, compared to novice users of the method. The practical reality in many organizations is that usability expertise may not be available, and if that is the case we need to find alternatives. Successful introduction of usability in organizations requires us to change attitudes and values of people working with systems development. One alternative might be to educate people with no previous usability expertise in usability methods hence impacting their sense making through experiences and understandings of the use situation. If usability methods are not regarded as useful by the developers, its prospect for successful deployment may be severely undermined [5]. Hence, this study examines how developers, with no previous knowledge of HCI, experience field studies in practice and discusses the possible implications it might have on system development. It does not focus on the documentation made in field studies, or the quality of observations, but on experiences and understanding gained by the system developers.

3 Theoretical Perspective

Examining the role and implication of usability methods in system development practice requires an understanding of human action and competence. Our research has been inspired by the perspective that human action is situated [6]. Consequently, our research must take place in practice. It consists of both "high, hard ground where practitioners can make effective use of research-based theory and technique" as well as "swampy lowland where situations are confusing 'messes' incapable of technical solution" [7]. Our research is based on a constructivist and interpretive perspective, where we create and understand our reality by using language through communication. Interpretations are flexible, situated, and socially constructed. We adhere to the quality criteria and principles established by Klein and Myers [8] as well as Rasmussen [9]. Such research based on case studies leads to contextual in-depth knowledge, which is hard to generalize. We as researchers, the context, the organization and the conditions, under which the research takes place, color the results. However, even though the organizations and the findings are unique they are not uncommon and therefore we believe that the reader may find the knowledge gained applicable in other settings. Our perspective also partly originates from Participatory Design (PD) [10-12] which stresses the importance of involving users in the design process, arguing their right to be involved in development of their future use situation. PD has evolved during the last twenty years and has distanced itself from its political heritage towards a pragmatic view focusing on the quality of the user experience [11, 13]. Moreover, our perspective adheres to user-centered systems design (UCSD) [14, 15], which is also an international standard [16]. UCSD provides methods, roles, processes and techniques addressing usability and users' needs in systems development in practice. These approaches emphasize among other things the necessity of involving users, addressing usability, and understanding users' needs and work practices. For instance, they suggest studying users in real work situations in order to understand the activity-oriented view of work [17] reflecting what people do in their work to meet organizational and individual goals. These theoretical views guide us as researchers as well as informing the practice we try to improve.

4 Field Studies

A well-known method of gaining knowledge about the users' work situation is ethnographically inspired methods, e.g. field studies. Several practical variants of field studies have been developed, such as contextual inquiry [18], analysis of information utilization [19] and the ADA method [20]. Even though field studies are valued within HCI, they do not seem to be widely used in practice since many believe them to be too time consuming and producing too much data. Hughes [21] presents different forms of ethnographic methods that inform the design process, of which "quick and dirty ethnography" aims at doing shorter, more focused studies. It was based on insights that understanding complex organizations is not possible through a traditional ethnographic study given limited time frames and that fieldwork will provide no further useful insights that could aid the requirements engineering process. Another method, rapid ethnography [22] is based on the notion that that field methods are too time consuming. The process is made more efficient by limiting the research focus and scope, using key informants, capturing rich field data by using multiple observers and interactive observation techniques, and collaborative qualitative data analysis [22]. Previous research on field studies in practice has focused on adapting and streamlining field studies to the limited time scale of systems development. One example is Kujala's strategy labeled the "Trojan Horse" [23] where field study findings are presented in a familiar form, easy to implement in system development.

There is an ongoing discussion within HCI about the role of ethnography and ethnographically inspired methods. Most of the criticism is directed against the HCI research field rather than how the methods are used in practice. E.g. Dourish [24] argue that the HCI field more or less see ethnography as a method that generates requirements for systems design, thus overlooking the wider value of ethnography. Anderson [25] claims ethnography to be misunderstood, and that unfamiliar users of ethnography tend to think that it is easy to report back findings without interpretation, thereby missing the analytical part of ethnography. This criticism can also be seen in the paper by Forsythe [26] where she presents 6 misconceptions about the use of ethnography in systems design. All in all these 6 misconceptions lead to the belief that ethnography is "common sense" that anyone can do, thereby miss the theoretical grounding and analytic work needed in order to produce reliable results. Ethnography leads to a rich in-depth material, and according to Plowman et al [27], it might be difficult for designers to make use of the results. All of the above-mentioned criticism is primarily focusing on the usage of ethnography or ethnography-inspired methods within research or by researchers to inform systems design. Bader and Nyce [28] aim their criticism mainly at practice and state that ethnography will not be a method that developers use frequently in systems development. Developers tend to "mistake themselves for their informants" thereby using data from ethnography to confirm their own beliefs. The authors' second reason is that developers' view of knowledge is rule bound and therefore the developers believe that work and social life can be explained with complex rules and principles. According to Nyce and Bader, developers will not value results stemming from ethnography or ethnographically inspired methods.

We argue that if field studies could be seen simply as a methodology to gather a scent of the use situation and the context of use, and not as a method to generate requirements or implications for design, we would be able to understand its value for developers in practice much better. We claim that it is crucial that developers broaden their perspective to make design decisions that may lead to better work environment and improved quality of work. In this sense, field studies can be a tool for developers to understand technology's impact on humans and work, and it might constitute a reflective tool that helps developers frame and reframe the problem space. One example of this would be to reconsider what system development is about, and perhaps see it more as designing future work, and less about creating perfect code.

5 The Case Setting

This research study was conducted in cooperation with the Swedish public authority that handles financial aid for students. The computer support used in the authority is mainly developed in-house, and a majority of these IT projects start because of changing legislations that influence the prevalent work practices. Consequently, time is a major factor controlling most IT development since the deadline of the implementation of a new legislation is fixed. IT projects have participants from different parts of the organization, as in a matrix organization, where each department has their role and responsibility. Officially, all projects are run according to methods provided by the business development department, and the IT architecture department. These include methods for acquisition, project management and systems development. In short, these methods provide a common framework, with descriptions of milestones, decision points, templates and role descriptions. Systems development is based on the waterfall model [29], and very few activities focus on usability. About two years prior to this study, a large 3-year action research project was launched within the organization with the purpose of increasing awareness and knowledge about usability and a healthy computerized work environment. The aim was to introduce and develop usability work practices based on our previous theoretical and empirical work [15, 30].

Focus is on increasing competence among all parties involved in developing computerized work through education and training, coaching and project cooperation.

6 Method

Our research is based on an action research perspective, see for example Rasmussen [9] and Avison et al [31]. As an action research project, it has two goals as described by McKay and Marshall [32], i.e. one research interest and one problem solving interest for the organization in which the research is performed. The interest in system developers and field studies is one good example where the organization and their concerns have guided the research, as they wanted to find ways to use usability methods despite the fact that there were few usability experts in the organization. We participated in project activities as members of the project group, at the same time as we observed certain aspects of the interaction and communication as researchers. In this study, data was generated during four months through participatory observations, a qualitative survey, interviews and the developers' own field study reports. The events and data generation methods are described in more detail below and summarized in Table 1.

Table 1. Data generation methods and events

Usability training course	Pilot project
• Participatory observation during course; field notes • Feedback survey after field study • Field study reports • Course evaluation • Interview with Brian, system developer	• Participatory observation in project meetings; field notes • Participatory observation debriefing meeting after field study; field notes • Interview with Scott, system developer
• Interview with John, system developer	

During fall 2006, one of us planned and implemented a usability training course for developers. The course was three days long with 47 participants. The first day contained basic knowledge on general usability, quality attributes and a practical introduction to field studies, based primarily on the ADA method [20]. The second day the developers planned conducted and documented field studies in pairs. A field study feedback survey was distributed to the 47 participants in the usability training course, of whom 36 responded. The aim of the survey was to get information about how developers experienced the field study. The feedback survey consisted of 10 questions ranging from general questions about background and previous experience to specific questions about the advantages and disadvantages experienced with field studies. The feedback survey contained multiple choices and open-ended questions in free text form. The third day the results obtained in the feedback survey was discussed and basic introduction to design workshops and evaluation methods were taught. Furthermore, we got access to the field study documentation that the developers had written, as well as the results from the course evaluation made by the authority. In addition to the researcher giving the course a second researcher participated, doing participatory observations, which resulted in field notes.

The field studies the developers in the usability training course performed were not part of any ongoing project, although some of the developers chose to look at systems they had been or were developing. To get another perspective of how developers experience field studies, we included results from a software development project into this study, which was a pilot for testing usability methods. In this paper this project is referred to as the pilot project. The project was followed from pre-study to project start between August and December 2006. None of the participants in the project had any previous HCI knowledge or experience, and therefore they participated in the usability training course described above. After the course they conducted field studies in the pilot project, and discussed the results at a debriefing meeting. One of us followed the pilot project and made participatory observations. This resulted in field notes. Moreover, we interviewed three developers in order to further understand the their experience of field studies. The developers were selected from different perspectives, one had done field studies in both the usability training course and the pilot project (John), one had only done a field study in the pilot project and had not participated in the usability training course since he was a consultant and not employed by the public authority (Scott) and one had only done a field study in the usability training course (Brian). The 45-minute interviews were semi-structured, used an interview guide and were audio recorded. The feedback survey was analyzed and summarized into written form with emergent themes highlighted. The summarized feedback survey was presented at the last course day, and the different themes were discussed with the course participants. These discussions were partly recorded, and partly written down in participatory observations. The different participatory observations all ended up in hand written field notes, and these field notes were read thoroughly and used as both inspirations for questions in the interviews as well as material for the article. The course evaluation and the field study reports written by the developers were read through and interesting parts were highlighted and categorized into themes. The interviews were listened through and subsequently transcribed, printed and cut in pieces, color marked and sorted into themes. The themes in the feedback survey, field study reports, course evaluation and interviews were compared and discussed between all three authors. When writing up the study all names were altered, and quotations used are not transcribed verbatim.

7 Results

The results mainly focus on how developers experience field studies. When defining their formal role in free text, the majority of the respondents (26 out of 36 that answered the survey) wrote system developer. However, 4 were system specialists, 3 technology development leaders, 1 web designer, 1 consultant and 1 process developer. The interviewees as well as the participants that conducted field studies in the pilot project were system developers.

7.1 Holistic Overview versus Concrete Problems

The interviewees were asked if anything from the field studies could be used in their work. The interviewees mention the importance of gaining a holistic overview rather than pointing at concrete problems. The holistic view can be divided in two aspects. First, the overview of the system development project, that is, knowing the common goal of the project, what is happening and take part in all the stages of the project. Scott gives an example: *"The best thing is when you've been there so you know what it is all about /…/ so that you know what we are supposed to do, where we're heading."* Second the overview gives knowledge about core business. The developers have a wish to understand the context and work of case handlers[2] in order to create good systems. For example, when asked about what he needs in order to do a good job, Brian first states that he needs good programming tools and then core business knowledge. When prompted for clarification of what he means by core business knowledge he says: *"I need to understand where I am, what I'm doing and why. Yeah that's it, to understand the whole flow, to understand the process when people work"*

A majority of the answers in the feedback survey concern different aspects of a holistic view. They report insight into the system functionality and its interfaces with other systems, insights into what the user's work actually is about. Other insights relate to the users' work situation, e.g. that users felt trapped and couldn't receive essential help from anywhere and that the users don't see the possibility of influencing the system design. Moreover, some respondents where surprised at how stressful the case handlers described their work with the computer systems, and the extent to which work was manual, dealing with a lot of printed material. One respondent was surprised at the extent to which users used informal information, e.g. "reading between the lines" and the degrees of freedom in how people made use of the computer system. Interestingly, the respondents also noted several things concerning the user's work process, for example difficulties built into the work due to a mismatch between business processes and the needs of the work tasks. The work process was also discussed during the debriefing meeting after the field studies in the pilot project. Participants had seen problems that were hard to classify as work process problems or non-supportive IT-product.

In the pilot project participants found more concrete problems than problems related to holistic view. At the debriefing meeting, large amounts of problems were discussed ranging from a need of a weekly calendar for case handlers (someone had actually glued a calendar onto the screen of the computer) to the intense use of the mouse (since the instructions for shortcut keys where hard to find). During the interview, John started to recount different details that he would like to change, certain aspects of the graphical user interfaces that forced the user to do a lot of work with the mouse etc, which is also something he had written in his field study report. When prompted if he could change these things in the current project, he said no, it was not planned, and there was no time for this. All in all, the participants in the pilot project

[2] Case handler refers to a civil servant working at a public authority with case handling.

were overwhelmed with all the data they had collected, and wondered who should receive a description of the needs for identified improvements. Incorporating the problems in the pilot project was impossible due to time constrictions.

The feedback survey shows that some developers gained knowledge about the more detailed uses of the system. As opposed to the pilot project participants, they were surprised at the extent to which shortcut keys were used, although it should be noted that they were all observing use of different systems. Another finding was the practical difficulties making use of scanned information and documents and the effect this had on the case handlers' work. Three respondents also noted the extent to which text templates needed to be reworked for every situation.

7.2 Seeing without Gaining an Insight

Some developers report explicitly that they did not see anything of value during the field studies. However, when looking through the material, several things could be discerned that trained usability practitioners would notice and bring up. From a sense making point of view this reveals a discrepancy between what developers see, and what they find relevant. There is more to it than simply see and record as they must also gain insight. Brian, for example, was negative to involving end users in systems development and doing field studies. He expressed that it was interesting doing the field studies to see if the case handler worked the way the development team had thought they should work. But he concluded that the field studies didn't give him anything that he could use in his work as a developer.

> *"Researcher: But did the field study give you something*
> *you could use when you are developing?*
> *Brian: No*
> *R: And you saw nothing you could improve or something like that?*
> *B: No*
> *R: Can you imagine another situation where you could have*
> *gained something from a field study?*
> *B: No"*

In the light of this, Brian's co-authored field study report is interesting. It states two concrete problems besides those a trained usability practitioner could find. First, the case handler did not have time to finish the case she was working with, whilst having the customer on the phone. She recorded data on post-its to finish later, which she found easier than registering everything in the system. However, even writing information on post-its seemed difficult to manage during the call. Second, there was not enough information in one of the windows in use, and the developers gave a brief solution to the problem. Later in the same interview, Brian explained that he didn't think his field study was beneficial because the system, (that he had been developing) had been in use for a long time. He would have preferred to look at the use of the system after 3-4 months. Furthermore few reports contain clear statements on what is wrong in the computer supported work. It is as if they have not really reflected on what is wrong, but on the other hand it was not made explicit in the usability training course that they should come up with any solutions in the field study report. Moreover, when problems are found, only two of the reports suggest a solution and in some cases the only problems that are mentioned are the ones that the case handler has pointed out.

7.3 Responsibility

Developers were prompted to discuss what to do if having to make a design decision that might affect the user when programming. Two of the interviewees, Scott and Brian, said they talk to the other project group members. Usually the question ended up at the project management level. John said he had not thought about whether his decision would affect the end-user in the end. But, when thinking about it, he would like to talk to the user in that situation, since they were the ones that would use the system in the end: *"We have done that before. 'You can decide since you're the one who will work with the system in the future.' Then you put the responsibility on them somehow"* Brian was less eager to give responsibility to the users. In his opinion, it was useless to talk to real end-users, since they are not able to express what they want and they do not know the possibilities available. *"We are as good as they are at guessing! And I don't buy all these things about working with the users because they know how things work. Because I have been talking to users and I know the rules better than they do since they are used to doing things in a certain way. I who don't work with these things say that I have read the rules and these are the rules. Then I can do it easier somehow. I wouldn't say that I am better than they are at saying what they want. But still somehow we guess at what they need and they are not better than we are at expressing that need. Of course there are some things that they can tell us, but I don't think it is a good idea."*

7.4 Documentation and Templates

Developers generally did not want an interview template as it might disturb the interview preparation, and risk including too general questions, hence being less usable for the developers. However, a quick interview guide might come at hand. When discussing the documentation, the developers favored hand-written notes instead of a document template. However, in their opinion, a template facilitates incorporation into the development process and makes the documents easier to read. The interviews also show that written program specifications are less important when having a holistic view. Scott: *"If you know where we're heading, broadly so to speak, yes that is good starting point. That's what's really important. And the specification, what's written, is less important then."*

7.5 Feedback and Motivation

In the field study survey feedback, developers write about feedback and contact with real end-users. Some could see future benefits from doing field studies, such as: *"Good to establish some real contact with end-users that we could make use of in the future"*. Furthermore, one of the respondents found it especially motivating to *"see things that I can actually change in the system construction to improve the user's work situation"*. In the interviews the developers experienced that the field studies had given them feedback and motivation in their work. They wanted more feedback from users in their work, especially critical feedback. John mentioned that there is usually no user participation early on in the development process, and the case handlers involved in the test phase in the end of the process, do as they are told, and do not give enough feedback. And then he concluded that it was too late anyway to get feedback from the case handlers doing tests, since there were no chances of modifying the system after the test phase. He had done so once and been criticized by the development project. In the interview Brian expressed dissatisfaction with the feedback from the users. He got too little feedback and wanted more. According to him, users do not complain enough, or the criticism does not reach developers. The developers involved in the pilot project were also truly motivated by the field studies. During the debriefing meeting all the participants talked animatedly about their field studies and seemed to have enjoyed it, one exclaimed when it was his turn to talk that: *"Directly after I thought: Damn, I should visit more case handlers!"*

8 Discussion

8.1 Holistic Overview

An appropriate understanding of the context of use is one of the foundations of a user-centered development process [16]. Our findings show that one of the major benefits of having developers do field studies is the increased knowledge and awareness of the context of use. This is also consistent with surveys of user-centered design in practice [33, 34]. The context of use is one part of a holistic overview, which was emphasized in all the data from the study. But the fact that developers talked about holistic overview, without noticing concrete problems to the same high degree, might have other explanations. Developers work in projects where time is short and where they have no opportunity to deal with the encountered problems. To make sense of the method they might react to these cues; lack of time and no possibility to change things, and when asked about the use of the method, the most plausible answer they have is that the benefit of the method is the holistic overview [35]. Other aspects that might affect what the developers gained from the field studies are their lack of experience and almost no HCI knowledge. The field studies were conducted during a day, or part of a day, and in some cases it was a system that they had not built themselves. Facing a complex work situation without experience, developers are overwhelmed with information, which lead them to notice the holistic overview before details. One could argue that we influenced the developers in their view of the benefits of the method. However, the course was filled with tips and discussions about the details, as well as the holistic overview of field studies to address this bias.

8.2 Developers' Background and Knowledge

When developers, often trained in a positivistic tradition, learn about field studies as a usability method they often position themselves as neutral observers, expressing a concern that they will inevitably distort the data and introduce subjectivity as their presence will affect the work situation. This is consistent with others observations, as *"many people confuse observations with inference"* [36]. As in this example from the survey: *"We need better tools for doing this, in order for us not to lead the users in a specific direction."* Qualitative methods, such as field studies, are based on a different perspective, where the subjective and interpretive elements are foundations of human understanding. The developers' understanding of the setting and context depends on his/her perspective of the world which is, shaped by background, social position, gender, etc. In relation to the results, we would like to elaborate around three different kinds of knowledge that affect what the developers see or not see during the field studies, HCI or ethnographical knowledge, knowledge about the core processes in the organization and knowledge about the IT-systems. It is quite clear that the developers lack HCI or ethnographical knowledge. One easily perceived example of how differently a developer interprets the situation from a usability expert is Brian, who believes that users become "ruined" and constitute bad respondents when they have worked with a computer system for some years. These results are consistent with that of [36] as the observer "fail to look for information that may contradict or challenge assumptions". What is also apparent in the results is that the developers did not see beyond the user and the context as place and situation. None of the developers elaborated around culture etc. of which those with ethnographical or anthropological background would argue is the central outcome from a method like field studies [25, 26]. Nor did they focus on "implications for design" [24], that is concrete things to be considered during the development. Instead they ended up praising the holistic overview that they gained from the studies. Knowledge about the core business is apparently something that the developers seek, and the field studies are one way of gaining this knowledge. However, some of the developers have core business knowledge and has worked a long time in the organization. Other developers noticed that the case handlers did not work in the way they had predicted, however, they could not say if it depended on the work process or a faulty IT-system. Finally, knowledge about the IT-system, and the underlying technology, makes the developers see small details during the field studies that can easily be changed.

8.3 Field Studies Helping Knowledge Transfer

Several of the developers expressed that they would bring the experience from the field study into development projects in the future. However, in development projects where time and resources are scarce, not all systems developers will get the chance to visit case handlers in every project. Furthermore, usability experts will still make field studies and report back their findings to the development group. As in many knowledge intensive organizations, there is a need to codify knowledge, and transfer the knowledge to those who has not been present when the knowledge was created [37]. Conducting field studies, albeit not in all projects, will give the developers a framework, which will help them to understand the information and knowledge conveyed to them by others. Or as Davenport and Prusak states it [37]: *"Knowledge is more likely to be absorbed if it adheres to the listeners' sense of ground truth, … and is placed in a context of frame that is at least partly shared by its audience"* .

8.4 Motivation

Field studies have the potential to motivate developers, and reduce their motivation. It motivates developers in that they understand that what they do affect others, and that they see the practical consequences of their work. We were intrigued by the joy the developers expressed over the field studies, and the enthusiasm most of them showed while talking about the field studies. Several expressed wishes to go back and do something about the problems that they encountered, because they know quick ways of dealing with the problems. But unfortunately the current project structure does not allow them to deal with problems they encounter, as they only may address problems that they have allocated time for. If they want to make changes based on their observations they must propose this as new system development projects that need to go through the annual work planning and prioritization process. This is frustrating and therefore the field studies might reduce their job motivation. The interview data indicates a wish for overview and early participation as developers wish to have a stimulating work. Simply being "coders" is not as stimulating as being a part of the whole process. Participation in different stages of the process gives some variation to their work. Moreover Brian states that: *"The requirement specification should be more on an overview level and not as detailed and controlling, because if that would be the case, whoever had done the requirement specification could do the coding as well, translating from a sheet of paper to a program, which is something anybody could learn. No, I'd much rather see a requirement specification on a more overview level, and if you do understand the business processes, you are able to turn it into a great system"*. Clearly Brian wants to be more than just a "coder"; rather he would like to take part in the design of the system. Field studies done in the pre-study or early in the system development process would give him some of the information needed to understand the business processes and hereby design a usable system.

8.5 Practical Consequences and Possible Implications

Higher Focus. We believe that field studies can give usability a higher focus in the organization. However, this implies that the methods with deliverables need to be specified in the systems development process; otherwise the projects will not allocate time and resources to it. Field studies would help promote the user situation as a part of the project goals and could then potentially lead to increased focus on quality. Unfortunately we have seen that usability has come to be synonymous with field studies in the organization. There is a risk that the sustainability of these methods in the long run will be affected if the usability issues are not safeguarded throughout the project. Yet another practical consequence is that the developers through field studies meet with case handlers and establish a contact that they can use throughout the development. This will make it possible for developers to contact case handlers for small questions that even though small, can impact the end result.

Potential Risk. Can the fact that developers are trained to do field studies, design and evaluation, but have little HCI knowledge potentially be harmful? There are always potential risks with adopting new approaches like this, as the organization might conclude that since these usability methods are potentially useful for everybody, then we do not need any usability expertise. These findings are similar to [38], who concludes that there seems to be an ideal that the HCI practitioners are not needed at all since every member of the system development project is an HCI practitioner. However, as [38] discusses "If HCI work is everybody's then nobody's responsible".

Does increased knowledge about users lead to increased power for developers in the development process, as knowledge is power? Gillian [39] indicate that user involvement can be used as a political tool within organizations and a tool for developers to “manage their decision making process”. In our study, developers showed few tendencies of using field studies to promote their own ideas. On the other hand, future misuse of this power is still possible. Some developers were surprised at the multiple ways in which technology was used. Potentially this surprise, and the discussion afterwards, make them question their own technological deterministic perspective, hence realizing that users are not passive recipients of technology but actors [40] who create the circumstances, context and consequences of technology in use. Practice gives meaning to technology, and a good work environment is created in the situated experience of technology in context. We argue that a small insight into the world is better than none. However, Forsythe [26] criticizes this view: “…surely some knowledge of a situation is better than none. The problem is that in ethnography as in some other pursuits, a little knowledge can be a dangerous thing: superficial social research may confer the illusion of increased understanding when in fact no such understanding has been achieved”. We argue that developers are not involved in research, and furthermore they are not trying to do ethnography.

Usability Expert. We do not discard the need for a usability expert, who should have the knowledge and experience to challenge assumptions that developers might have. The usability expert’s knowledge is essential for making more complex observations and for keeping the focus on users and usability throughout the development project. We agree with Anderson, that: “This is not to say that getting to know users and their knowledge and practices is unnecessary or irrelevant or that observational fieldwork and impressionistic reportage can be of no value in this. Far from it! It is simply that you do not need ethnography to do that; just minimal competency in interactive skills, a willingness to spend time, and a fair amount of patience.” [25]

Impact on practice. Most of the developers who conducted field studies as a part of the course experienced it as useful, interesting and motivating. Our mission as action researchers is to make usability have a greater impact in organizations. One way of doing so is to involve usability expertise. But from previous studies we know that usability professionals may not have the impact in practice that one would have expected [2]. Or as Siegel and Dray put it: "UCD professionals who focus on doing "studies" as opposed to generating designs and products, will always be perceived as peripheral." [41]. We believe that the HCI research community partly is to blame for this. It may have been too focused on quantitative empirical studies, and scientifically validating the findings rather than making an impact on practice. Hence, HCI has produced a vast number of analysis techniques and usability evaluation methods, mainly to be applied by usability professionals. From a scientific point of view this is excellent, but does it really contribute to the development of practice? As HCI researchers we should be more concerned about making our methods, tools and knowledge used by practitioners who generate designs and products, and some of these practitioners are developers. As an interesting epilogue, we can report that inspired by the field studies conducted by the developers, the case handlers wanted to see how the developers worked. In the work exchange program organized by the union, case handlers visit headquarters and the developers' work setting, in order to get an understanding of their situated work and not only other case handling offices.

Acknowledgement. This work performed was sponsored by The Swedish Development Council for the Public Sector (Utvecklingsrådet). Moreover, we would like to thank the organization CSN (the Swedish public authority that handles financial aid for students), and all the participants in the study. We would like to direct our thanks to Paul Dourish and Inger Boivie for valuable comments on the manuscript, as well as anonymous reviewers who came with constructive critic on earlier drafts of this paper.

References

1. Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J., Shneiderman, B.: Determining Causes and Severity of End-User Frustration. International Journal of Human-Computer Interaction 17, 333–356 (2004)
2. Boivie, I.: A Fine Balance: Addressing Usability and Users Needs in the Development of IT Systems for the Workplace. Acta Universitatis Upsaliensis, Uppsala 85 (2005)
3. Rajanen, M., Iivari, N.: Usability Cost-Benefit Analysis: How Usability Became a Curse Word? In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4663, pp. 511–524. Springer, Heidelberg (2007)
4. Boivie, I., Gulliksen, J., Göransson, B.: It's all in a Days Work of a Software Engineer HCI International (2003)
5. Riemenschneider, C.K., Hardgrave, B.C., Davis, F.D.: Explaining software developer acceptance of methodologies: a comparison of five theoretical models. IEEE Transactions on Software Engineering 28, 1135–1145 (2002)
6. Suchman, L.: Plans and Situated Actions: The Problem of Human-Machine Communication. Cambridge University Press, Cambridge (1987)
7. Schön, D.: The Reflective Practitioner - How Professionals Think in Action. Ashgate Publishing, Aldershot (1983)
8. Klein, H., Myers, M.: A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. MIS Quarterly 23, 67–94 (1999)
9. Rasmussen, L.B.: Action research—Scandinavian experiences. AI & Society 18, 21–43 (2004)
10. Greenbaum, J., Kyng, M.: Design at work: cooperative design of computer systems. Lawrence Erlbaum Associates, Inc, Mahwah (1992)
11. Asaro, P.M.: Transforming society by transforming technology: the science and politics of participatory design. Accounting, Management and Information Technologies 10, 257–290 (2000)
12. Bødker, S., Iversen, O.S.: Staging a professional participatory design practice: moving PD beyond the initial fascination of user involvement. In: Second Nordic conference on Human-computer interaction, pp. 11–18. ACM Press, New York (2002)
13. Bødker, S., Ehn, P., Sjögren, D., Sundblad, Y.: Co-operative Design - perspectives on 20 years with The Scandinavian IT Design Model. In: NordiCHI (2000)
14. Norman, D.A., Draper, S.W.: User Centered System Design; New Perspectives on Human-Computer Interaction. Lawrence Erlbaum Associates, Inc., Mahwah (1986)

15. Gulliksen, J., Goransson, B., Boivie, I., Blomkvist, S., Persson, J., Cajander, A.: Key principles for user-centred systems design. Behaviour and Information Technology - BIT 22, 397–409 (2003)
16. ISO 13407: Human-centred design processes for interactive systems. International Organisation for Standardization, Geneva (1999)
17. Sachs, P.: Transforming work: collaboration, learning, and design. Communications of the ACM 38, 36–44 (1995)
18. Beyer, H., Holtzblatt, K.: Contextual Design: Defining Customer-centered Systems. Morgan Kaufmann Publishers Inc., San Francisco (1998)
19. Gulliksen, J., Lif, M., Lind, M., Nygren, E., Sandblad, B.: Analysis of Information Utilization(AIU). International Journal of Human-Computer Interaction 9, 255–282 (1997)
20. Åborg, C., Sandblad, B., Gulliksen, J., Lif, M.: Integrating work environment considerations into usability evaluation methods—the ADA approach. Interacting with Computers 15, 453–471 (2003)
21. Hughes, J., King, V., Rodden, T., Andersen, H.: Moving out from the control room: ethnography in system design. In: 1994 ACM conference on Computer supported cooperative work, pp. 429–439 (1994)
22. Millen, D.R.: Rapid ethnography: time deepening strategies for HCI field research. In: Conference on Designing interactive systems: processes, practices, methods, and techniques, pp. 280–286. ACM Press, New York (2000)
23. Kujala, S.: User involvement: a review of the benefits and challenges. Behaviour & Information Technology 22, 1–16 (2003)
24. Dourish, P.: Implications for design. In: Proceedings of the SIGCHI conference on Human Factors in computing systems, pp. 541–550. ACM Press, New York (2006)
25. Anderson, R.J.: Representations and requirements: the value of ethnography in system design. Human-Computer Interaction 9, 151–182 (1994)
26. Forsythe, D.E.: "It's Just a Matter of Common Sense": Ethnography as Invisible Work. In: Computer Supported Cooperative Work (CSCW), vol. 8, pp. 127–145 (1999)
27. Plowman, L., Rogers, Y., Ramage, M.: What are workplace studies for?: The fourth conference on European Conference on Computer-Supported Cooperative Work, pp. 309–324 (1995)
28. Bader, G., Nyce, J.M.: When only the self is real: theory and practice in the development community. Journal of Computer Documentation 22, 5–10 (1998)
29. Royce, W.W.: Managing the development of large software systems: concepts and techniques. In: 9th international conference on Software Engineering, pp. 328–338 (1987)
30. Sandblad, B., Gulliksen, J., Aborg, C., Boivie, I., Persson, J., Goransson, B., Kavathatzopoulos, I., Blomkvist, S., Cajander, A.: Work environment and computer systems development. Behaviour and Information Technology - BIT 22, 375–387 (2003)
31. Avison, D., Lau, F., Myers, M., Nielsen, P.A.: Action Research. Communications of the ACM 42, 94–97 (1999)
32. McKay, J., Marshall, P.: The dual imperatives of action research. Information Technology and People 14, 46–59 (2001)
33. Vredenburg, K., Mao, J.Y., Smith, P.W., Carey, T.: A survey of user-centered design practice. In: SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves, pp. 471–478. ACM Press, New York (2002)
34. Gulliksen, J., Boivie, I., Persson, J., Hektor, A., Herulf, L.: Making a difference: a survey of the usability profession in Sweden. In: Third Nordic conference on Human-computer interaction, pp. 207–215. ACM Press, New York (2004)
35. Weick, K.E.: Sensemaking in Organizations. Sage, Thousand Oaks (1995)

36. Wixon, D.R., Ramey, J., Holtzblatt, K., Beyer, H., Hackos, J.A., Rosenbaum, S., Page, C., Laakso, S.A., Laakso, K.P.: Usability in practice: field methods evolution and revolution. In: SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves, pp. 880–884. ACM Press, New York (2002)
37. Davenport, T.H., Prusak, L.: Working knowledge. Harvard Business School Press, Boston (1998)
38. Iivari, N.: Understanding the work of an HCI practitioner. In: 4th Nordic conference on Human-computer interaction: changing roles, pp. 185–194. ACM Press, New York (2006)
39. Gillian, S.: The Work of IT System Developers in Context: Organizational Case Study. Human-Computer Interaction 13, 37–71 (1998)
40. Bannon, L.: From Human Factors to Human Actors: The Role of Psychology and Human-Computer Interaction Studies in System Design. In: Greenbaum, J., Kyng, M. (eds.) Design at Work: Cooperative Design of Computer Systems, pp. 25–44. Lawrence Erlbaum Associates, Hillsdate (1991)
41. Siegel, D., Dray, S.: Living on the edges: user-centered design and the dynamics of specialization in organizations. Interactions 10, 19–27 (2003)

Supporting Worth Mapping with Sentence Completion

Gilbert Cockton[1], Sari Kujala[2], Piia Nurkka[2], and Taneli Hölttä[3]

[1] University of Sunderland, Faculty of Applied Sciences, Sunderland, SR6 0DD, UK
[2] Tampere University of Technology, P.O. Box 589, FI-33101 Tampere, Finland
[3] Paf, Unioninkatu 24, FI-00130 Helsinki, Finland

Abstract. Expectations for design and evaluation approaches are set by the development practices within which they are used. Worth-Centred Development (WCD) seeks to both shape and fit such practices. We report a study that combined two WCD approaches. Sentence completion gathered credible quantitative data on user values, which were used to identify relevant values and aversions of two player groups for an on-line gambling site. These values provided human value elements for a complementary WCD approach of worth mapping. Initial worth maps were extended in three workshops, which focused on outcomes and user experiences that could be better addressed in the current product and associated marketing materials. We describe how worth maps were prepared for, and presented in, workshops, and how product owners and associated business roles evaluated the combination of WCD approaches. Based on our experiences, we offer practical advice on this combinination.

Keywords: User Values, Sentence Completion, Worth-Centred Development (WCD), Worth Mapping, Monetary Gaming.

1 Introduction: Holistic Worth-Centred Interaction Design

HCI needs to develop more *holistic design practices*, i.e., combinations of approaches that give equal attention to all aspects of design and evaluation and work well together. olistic practices balance attention to choices of means, ends, beneficiaries and evaluations [1], but more importantly create and maintain synergies between and across them. This focus on synergies, or links between choices of means, ends, beneficiaries and evaluations, results in *designing as connecting* [1], rather than *designing as crafting*. Design choices extend beyond choices of means (e.g., features) to choices of ends, stakeholders ('beneficiaries') and validations ('evaluations'). By covering the classes of design choice, and their inter-connections, holistic design practices are supported by a wider frame of reference for designing.

Figure 1 shows six connections between four classes of design choice. The three dotted connections associated with evaluations are not in scope for this paper. Of the remaining two connections, the lowest one between *means* and *ends* is the main focus of worth maps [1], which are networks of elements and associations that contain *means-end chains* that hopefully causally link design elements to human elements [1]. Solid connections from *beneficiaries* to both means and ends are the main focus for a *sentence completion* approach. These two WCD approaches interface via a *design's ends*, that is, the explicit purposes of a design. Sentence completion provides direct

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 566–581, 2009.

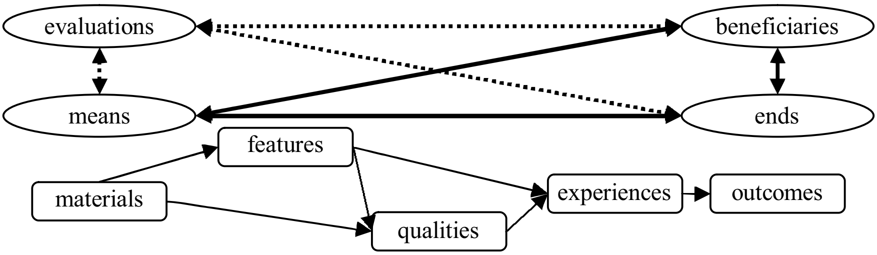


Fig. 1. Designing as Connecting, with means-end chain of worth map elements below

access to people's values in specific usage contexts, which values can be mapped onto outcome elements in worth maps. Sentence completion provides user data on benefits (a.k.a. value/s) that motivate users, and on the costs (perhaps adverse) that demotivate them. Once costs and benefits are identified, likely balances of worth can be assessed.

Sentence completion and worth mapping are called 'approaches', rather than 'methods' or 'techniques' because we cannot rigidly direct human activities (especially creative ones such as design). Instead, approaches combine flexible representations with suggested modes of use, theoretical and practical knowledge, perspectives, insights, and guiding values. Approaches leave project teams to find their own way through product or service development.

Worth maps have already been used in a range of explorations. They are network diagrams that use lines to indicate associations between human and design elements. These associations build vertically into means-end chains (an abstract one runs horizontally across the bottom of Figure 1, (see Figure 3 and [3] for example worth map structures). Worth maps explicitly connect design means with design ends, indicating 'why' people would (not) use a system. User Experience (UX) elements are the interactive means that link design ends (outcomes) with design means (qualities, features, materials). Good UXs satisfy motives.. Sets of worth maps can visualize a complete product context [2].

As with all approaches within a WCD framework, worth maps only support a subset of six design meta-principles [1]. Relationships between WCD approaches and meta-principles are important, since the latter scope the former. An approach needs to be evaluated on the bases of the meta-principles for which it provides support. Worth maps primarily support the *committedness* meta-principle through a single representation that can reference a project team's current set of choices of means, ends, beneficiaries, evaluations and their inter-connections. Worth maps also support *expressivit*, through visual representations of means-end chains. However, they do not support *receptiveness*, which requires openness to alternative design means, ends, evaluations, and a range of beneficiaries. Sentence completion can thus complement worth maps through receptiveness to design ends as user values. By providing data and analyses to identify (un)desirable design ends (outcomes), sentence completion also supports *credibility* by grounding design purpose (ends) in inspectable user data.

2 Research Instruments for User Values

Values are abstract constructs that cannot be directly observed. People can talk about them, or they can be inferred from observations of people, places and/or things. However Hoyer and MacInnis [4], note that people do not often think about their values and can not easily verbalise what is really important to them. This presents different research challenges to understanding users' activities, which has been the dominant focus in HCI, although this is not to say that such observational studies are straightforward and without challenge. Even so, observations may not be the most effective research instrument when trying to access user values.

Some research instruments directly access a pre-determined context-free set of values, but these questionnaires make assumptions about values (e.g., the Rokeach RVS [5] and Schwartz SVS [6] value surveys). However, people can *value* something without being able to associate motives with pre-identified named values. Feelings tell people what is (not) worthwhile, but without always exposing value drivers. Thus validated psychometric instruments may limit accessible user values, perhaps to either too narrow a range of values, too abstract a set of values, or both.

Projective research instruments are more open than closed value surveys: examples include word association tests, sentence completion, drawing, writing and story completion [7]. These elicit qualitative data that people cannot or will not verbalize via more direct instruments. Many were developed in therapeutic or child research settings with limited verbal fluency and/or difficult access to feelings. Projective instruments elicit indirect responses, either to ambiguous stimuli or talking about objects, situations, or other people's feelings, attitudes and opinions. Presented stimuli are very ambiguous; individuals' responses can reveal fundamental modes of thinking and behaving. Also, in talking about a third party or object, respondents may project covert feelings, which may be subsequently discussed [7]. Banister and Booth [8] experimented with different projective techniques to explore innovative methodologies for child-centric consumer research. In drawing tasks, they provided templates blank except for an outline. Given a tree outline, children were asked to decorate it as a tree of disgusts, tree of not very good taste, and tree of very good taste. This prompted insightful data on likes and dislikes without being intrusive, with children's experiences communicated through their drawings.

Projective techniques are now more common in HCI [10]. Cultural probes aim to stimulate participant reactions [9] through a probe pack of artefacts for use in directed activities, such as disposable cameras for taking photos, or a voice recorder for capturing dreams. Cultural probes took advantage of rapidly falling prices of recording devices and for custom printing, extending the range of respondent activities with projective instruments. Voida and Mynatt [11] modified cultural probes and combined them with RVS [5] to elicit values from two families. Their value probes directed families to complete several value manifestations, for example, a Family Album (how they portray themselves), a Day Planner (how they spend their time, actually then ideally), a Map (how they use their space), a Budget (how they spend their money), and a Scrapbook (what they surround themselves with). Four student design teams each completed a RVS based on returned probes (two teams were assigned to a family). The teams' RVS results were then compared to their family's, with good matches at the extremes (most/least important values). However, while the RVS did prompt students to re-examine probes for manifestations of initially overlooked values, probes also manifested values that could not be associated with any RVS value ([11] p. 2013). This supports the claim that existing validated closed value surveys may not be comprehensive enough to connect with the tacit 'value' that can be provided by technological innovation.

Projective research instruments can thus complement observation, interviews and closed format questionnaires. However, an instrument's worth is inherently situational, and thus we cannot establish the relative worth of one instrument over another. There are no absolute orderings, with one approach inherently superior to another. As worth is situational, a project may not be able to afford extensive observations and/or triangulation interviews; or a project sponsor may prefer quantitative methods over interpretative methods that make inferences from informants' accounts of actions. Thus, for example, cultural probes have high production values, requiring highly creative design and selection of artefacts for the probe pack to ensure high aesthetic standards. Not all project teams can bear such costs. In addition, interpretation of returned probes should be creative and open, which again may be unsuitable for project milieux where cultural probes would be misused through systematic quantitative analysis [10].

3 Using Sentence Completion to Elicit Values

Our main case study sought to develop a user research approach to provide quantitative data on user values that could be affordably scaled to large samples and could integrate well with worth maps. We thus explored projective *sentence completion* as a research instrument for credibly identifying common user values.

3.1 Background

In sentence completion, a person is asked to complete incomplete sentences with their first reaction, since they are in a written form. Sentence completion is well established in consumer psychology. For example, Hoyer and MacInnis [4, p. 60] asked cigarette smokers why they smoked. Most said they enjoyed it and believed that smoking in moderation was fine. However, when given incomplete sentences like "People who never smoke are _____", they filled in the blanks with words like 'happier' and 'wiser'. And, given sentences like "Teenagers who smoke are _______", respondents answered with words like 'crazy' and 'foolish'. Smokers were clearly more concerned about smoking than their explicit answers indicated [4].

3.2 Theoretical Support for Values Elicitation

Sentence completion is a suitable questionnaire format for identifying user values. Unlike closed questionnaire formats (e.g., [12]), sentence completion tasks do not prime respondents with direct questions about specific emotions or values and their product or service associations. Sentence completion makes no assumptions about fixed sets of human values (as do RVS [5] and SVS [6]), but can still be guided by various literatures on human values, which we surveyed to identify ten categories of

Table 1. Categories of Values for Receptiveness in Sentence Completion (based on [13])

Category of value(s)	Examples
Social Relatedness [5,6,12,14]	Esteem, status, power, control and dominance, achievement, conformity, equality, helpfulness, honesty and loyalty
Emotional and hedonistic [5, 12,14]	Aroused feelings or affective states, pleasure, fun, sensory enjoyment
Stimulation and epistemic [5,12,14]	Excitement, experienced curiosity, novelty and gained knowledge
Growth and self-actualization [5,14]	Independent thought and action: choosing, creating, exploring
Traditional [6,14]	Respect, commitment, and acceptance of customs and ideas that traditional culture/religion impose on the self
Safety [12,14]	Security, social order, health, comfort, free from fear
Universal values [12]	Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature

potential user value(s) from a broad range of theories of human motivation, social psychology and consumer behaviour.

Psychology and sociology have investigated the nature of values. For example, SVS [6] value categories have been shown to be valid in 21 countries. The relative importance of these categories constitutes an individual's system of value priorities. However, consumer behaviour and marketing identify different value categories, which correspond to unnamed concrete value, and not to named abstract values.

Table 1 summarizes six of ten value categories that we have identified, with sample source citations (for an extensive survey, see [13]). They are mostly based on psychological literature [5,6,14] or consumer psychology [12]. Although Maslow [14] writes about *needs* rather than values, we include these, reflecting extensive overlaps between concepts of not only needs, motivation and values, but also emotions and feelings. In Table 1, similar values are grouped into main categories. The most extensive is social; which values influencing people and relationships. Other categories relate to self, traditions and universal welfare.

There are two further categories of perceived value specific to human product relationships. Pura [15] mentions *monetary* and *convenience* value categories in her review. For monetary value, the product is seen as means of fulfilling tasks to derive monetary value. Convenience value gives a person ease and speed for achieving a task effectively and conveniently. Also, Boztepe [16] stresses the *utility* value of products alongside other value categories. A similar specific category, *conditional* value [12] only arises in a specific context, e.g., buying Christmas cards.

Knowledge of these ten categories of potential user and product related value(s) identified above can guide design of sentences for completion, illustrating the value of *receptiveness to theories* in WCD [1]. These categories can guide development of introductory sentences that probe general reactions to life and the focus topic, but are open enough to not prime responses. However, more focused sentences must probe specific value categories that are important to potential users. Before combining sentence completion with worth maps, we explored the former in a different setting.

3.3 A Familiarisation Study in Sentence Completion

A familiarisation study focused on user values towards exercise, which interested an industry collaborator in the Finnish VALU research project. Table 1 supported brainstorming to develop 50 incomplete sentences, which were pilot tested with a 41 year old man and two women aged 35 and 30, revealing different values towards exercising. The man did not like exercising, but instead shared children's joy as he coached football. One woman enjoyed the physical experience and associated relaxation of exercise. The other woman was more goal-oriented, trying to improve her performance. After this pilot, ten families were recruited to try out sentence completion alongside a familiar established research instrument, semi-structured interviews, with well documented advantages, insights, disadvantages and pitfalls in the research methods literature. The 50 incomplete sentences from the pilot test were reduced to 25. Table 2 shows examples of incomplete sentences and parental responses, translated from Finnish. We recruited selecting different sizes, gender/age compositions and types of family (e.g., single parent). Participants were aged 32-43 years, had 1-4 children, and were given movie tickets to compensate participation.

Accompanying interviews lasted from one to one and a half hours, with all but one at respondents' homes, with assistance from an industrial collaborator. Before each interview, family members were told about the research in general and the procedure, then one parent was asked to fill out the sentence completion task and a background information form. During interviews, we discussed the family's, and particularly each child's, ways to spend time and exercise. There were three main interview themes: family-related; parent-related and child-related questions.

Interviews were audio-recorded and transcribed. Values identified by parents were coded using Table 1 as a tool. For example, if a parent said that it is very important that her child has fun while exercising, this was coded as emotional and hedonistic. To control bias, two researchers independently analyzed results and counted the

Table 2. Examples of incomplete sentences and example parent sentence completions

Category	Example incomplete sentences	Completion by a parent
General	The most important thing to me is...	... the well-being of me and my family.
	My children exercise …	... irregularly, but willingly when they have company.
	It is important in my children's spare time activities that …	... she enjoys it and it is good for her health
Social	My children receive positive attention in spare time activities if …	... he participates or succeeds e.g. meets a goal
Stimulation /epistemic	About my children's way of spending time I want to know…	... she enjoys it and it is good for her health.

values mentioned, both in sentence completion and the interviews. Counts were based on judgements, as interpretation was needed to associate phrases and values. Sentence completion provided slightly more instances of values and more detailed descriptions than interviews, with better foci across a range of value categories.

Sentence completion often uncovered values unreported in interviews. For example, one parent's sentence completion revealed that, although her child's friends were important in encouraging him to exercise or join in a certain hobby, sometimes they teased him, making him unsure and passive. However, during interview, she only mentioned teasing at a general level. Also, in sentence completion, nine parents mentioned well-being and health as an important to pursue, but only three mentioned well-being in interview, and none health. Sentence completion was thus slightly more receptive and also required less time to analyse and administer than interview data: each interview lasted from 60-90 minutes, while sentence completion took 10-15.

Sentence completion was focused, direct and relatively efficient in terms of design, administration and analysis. Its speed, straightforwardness and focus would benefit our main design context, where it would be adequate compensation for losing out on the creative insights and rich responses of cultural or value probes. Again, no absolute comparison is being made here. Rather, different values in different development contexts favour one approach over another. Thus use of sentence completion in a family exercise context indicated that the balance of benefits over costs would be adequate for our intended context, where quantitative methods were more acceptable over open, creative interpretative ones. The use of cultural probes in such contexts results in potential misuse through overriding the core principles that guided their invention [10]. From the perspective of development organisations and project sponsors, lower potential costs of sentence completion contribute to its worth.

We concluded that sentence completion was a very acceptable alternative to interviews, especially as once designed, sentence completions can be used with larger samples with far fewer additional skilled researchers (analysis costs will still rise, albeit less steeply than those for interviews). Additionally, online sentence completion surveys are less expensive to administer, easy to modify, generate fast results (usually with good response rates), with data easily exported into suitable formats for analysis.

4 Combining Sentence Completion and Worth Mapping

An industrial collaborator in the VALU research project on user values wanted to explore worth mapping, so we used sentence completion to elicit user values that could be associated with proposed new product attributes for online monetary gaming, now a very popular leisure activity worldwide. The partner, Paf, is a gaming association governed by public law, whose goal is to raise money for the public good by offering gaming to the public. Paf wants to know more about their customers and their deeper motivations to play online monetary games, and thus support further development of online gaming services, while avoiding adverse gambling outcomes in a responsible way. A further objective was to gather more details on players who are less active on Paf's site and compare them to more regular players. Research topics included e-commerce, online gaming behaviour, and selection of the online games and gaming sites. Most published research focuses on exploring pathological gambling (e.g., [17]): few addresses why unaddicted people play monetary games.

While the commercial context is online gaming, the research goal was to underpin worth mapping with sentence completion, leveraging its receptiveness to user views, and to credibly ground design ends in user data. The case study combined independent research and analysis with three collaborative workshops over three months. Two workshops focused on results of sentence completion, prepared initial worth maps, and decided how to present them for a third workshop focused wholly on worth mapping. Formative evaluation insights for the two approaches were gathered during the first two workshops, with a summative evaluation at the end of the third. As with the familiarisation study, close collaboration with an industrial partner with a strong commercial interest in results brought well motivated participants, with approach usage and evaluations going beyond 'academic' exercises.

4.1 Sentence Completion in Support of Worth Mapping

The participant sample was based on Paf's earlier player segmentation studies. It was the first study of its kind, and involved online monetary players in Finland. The two most interesting player groups to Paf were selected. They have differences in age range, income, residential area, and education definitions, as well as with – most importantly – various specific playing behaviours and preferences. Invited respondents were selected from Paf's customer database, based on gaming behaviour on Paf's site matching a segment profile. The participation incentive was a prize draw. Prizes were appropriate to interests for the segmentation description (Group 1, sports item; Group 2, gift vouchers).. The first group included 8 less active players, and the second 34. Note that, referring to the WCD meta-principle of *inclusiveness* [1], only values from two user groups were considered; business value was not systematically addressed through any receptiveness to Paf's product stakeholders.

40 sentences for completion were designed to meet case study objectives, using the approach in the earlier pilot study. Incomplete sentences (translated from Finnish) included: (a) "In my opinion, online monetary games__", (b) "An unpleasant gaming experience emerges__", and (c) "I find myself playing mostly games that__". 17 background questions were answered. Sentences were completed online using the Webropol service, since respondents played monetary games online.

Once online responses had been gathered, a basic quantitative analysis (SPSS t-test) revealed statistically significant differences between active and less active players. Less active players preferred playing online monetary games alone, and their regrets suggest aversion to losing more money than planned. Active players, on the other hand, seem to have fewer regrets and get over losses more quickly. ontent analysis was used to categorize open answers for each sentence completion. Many qualitative differences were found among the player groups, replicating existing customer segmentation studies, but also providing new and more detailed information about players, revealing that each group's gaming behaviours differed in many ways, with different motivations. A chart was created for each sentence completion, with percentages for frequency of responses. Some respondents mentioned several issues and others none, so bar totals could be either side of 100%. Figure 2 shows response frequencies for completing "When playing online, I feel myself _". Bar length indicates frequencies, and can prioritize user values for feature review and innovation, as well contrast gaming behaviour and reasons to play online monetary games within different market segments. For example, in Figure 2, Group 1 players experience more normal, calm, or satisfied feelings when gaming online, while Group 2 players seem to feel very excited or a regretful idiot. Such differences could be explained by different expectations or playing behaviours and their experiential consequences. However, quantitative rankings and comparisons must be used in combination with qualitative data. Numbers can inform design priorities, but should not dictate them.

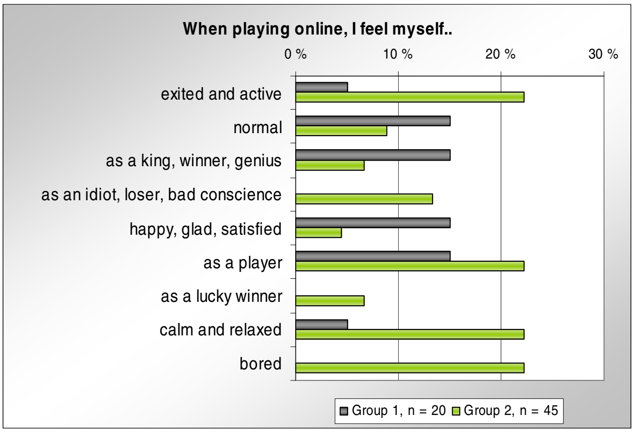


Fig. 2. Example Results of a Sentence Completion

Sentence completion was seen to support credible and expressive receptiveness to user values, although there were some later negative responses within Paf to using percentages with (for them) relatively small samples. Counts on horizontal scales of summative chart may be better. This aside, sentence completion via an existing web survey tool allowed relatively low cost collection of data without researchers telephoning or visiting participants. Sentence completion results extended Paf's existing understanding of users' needs, motives, values, and especially aversions. This broader view of customer value called for new product features, but receptiveness to such features would have to be come from other approaches.

4.2 Initial Worth Map Creation

Figure 3 illustrates a worth map's structure, with a key to symbols on the left (Motives and Aversions respectively correspond to [3]'s Worthwhile and Adverse Outcomes). Upper and lower boxes (motives, UX, aversions) are human value elements that can be grounded in sentence completionI, indicating 'why' people would use a system (motives/benefits) or not use it (aversions/costs). Motives correspond to benefits, and aversions to costs. UX elements are the interactive *means* that link design ends (outcomes) with the design means (qualities, features, materials).

Worth maps support expressivity and credibility by explicitly connecting design elements and UXs to the underlying human motivations to (not) use them. A set of worth maps can visualize a complete product context, highlighting how user or supplier motives associate, via UXs, with design elements to *enable* (solid lines) or *disable* (dashed) positive or negative outcomes [2].

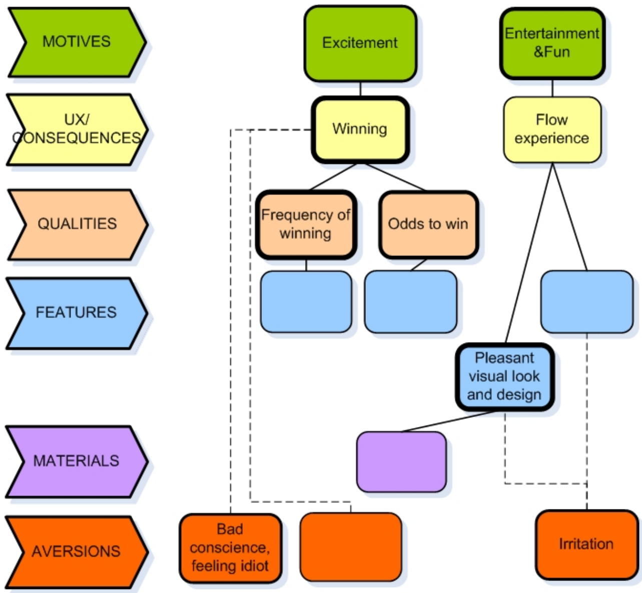


Fig. 3. Example Worth Map Layers

After researcher analysis and internal reporting of sentence completion results, two collaborative workshops further reviewed results and generated initial worth maps for each user group. Researchers and Paf representatives from User Experience and Player Intelligence roles attended. Sentence completion results were reviewed, and identified user values were transformed collaboratively into worth map elements, e.g., 'excitement' became a Motive element for both groups, and 'feeling loser and idiot' an Aversion element for Group 2. Sentence completion frequencies and existing player intelligence identified the most important issues for each group. Some sentence completion data covered design elements, which combined with existing ideas for new features to identify potential design elements to remove specific aversions or deliver specific UXs and their associated (connected) motivating meaning.

Several relevant product features were added to worth maps, connected to specific motives to play based on user data. To identify such connecting causalities between elements, answers of selected respondents in each group were examined question by question. Further new elements were added as they emerged. Elements were connected whenever associations were obvious. A guiding heuristic was that when respondents mentioned two consequent issues in their answers, these could be connected. Also, where workshop participants agreed on logically obvious – or otherwise known – associations (connections between elements), these were added, even without direct support from user data. The aim was to prepare partial worth maps for discussion, revision and extension in a third workshop. However, several identified worth elements were left with no associations connecting them to other map elements. To highlight these, and to aid map subsequent extensions, blank elements were added the partial worth maps to highlight the lack of associated product features.

In [3], manual card sorting and arranging is proposed as a practical basis for worth mapping, where project teams gather around a large table to manipulate an evolving sketch or map. In this case study, a computer drawing tool was used. Once identified, each map element was added to a Microsoft Office Visio diagram, colour coded by category as in Figure 3 (thus creating a worth sketch [3] of unconnected elements). Worth maps were constructed for both player groups. Group 2's map had over 50 elements. Comments indexed by question and respondent number were added to elements to ground them in user data, e.g., "22% of the respondents mentioned they feel excited and active after playing online" for the motive element 'Excitement'. Element boxe edges were given different line weights in Visio corresponding to frequency of mentions, thus highlighting potential high priority elements. This creative use of Visio extended the credibility and receptiveness of worth maps by directly and inspectably grounding human value elements in user data.

Another very effective innovation was a researcher's use of Visio drawing layers to modularise complex associations. All worth map elements were placed in the base layer, with each subsequent layer containing associations for means-end chains passing through a single UX element, making such chains easier to follow by hiding other layers. Element frequency coding, respondent comments, line weights, blank elements and means-end chain layers combined to improve worth map expressivity, and thus communicate them better to people not involved in the mapping process.

Figure 3 shows two layers for Group 2, showing means-end chains for the *Winning* and *Flow* UXs, which should lead to motivating outcomes of *Excitement* and *Entertainment & Fun*. These UXs were thought to set requirements for the gaming site. To experience winning, frequency of winning and odds to win are key design qualities. Similarly, to experience flow, the site needs a pleasant visual look and design. No other design elements in support of positive UXs have been identified in Figure 3, hence blank elements draw attention to unsupported values and experiences. The 'pleasant visual look and design' element is a place holder for more extensive understandings of why existing layouts and sound may be inadequate. Worth maps support expressivity for connections between design choices, but not for choices themselves, which must be expressed by other approaches (e.g., Figure 2).

Identified aversions were also added as map elements, e.g., when not winning, but losing much and suffering *a bad conscience*, the excitement motive is not fulfilled. Similarly, *irritating* media usage will not lead to flow, entertainment and fun, as indicated by a dashed 'aversion block', indicating that a design element can help to avoid an adverse outcome. A dashed aversion block to 'irritation' requires subsequent close review of associations between multimedia features and usage outcomes. The extent of required creativity and sustained design management in WCD should not be underestimated. Design elements in initial worth maps are only possible means to achieve desired ends. Subsequent iterative design and evaluation is needed to show actual value achievement when users interact with new features. Figure 3 is thus one starting point for the third workshop. It is only one example of what worth mapping can achieve. The partial worth maps were extended further in the third workshop.

After two workshops and a familiarisation study, researchers and collaborators proceeded confidently to a third workshop, having seen sentence completion reveal what is (not) worthwhile for users. When presented as charts, results supported worth mapping, facilitating familiarity with data and staying anchored in it. However, with larger samples and sizable question sets, comprehensively walking through all data is too labour intensive. Data should be progressively sampled in reasonable chunks, with collaborative interpretation focused on qualitative insights and new product ideas. It is better to exploit sentence data in prioritised chunks, letting worth mapping proceed as collected data is being considered.

There was also early evidence that separation of design elements into materials, features and qualities was not helpful in this context. Such distinctions make sense to Interaction Designers and hardware and software engineers, but are largely irrelevant in high level mapping, especially when few developers or engineers are involved. Also, in the target web context, materials had little relevance (unlike in mobile and ubicomp settings). Hence materials were rarely used, and qualities and features were sometimes confused, e.g., 'a pleasant visual look and design' is both a *quality* ('pleasant') and a *feature* stub ('visual look and design'). Similarly, another map element, 'diversified, interesting games' (not in Figure 3), combined two qualities and a feature stub. A sensible appropriation here would have been to use a *single unified product attribute element* to avoid distraction until there was a real need to distinguish materials, features and qualities. Brainstorming could just identify 'design ideas'.

4.3 Worth Map Review and Extension

A key research objective of the third workshop was a summative evaluation of worth mapping. Paf's commercial goals were to share knowledge, understand how current services deliver identified customers' values, and how they can be improved (via brainstorming design ideas), as well as developing an understanding of worth maps' usefulness for Paf. Worth maps supported brainstorming on possible innovations for product design and marketing/customer relationship management (CRM).

The third workshop began with a 90 minute session before lunch that presented sentence completion results and introduced worth maps. In three hours after lunch, two teams reviewed, discussed and extended partial worth maps for each user segment, each time making a presentation to the other group. Each reviewed a worth map for both market segments of interest. Each team was facilitated by a researcher, with two Paf staff in each (overall, there were three product owners who had not taken part in previous workshops, plus a player intelligence analyst who had).

Understanding the worth maps was initially challenging, highlighting a need to explore ways of introducing prepared worth maps to people who had not constructed them. However, as participants came to understand how worth maps associate design features with user values, they were able to reassess existing ideas for product innovations, and also re-evaluate how products could be positioned in terms of a wider range of values, or to highlight alternative value propositions. Reconsidering existing proposed innovations within the context of worth maps did reframe how they were assessed. Also, ideas for additional design elements were proposed during discussions, spanning a wide range of design features from visual re-design, user control over some game attributes, and social computing. Again, the details of these features are not relevant to worth mapping, nor is their originality. Most additional features had already been identified in previous product review sessions before the worth mapping study. However, these features could now be closely associated with user motives and/or aversions. The credibility of such features changed, as did the extent of

Table 3. Summative Evaluation Ratings

Question (some slightly reworded)	**Av.** (max 4)
1. Goals of mapping are understandable	3.50
2. Understanding user values is important in our business	3.75
3. Worth maps help to make user values, needs and motivations explicit	2.75
4. Worth maps help to design more successful products	2.67
5. After worth maps, I have a better conception of value of products/services	3.50
6. The worth map method is easy to use	2.75
7. Will use worth maps in the future	2.33
8. With worth maps, I can better connect product features with what really matters to users	3.25
9. With worth maps, I can better brainstorm	2.75
10. With worth maps, I can better develop a suitable marketing communications and manage customer relationships	3.33
11. With worth maps, I can estimate business value and use this in strategic decisions	1.75
12. Worth mapping results are easy to communicate further on	2.50
13. Worth maps suit company purposes well	2.75
14. Worth maps are useful in my work tasks	3.00
15. Would like to learn more on worth maps	2.67

commitment to them by the project team. Features that were previously inadequately motivated could now be directly related to user costs and benefits. Worth mapping thus supported re-evaluation of existing or novel product features. Surprisingly, it also supported marketing and CRM innovations. Some marketing communication samples were briefly reviewed by each team, leading promisingly to new ideas about content and style. The lower relative cost of changing marketing communication materials meant that this could quickly translate into action.

Much of the above assessment is autobiographical. However, given the expectation that approaches are designed for appropriation, that is a key to any assessment of WCD frameworks. Thus innovative uses of Visio extended the expressivity of worth maps. In particular, use of layers avoided crowded, impenetrable diagrams in a way that is not possible with large single representation of 'table top' maps as advocated in [3]. We ended the third worth mapping sessions with a short participant questionnaire. The four commercial participants completed an evaluation form comprising 15 questions with a don't know option plus 4 point Likert scale (1=Strongly Disagree; 2=Disagree; 3=Agree; 4=Strongly Agree), followed by six open questions, which were not restricted to the scope of worth maps and sentence completion as understood in terms of WCD meta-principles. Thus despite low receptiveness to business value, we evaluated this and other 'out of scope' factors: Question 11 addresses inclusiveness of business value, and Question 4 addresses improvability. The results are shown in Table 3. As would be expected, the least favourable response is to Question 11. More inclusive use of maps needs (perhaps additional) user experiences, motives and aversions to have (additional) explicit associations with business beneficiaries. Business outcomes thus must be addressed separately from user outcomes, but must still be associated with them. Explicit direct receptiveness to business value would be required to improve responses to Question 11.

For success in practical settings, only means over 3 can be taken as good indicators of suitability for widespread use, which have been achieved for worth maps' main purpose of supporting *designing as connecting* ([1,3] Question 8) and *committedness* (Question 5). The mean score for Question 7 is a fair response overall, Improving this requires improvements on other responses, e.g., to Questions 3, 4, 6, 9, 12, 13 and 15, through use of approaches beyond worth maps and sentence completion, ideally making more use of existing Agile and relevant business approaches within Paf. Although almost 'good enough', brainstorming (Question 9) could be further improved by the use of other projective instruments reviewed above. Even so, the mildly positive responses on outcomes beyond the intended purposes of worth mapping and sentence completion are encouraging. Also, there was a fairly even distribution of positive and negative responses across roles, suggesting a lack of bias for product owners or marketing/CRM. Responses to open ended questions were mostly (very) positive, e.g., "Good tool to visually work with information", "good overview", "helps to connect abstract and concrete", "offering fine tuning and new products are found with it", "good for boosting workshops", "tool to brainstorm and to try to innovate new features", "possible to see [product] strengths and weaknesses". One participant preferred 'traditional' closed questionnaire formats to sentence completion, believing the former to be less leading, requiring less time for analysis and less researcher interpretation. However other participants valued sentence data as complementing existing marketing data, strongly indicating values. Three participants had not attended the first two workshops: one found worth maps difficult and confusing, especially the map element categories. Another found the density of connections challenging. Both could be avoided by further innovative use of drawing tools and improved workshop preparation and tutorial materials.

4.4 Later Practical Utilisation of the Results

Since the workshops, all project deliverables have been published via the company Intranet and Wiki. Three presentations, open to all interested personnel, were held: two on site in two of Paf's biggest office locations, and the third via overseas video conference. Following each on site presentation, a further workshop was held, where worth maps were walked through layer by layer, each showing a part of the whole map that was related to one of 4-6 'top issues', as identified through each player group's values. One workshop was more markets and sales focused, the other more product focused, in terms of participants and facilitation. In both, many evaluative notes were produced and discussed: some were evaluations of the existing offering; some were related 'design ideas' – some new, and some previously conceived but now further supported or differently anchored. Many ideas were quite high level (or epic!), which is in line with worth maps' expressivity, but some were specific. After the workshops, a collection of notes (as high resolution digital images), and a written summary of emergent topics were shared among the invitees. The third presentation drew most of all country organizations' management. Besides introducing all the project deliverables, a summary of emergent topics was briefly presented. It is now up to each business manager, product owner, and other key people to form their overview of the deliverables, and use this as a part of various 'knowledge building blocks' supporting their daily work. However, the investment of resources in follow up presentations and workshops indicates that worth mapping and sentence completion provided genuine business value to Paf. Of course, more work must be done to realise commercial value, and no payback can be assessed until new features and CRM/marketing are implemented, but there is clear competitive potential.

5 Conclusions and Future Work

The two combined WCD approaches can be assessed relative to the meta-principles that they aim to support. Worth maps aim to support committedness and expressivity, while sentence completion supports receptiveness, credibility and (through bar charts) expressivity. *Committedness* is the extent to which a project team has made explicit their commitment to design means, design ends, intended beneficiaries and evaluations. *Receptiveness* is indicated by openness to alternative design means and ends (i.e,, user values), beneficiaries, and approaches to evaluation. *Expressivity* is assessed in terms of how well design approaches in use communicate design options and choices. *Credibility* is assessed in terms of the feasibility or groundedness of a design choice. Worth maps fared well on committedness (Question 8) but less well on expressiveness (Question 12), although there were some positive results (Question 3).

Both worth maps and sentence completion are WCD approaches that let project teams reflect on worth as the likely balance of costs (as adverse outcomes/aversions) and benefits (as worthwhile outcomes/motives). Both performed well here. There were very positive responses on understanding the goals of worth mapping (Question 1) and their ability to improve understandings of product or service value (Question 5). Sentence completion was evaluated indirectly, but had a role in positive responses on Questions 3 and 10. The two WCD approaches thus combined well to provide good bases for understanding user values. Sentence completion helped to elicit values that users often find difficult to articulate. It also surfaced experiences and emotions, and not just goals and needs, as well as aversions, which were particularly fruitful when brainstorming product improvements. Worth maps linked user values, needs and motivations explicitly to design elements. Explicit sentence completion data on user values and aversions lets worth maps be grounded in user viewpoints.

The practical efficiency and effectiveness of worth mapping was improved through the use of Visio, but further innovative tool support and tutorial materials could reduce the time taken to understand worth maps authored by others. This would better support combined industrial use of worth mapping and sentence completion to let project teams innovate and appropriate methodologically to deliver worth.

Simplifications are also possible, such as a single 'design idea' category to replace qualities, features and materials. This should make worth maps easier to produce, use and communicate to colleagues who were not involved in constructing them. Separate worth map elements for qualities, features and materials are better suited to R&D teams with a predominantly creative and technical focus (i.e., user experience researchers, interaction designers, software/ hardware engineers). Worth maps with only ‘design idea’, UX and outcome elements could be initially developed by project teams with mostly business roles (e.g., product or service owners, marketing and CRM specialists, brand strategists), for subsequent refinement of ‘design ideas’ by a more R&D focused creative and technical team. Modifications to approaches and project practices should be able to make worth mapping more worthwhile, but even so, its first use without its inventor present did result in worthwhile outcomes.

References

[1] Cockton, G.: Getting There: Six Meta-Principles and Interaction Design. In: Proceedings of CHI 2009, pp. 2223–2232 (2009)
[2] Cockton, G.: Designing worth is worth designing. In: Proc. NordiCHI 2006, pp. 165–174 (2006)
[3] Cockton, G.: Designing Worth – Connecting Preferred Means to Desired Ends. Interactions 15(4), 54–57 (2008)
[4] Hoyer, W.D., MacInnis, D.J.: Consumer Behavior, Houghton Mifflin, USA (2007)
[5] Rokeach, M.: The Nature of Human Values. Free Press, New York (1973)
[6] Schwartz, S.H.: Universals in the content and structure of values: theoretical advances and empirical test in 20 countries. Adv. in Experimental Soc. Psych. 25, 1–65 (1992)
[7] Donoghue, S.: Projective techniques in consumer research. J. Family Ecology and Consumer Sciences 28, 47 (2000)
[8] Banister, E.N., Booth, G.J.: Exploring innovative methodologies for child-centric consumer research. Qualitative Market Research 8(2), 157–175 (2005)
[9] Gaver, B., Dunne, T., Pacenti, E.: Cultural probes. Interactions 6(1), 21–29 (1999)
[10] Boehner, K., Vertesi, J., Sengers, P., Dourish, P.: How HCI interprets the probes. In: Proc. CHI 2007, pp. 1077–1086 (2007)
[11] Voida, A., Mynatt, E.D.: Conveying user values between families and designers. In: CHI 2005 Extended Abstracts, pp. 2013–2016. ACM, New York (2005)
[12] Sheth, J., Newman, B., Gross, B.: Consumption Values and Market Choices, Theory, and Applications. South-Western Publishing, USA (1991)
[13] Kujala, S., Väänänen-Vainio-Mattila, K.: Value of information systems and products: Understanding the users’ perspective and values. JITTA 9(4), 23–39 (2009)
[14] Maslow, A.: Motivation and personality. Harper and Row, New York (1970)
[15] Pura, M.: Linking perceived value and loyalty in location-based mobile services. Managing Service Quality 15(6), 509–538 (2005)
[16] Boztepe, S.: User Value: Competing theories and models. Int. J. Design 1(2), 57–65 (2007)
[17] Lee, C.-K., Lee, Y.-K., Bernhard, B.J., Yoon, Y.-S.: Segmenting casino gamblers by motivation: A cluster analysis of Korean gamblers. Tourist Management 20(5), 856–866 (2006)

What Is an Activity? Appropriating an Activity-Centric System

Svetlana Yarosh[1], Tara Matthews[2], Thomas P. Moran[2], and Barton Smith[2]

[1] Georgia Institute of Technology, Atlanta, Georgia, USA
[2] IBM Research, Almaden Research Center, San Jose, California, USA
lana@cc.gatech.edu, {tlmatthe,tpmoran,n6hdn}@us.ibm.com

Abstract. Activity-Centric Computing (ACC) systems seek to address the fragmentation of office work across tools and documents by allowing users to organize work around the computational construct of an *Activity*. Defining and structuring appropriate Activities within a system poses a challenge for users that must be overcome in order to benefit from ACC support. We know little about how knowledge workers appropriate the Activity construct. To address this, we studied users' appropriation of a production-quality ACC system, Lotus Activities, for everyday work by employees in a large corporation. We contribute to a better understanding of how users articulate their individual and collaborative work in the system by providing empirical evidence of their patterns of appropriation. We conclude by discussing how our findings can inform the design of other ACC systems for the workplace.

Keywords: Activity-Centric Computing, Appropriation, Office & Workplace.

1 Introduction

Professionals whose work involves the creation, management, and distribution of information (known as *knowledge workers*) expend considerable effort managing work that has been divided across teams of contributors and a diverse collection of tools for handling documents, communicating, updating status, etc. [1,2]. Activity-Centric Computing (ACC) is an approach to address work fragmentation by allowing users to structure their work around the computational construct of an Activity [1,3-5]. This approach is loosely inspired by Activity Theory, a framework for studying socially situated phenomena originally developed by the Soviet school of Psychology. Adopting this framework for HCI, the term "activity" is used to refer to the basic unit of social analysis broadly defined as collective action towards a goal [6]. We use "Activity" (capital A) to refer to the computational construct and "activity" to refer to the social unit of analysis. Thus, an Activity is a locus for uniting information relevant to carrying out an activity and a persistent representation of its status, members, resources, constraints, and plans [1,3]. In this paper, we present an empirical investigation of how users successfully structure their work activities as system Activities in an ACC system.

The ACC approach leverages the Activity construct to support the informal and emergent situational work of small groups, as has been validated through prototype

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 582–595, 2009.

deployments (e.g., [1,3]). ACC systems have several distinguishing characteristics. They focus work around activity goals, which separates them from communication systems like email. Unlike workflow systems, they allow the flexibility of on-the-fly (re)organization, are generic enough to represent a wide range of activity, and support lightweight creation and editing. To provide this generic and flexible support, ACC systems leave the responsibility of defining and structuring Activities to the user. However, prior studies have shown that this poses a major challenge: it is hard to delineate an activity's purpose, boundaries, and representation within a system, especially at an early stage [1,2].

Representing work in an ACC system involves thinking about how the Activity fits into existing working practices. Dourish [7] defines *appropriation* as the process of adopting and adapting systems to fit and evolve existing work practices. Workers must appropriate the Activity construct and the ACC system to benefit from its use. Yet, we know little about how they define, structure, and use an Activity in the process of doing real work. We address this gap by presenting empirical evidence of usage patterns and strategies emerging from the appropriation of one ACC system by some of its most active users.

We studied a production-quality ACC system, Lotus Activities. We present an investigation of 15 knowledge workers who have appropriated Lotus Activities for their everyday work. We focus on highly active users because they are most likely to have developed successful patterns and strategies of use that can inspire potential design directions of ACC systems. We address two questions to inform the design of ACC systems:

1. What use patterns do active users develop in appropriating Lotus Activities?
2. How do they divide and structure their work into effective Activities?

We begin by positioning our work in the context of other investigations of ACC systems. We describe Lotus Activities and its usage in one company. We then describe our interviews with 15 active users, and a content analysis of 120 of their Activities in the system. We present the patterns of use that emerged from this investigation. Finally, we discuss the appropriation of Lotus Activities and the Activity construct as evidenced by these patterns, highlighting the implications of our findings for ACC systems.

2 Related Empirical Work

Several activity-centric systems have explored this paradigm prior to Lotus Activities. Christensen and Bardram [4] promoted an activity-centric approach, grounded in the evaluations of an off-the-desktop prototype with health care workers in a series of workshops. Bardram *et al.* [1] developed operating system extensions to provide personal computer users with an activity-centric experience. Though feedback for the activity-centric approach was positive, they found that users had difficulty in defining activity boundaries in the system. The UMEA system [5] attempted to address the problem of partitioning work by inferring personal project categorizations from the user's interaction with other tools. While the automated approach reduced the work needed to define Activities, it introduced substantial overhead to remove incorrectly associated items from projects. Over a long-term deployment of Giornata [8], a system supporting reframing one's desktop into manually declared activities, Voida & Mynatt found that users had trouble defining Activities because of the interdependent and complex nature of knowledge work. Bellotti *et al.* [9] explored the use of an Activity construct for email and to-do management to find that the construct was flexible enough to support individual PIM needs, but that users could benefit from suggestions of potential Activity structures. Our work can be distinguished from these systems because it focuses on collaborative rather than individual Activity management, but all of these explorations reveal a common problem: dividing and structuring work into Activities is a challenge for users.

Most closely related to our work is a study of ActivityExplorer [3], an ACC system organized around flexible threaded collections of diverse media objects (e.g., files, chats, tasks). A deployment with 33 researchers and interns over the course of 3 months provided an initial validation for ACC as a design paradigm. Muller *et al.* [10] found that interns appropriated the system in surprising ways, e.g. as a community forum. We extend this line of work with a deeper exploration of emergent use patterns of a production-quality ACC system by knowledge professionals, who adopted the system of their own accord (i.e., our users were not recruited). Balakrishnan *et al.* [1] have recently examined the adoption of this same ACC system in the context of other tools available to teams. In contrast, we focus on another pervasive open question in activity-centric work: How do users successfully structure their work as Activities? We address this through an empirical investigation of active users of one ACC system.

3 Overview of Lotus Activities: An ACC System

Lotus Activities is a Web-based service. On the entry page the user is presented with the following text explaining the system:

> *Organize all your people, work, and collaborations associated with a business activity in a single place. Share files and web links, post messages, track to-dos, and manage your deadlines with Activities. Access the web-based dashboard from anywhere, or use plug-ins to extend your existing collaboration tools.*

Each Activity is described by a title, a set of tags, a list of members, an optional due date, and content entries posted by the members. To initiate a new Activity the user only has to provide a title. Members can be added at any time: "owners" who have complete access; "authors" who can add entries but not edit those of others; and "readers" who can look but not contribute. Activity entries can be of various types: basic text posts, to-dos, and threads of comments. Any entry can have attachments, links, and tags. Only the members of an Activity have access to its contents, unless the Activity is made "public" to all Lotus Activities users. Figure 1 shows a screenshot of an Activity.

After logging in, the user sees a "dashboard" page, listing her current Activities. To manage the list, the user can sort, view by tag, mark as complete, or hide any Activity. The user can also view a combined list of all to-dos across all Activities. Users can send email notification when adding or editing an entry. To organize entries, users can group them into sections (like folders). To enable work reuse, any Activity can be converted to an Activity Template. When creating a new Activity, a user has the

Fig. 1. Screenshot of the web interface to an Activity

option to instantiate a Template. This new Activity will have the structure and content of the parent Template.

A total of 38,719 Activities (containing at least one entry) have been created by employees of a large, global company since May 2006. The average activity has 3.4 members (*SD*=9) and contains 12.1 entries (*SD*=29.9). There are 31,913 users who are members of at least one Activity. Most users are members of 10 or fewer: 7,286 users have 1 Activity; 7,873 have 2 to 5; 20,073 have 6 to 10; 2,234 have 11 to 50; and 296 have 51+.

About a third (33.5%) of Activities were modified at least once during the course of our study (May-July, 2008), indicating a large number of in-progress Activities, as well as a large history of completed (or abandoned) work. We conducted an in-depth qualitative investigation of how 15 active users employed Lotus Activities for their everyday work.

4 User Study Methods

We conducted interviews with 15 users and analyzed the content of 120 publicly-shared Activities to investigate patterns of use and strategies for appropriating Lotus Activities.

Table 1. Demographics and Lotus Activities use of each participant

Code	Gender	Formal Job Title	# Activities	# Entries
P1	M	Techn. Enablement Spec.	61	999
P2	M	Technical Sales Manager	13	193
P3	M	Certified IT Specialist	23	570
P4	M	Software Developer	80	1579
P5	M	Software Sales Leader	17	281
P6	M	Sales & Distr. Consultant	45	826
P7	M	Technical Sales Spec.	17	212
P8	M	Support Ops. Analyst	16	76
P9	M	Research Staff Member	43	230
P10	M	CIO Architect	25	408
P11	F	Client Relationship Rep.	226	3559
P12	M	Senior IT Specialist	11	127
P13	M	DB Support Team Leader	14	91
P14	M	Software Architect	83	173
P15	F	Level 3 Service Specialist	73	649

4.1 Participants

We focused our investigation on highly *active users*, as they represent workers who have appropriated the system in ways that are useful to their work. We solicited by email users who owned the most public Activities or authored the greatest number of entries in these. We continued interviewing users until we reached a point of data saturation (as defined in [12]). These participants represented the variety of job roles, work locations, and experience levels typical of their company (see Table 1). To understand how these users structured and divided their work in Lotus Activities, we studied their public Activities through a content analysis and their private Activities in interviews and system tours.

4.2 Interviews and System Tours

We conducted 1-hour, semi-structured interviews by phone. Each participant was asked to provide an overview of their job responsibilities and current projects. We asked about their initial motivation for adopting Lotus Activities and current system use. Finally, we asked users to show us a few of their Activities using web-conferencing software and to discuss how they organized and managed their work in the system. These tours allowed us to look at Activities that were not publicly shared. We analyzed the interviews using a grounded theory [12]. Distinct themes began to emerge after two independent passes of open coding and memoing followed by axial coding.

4.3 Content Analysis of Participants' Public Activities

Following the interviews, we created a list of all public Activities and Templates in which our participants authored entries or were listed as members. We did not look at any private Activities outside of the interview to preserve their privacy. Our participants had 152 public Activities/Templates, representing 22% of the 692 such Activities/Templates in the system. We excluded 7 Activities that represented blank or unedited instances of existing Templates and those where participant membership

Table 2. Activity patterns found, indicating the percent of the 120 public Activities for which the pattern was present (some Activities contained multiple patterns) and the number of participants (out of 15) who took part in the Activity of this pattern or discussed it in an interview

Pattern	% Activities	# Participants
Set of Job-Specific Work Units	48 %	12
Create a Document	14 %	13
Organize an Event	9 %	10
Solve a Problem	48 %	9
Collect Information	19 %	13
Formalized Exchange	3 %	4
Personal Tasks	23 %	11
Team Space	5 %	11
Create a Tutorial	13 %	10

was in question. One public Template led to 25 public Activity instances. Including all 25 would have incorrectly overrepresented this type of Activity, so we only included the original Template of this set in the analysis. We analyzed the remaining 120 Activities/Templates to complement the interviews as a source of evidence on the participants' patterns and strategies of use. We performed initial open coding, focused on the patterns of use and strategies evidenced in the structure and content of these Activities and Templates. In the process of axial coding, we found that the use patterns evidenced in the interviews and the content analysis converged. We proceeded to assign one primary and up to two secondary pattern codes to each of the 120 public Activities/Templates.

5 Results: Observed Patterns of Use

In the course of our analysis, we observed that Lotus Activities users developed specific structures and uses for the generic Activity construct. Our interviews and content analysis revealed nine patterns of use, summarized in Table 2. The patterns were informed by our participants' public *and* private Activities, so they represent the types of patterns employed by active users, overall. As these patterns emerged, we considered the open question of how users go about dividing and structuring work into Activities. This section provides the empirical groundwork needed for us to revisit this question in the discussion section, focusing on describing the patterns and strategies of use. We discuss the most interesting patterns individually, grouping others with similar structure.

5.1 A Set of Job-Specific Work Units Pattern

The "set of job-specific work units" pattern was characterized by a set of identically-structured Work Unit Activities, with a single Activity representing an active unit of work for which the user is responsible (e.g., a customer account, a software version). Multiple Work Unit Activities of the *same type* are often active at once, and they are used to track status and store content. P11, a sales representative, had over 50 active Units, each representing an account containing different sections for brands and opportunities. Other examples included work requests (P8, see Figure 2), proof-of-concepts (P5), sales plans (P2), customer reports (P15), invention disclosures (P1), and product-client pairs (P12).

Highly specialized tools have arisen to support specific types of common, repeated activities. For example, CRM tools like Sales Force[1] help manage customer accounts, and development tools like Rational Jazz[2] help users coordinate software debugging tasks. Lotus Activities does not replace these tools for knowledge workers, but rather supplements them with a space for ad-hoc collaboration and everyday planning. For example, P11 continues to use the company-prescribed CRM tool (Siebel) to account for her work, but also uses Lotus Activities to define a structure for each account that better matches her everyday needs:

> *I still definitely use Siebel, because we have to. But, like most CRM systems, they are for reporting only… It's great for management, but it's not so great for the sales reps, who have to track their own opportunities and their own communications with the client.*

Users transferred work to Lotus Activities because they saw their current tools as inadequate in some way, borrowing structure from those prior tools. Participants cited ease of maintaining and sharing Unit status and resources as the main reason for adopting Lotus Activities. Before Lotus Activities, P11 supplemented Siebel with a complex combination of file folders, email folders, and spreadsheets for each account. Transferring her organizational system to Work Unit Activities allowed her to consolidate all of this information and easily keep team members updated of changes to accounts through notifications. Starting with the established structure in Siebel made it easier to decide on an initial Activity organization then tailor it to specific needs.

A common strategy users employed to create Work Units was to design a Template to formalize the structure and process for a single Unit. Because the Work Units were

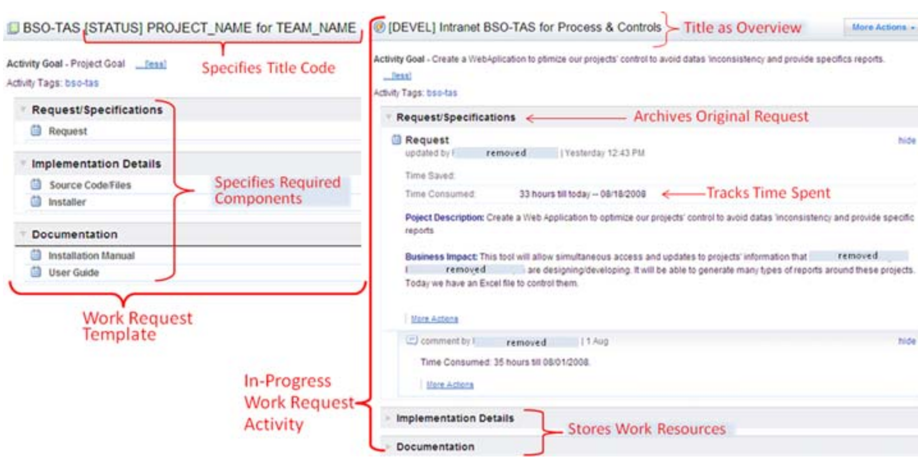


Fig. 2. Template for a work request work unit and an in-progress Activity of this type

[1] www.salesforce.com
[2] www.ibm.com/rational/jazz

persistent structures to which users referred often, they were willing to spend more time organizing the information. Dividing work into multiple Activities required developing strategies for understanding the current status across Units. Five participants developed conventions for including status codes in Activity titles, making status visible from the list of Activities on the dashboard.

Transferring process knowledge to others was another advantage of structuring Activities as sets of Work Units. Users appreciated that Activities exposed the *process* of carrying out the activity to new teammates by saving to-dos, intermediate drafts, and comments, as P15 related:

> *I had to show [a new team member] how to properly do the [customer report]… I just pointed her to the Activity… I didn't have to do a lot of explaining of the [customer report] process. It was all kind of self-explaining.*

The Set of Work Units pattern demonstrates a non-trivial translation of work activities into the system, splitting multiple threads of a job responsibility across many highly-structured Work Unit Activities.

5.2 Complete Work Activity Patterns: Documents, Events, and Problem Solving

The next three patterns—Create a Document, Organize an Event, and Solve a Problem—are Activities bounded by a clear goal and a well-defined endpoint. Thirteen users employed Lotus Activities as a space for coordinating the creation of a document, such as a report or a presentation. Ten participants organized events like workshops, business conferences, and sales demonstrations. Nine users had Activities aimed at resolving a problem in software development, customer issues, and policy decisions. These patterns all involved taking notes or collecting ideas/information, brainstorming, and/or making decisions. Having a shared space for coordinating production, planning, or problem solving was important to users because it helped improve awareness of the activity's status and reduce email volume during the process, as P12 explains about a document creation Activity: "…rather than sending these different files around, I can just stick it into an Activity and send a link around and… then everybody has access to the most current version."

5.3 Partial-Activity Patterns: Collecting Info and Exchanging Formalized Info

The next two patterns, Collect Information and Exchange Formalized Information, served to manage content intended as a resource for a larger activity (which may or may not have been represented in Lotus Activities). The Collect Information pattern involved using an Activity as a space to collect sources of information (text, links, attachments, etc.) on a single topic. Used by 13 participants, Collections ranged from roughly task-focused (e.g., use cases for a software product) to loosely connected around a topic (e.g., list of Web 2.0 development resources). A common strategy for generating Collections was to add ideas in an ad-hoc manner, as P15 noted: "I just started putting it here whenever I learn something… Activities is a good place for me to use kind of as my brain."

In contrast, Formalized Information Exchanges, which involved requesting and managing information from multiple people, were planned and usually made into

Templates. These Activities, used by 4 participants, included instructions requesting specific information. For example, P14 created a Template to collect demo proposals for a conference he was organizing, asking interested users to instantiate the Template and fill it out.

5.4 Cross-Activity Overview Patterns: Personal Tasks and Team Spaces

Participants found it valuable to create or maintain overview perspectives that spanned multiple Activities. Managing Personal Tasks and Team Spaces represent two views of work that cut across Activities: one unifying individual work, the other unifying work across a team. About two-thirds of users reported having Activities which only they could access, containing personal to-dos. These Personal Task Management Activities were developed by users who wanted the Lotus Activities to-do list view to contain a complete overview of their commitments.

Team Space Activities sought to unify content related to managing a team to provide an overview across projects and simplify bringing new members up-to-date. They were often organized into sections or discussion threads that included elements like meeting minutes, project timelines, accomplishments, and links to other team Activities. They also often included elements of other patterns, such as solving a problem, planning an event, or creating a document. These Activities represented a team-level view of past and ongoing work, usually maintained with considerable effort by the team's manager or coordinator.

5.5 Teaching and Reuse Pattern: Tutorials

While other Activity patterns focus on performing work, the Tutorial pattern was created explicitly to guide other users (10 participants had these). We saw two types of Tutorials: those intended to be read for educational purposes, and those intended to be instantiated to guide users through a process. P15 described the former: "I take entries from several different problems… [so that] others can now see how to solve this problem. Instead of coming back to me with questions, they can just refer to that."

One strategy for creating a new Tutorial was copying a successfully completed Activity. Another strategy was reproducing the steps in existing policy guides or paper checklists as a Template. Tutorials created as Templates provided an advantage over static guides or paper checklists: they could be used to create an independent Activity and tailored to the needs of the user carrying it out, as noted by P7: "I wanted to give people the freedom to change [their Activity]… and let people have more control over it."

It was difficult to make changes to tutorials over time, because instantiated Activities do not reflect Template updates. P12 was anxious about this limitation and "… felt like I really had to get it right before I put it out for people to use."

Once content was created, authors often publicized their Tutorials by posting links to them on wikis, blogs, and other resources that were indexed for search across the enterprise. This step was necessary, because the search function within Lotus Activities was too unreliable for users to effectively discover Activities and Templates.

5.6 Relationship and Evolution between Patterns

Our content analysis enabled us to observe relationships between patterns commonly found together in the same Activity. Creating a Document and Collecting Information were often components of other patterns, such as Work Units, Team Spaces, Solving Problems, and Planning Events. Work Units and Team Spaces were most likely to include other patterns, because they served as organizing Activities.

Patterns that were loosely-organized and unbounded by time (e.g., Information Collection or Personal Task Management) could evolve into more structured Activities. For example, P7 described an Activity created with a client to "share information" and "reduce the amount of email" (an Information Collection) that, with time, began to represent a history of work with the client that included meeting notes and to-dos (i.e., a Team Space): "We brainstormed some work items… We would use the Activity to track them and define actions and dates so that we would end up with a kind of project timeline." Next, P7 planned to remove client-specific information, make a Template of this Activity, and reuse it. Thus perhaps, we were seeing the evolution of a new Work Unit.

Tutorials were patterns that commonly evolved from other Activities by rearranging or merging content, adding directions, and/or making Templates from them. For example, P3 created a Tutorial guiding soon-to-be parents through company paperwork by making a Template from a Personal Task Management Activity, where he had collected his own to-dos and resources as he was going through the process.

Activities often represented multiple patterns and could evolve over time, so deciding on a structure for an Activity was not about just identifying the "best" pattern to use. In the next section, we use these results to return to our driving questions, discussing how knowledge workers go about dividing and structuring work.

6 Discussion

In this section, we discuss how our participants appropriated the Activity construct, propose three strategies users employ to structure Activities, and identify emergent behaviors in their use of the system. We also highlight implications for design of ACC systems.

6.1 What Is an Activity?

As a unit of analysis, an activity is defined as a coordinated set of actions by people towards a common objective, mediated by tools and subject to situational constraints (e.g. [6]). But in work settings, knowledge workers have objectives at different levels of granularity, some objectives being in service of other objectives (e.g. creating a demo as part of preparing for a review as part of managing a project). However, the Lotus Activities system does not represent in-service-of relationships between Activity constructs. Thus, the empirical question is: what granularity of objectives do people associate with Activity constructs. Our overall finding is that the Activity construct is used to support objectives at various levels of granularity. For some usage patterns, such as Creating a Document, Planning an Event, or Solving a Problem, Activity constructs clearly correspond to concrete, achievable objectives.

On the other hand, the Work Units pattern reveals a single objective of managing all the work represented with multiple Activity constructs. Still other patterns, like Collecting Information or Exchanging Formalized Information, are represented as Activity constructs even though they are in service of larger objectives. Overview patterns—managing a Team Space or Personal Tasks—bring together the work of many objectives within a single Activity construct. Finally, multiple usage patterns may occur in the same Activity construct, which may evolve from one pattern to another. Thus, the analytic concept of an activity does not map in a simple way to the way people organize their work in an ACC system. Our findings identifying these patterns show that active users of Lotus Activities *were* able to identify effective divisions of their work in an ACC system, which confirms the value of a flexible, generic Activity construct. However, P13 emphasized that "the thinking is the startup cost" for appropriating Lotus Activities, and this cost forms the main barrier for new users. We believe that ACC system designers could aid novice users by providing explicit support for the common patterns identified in this study.

6.2 Structuring Activities

We have identified some common patterns of use for Lotus Activities, but users often structured their instances of these patterns in different ways. This presents a challenge to designers trying to provide guidance for defining Activities. However, users followed a common process in structuring their Activities: (1) they seeded structure from previous tools, (2) they evolved existing Activities, and (3) they reused successful ones. Understanding this process leads to design implications for ACC systems.

1. Seeding. Users seek out new tools when current ones fail in one way or another. Thus a user seeking to structure a piece of work in an ACC system is often transferring that work from another tool. We saw that users leveraged the structure of their old tools to seed new Activities. This strategy may explain why Work Units in a similar domain had different structures—users were migrating from *different tools*. P11 used an "account" as a Work Unit Activity to track sales opportunities, because this was how she previously structured her file and email folders. On the other hand, P12 transitioned from internal documentation forms that needed to be filled out for each customer-product pair, so his sales opportunity Work Unit Activities mirrored this structure. Users continued using company-mandated, team-specific, and generic tools in parallel with ACC, so it is critical for the designer to integrate ways to transition data between the previous tools used for the activity and Activity constructs of an ACC system.

2. Evolving. While initial Activity structure may be suggested by a prior tool, the Activity may still need significant modification to address a user's needs, requiring incremental evolution over time. For example, P8's team wanted to use Lotus Activities to document work requests in more detail than the spreadsheet they currently used (see Figure 2). They structured each row of the spreadsheet (each request) as a new Activity with the spreadsheet field data (project name, status, and target team) becoming the Activity title. The Activity started out with no entries. The team iterated multiple times on the internal structure, adding entries prompting for certain information and organizing them into sections. We were able to see this process by looking at the older versions of the work requests. Since evolution and reuse of activity-related documentation is a common practice [13], supporting this process of “incremental formalization” [14] can be one of the greatest benefits of an ACC system. To do so, ACC systems should provide effective tools for reorganizing existing Activities. Additionally, Activity Templates must be able to selectively propagate changes to existing Activity instances in order to support incremental evolution.

3. Reusing. It may take significant effort for a user to achieve a successful structure for a given Activity; but once it is identified, it can be easily reused. This explains why our users were likely to have many Activities with the same structure (i.e., Work Unit Activities). Tracking status across sets of Activities currently presents a challenge for Lotus Activities users, but designers can leverage the structured and repeated nature of these Activities to create overview perspectives, e.g., as visualizations, filtered or aggregated views, or overview Activities.

6.3 Emergent Behavior

Users exhibited emergent behaviors to realize unexpected benefits of using an ACC system: sharing knowledge and reflecting on work. In identifying these emergent behaviors, we suggest features to enhance the sharing and reflection benefits.

Activity constructs represent records of how activities were *carried out,* which can be educational for others. Lotus Activities lowered the overhead of creating sharable content, since it was created in the process of doing work. Unlike workflow systems, ACC systems can provide rich accounts of work as it was actually accomplished rather than as it was prescribed [15]. Among our participants, successful Activities were used to train new members, transition work between employees, and share ideas with other teams. However, we also saw that currently users could learn about these resources only through explicit advertising efforts (e.g., posting the link to a blog, mentioning in a presentation, etc.). To capitalize on knowledge sharing and work reuse opportunities, an ACC system should increase “discoverability” by providing comprehensive search facilities and by highlighting successful tutorials, e.g. with social software techniques such as ratings.

Some of our users found that the act of creating content in the ACC system helped them reflect on their work. Six users said that capturing work as they did it made reporting current status and accounting for their time easier. ACC systems should help users create accounts of their time and work progress for reflection or reporting purposes. Four users found that combining information from multiple Activities was a way to reflect on content that assisted them in identifying solutions to problems. This kind of improvisation is consistent with practices of other successful knowledge workers [16]. To support this behavior, reorganizing content across many Activities should be facilitated.

7 Conclusion

For any ACC system user, developing effective ways to structure Activities will be a challenge. The usage patterns, strategies employed, and challenges faced by active users have revealed some ways to mitigate this adoption barrier in ACC systems. First, the designer could aid novice users by providing explicit support for the common patterns identified in this study. Second, the designer could aid in the transition to the ACC system by suggesting ready-made structures similar to those present in the existing tools favored by the user for targeted activities. Third, the designer could make it easier to continually improve and reuse the structure of Activities by supporting lightweight reorganization and by allowing selective propagation of Template changes to instantiated Activities.

We have shown that ACC systems have the potential to support knowledge workers by providing a shared Activity constructs for coordinating work units, producing deliverable outcomes, gathering information, and sharing knowledge. By investigating both the successful usage patterns and the challenges users faced while structuring their work, we believe that future ACC systems, as well as future versions of Lotus Activities, will overcome one of the major adoption barriers for ACC systems: helping users understand what an Activity is and how it can help them be more effective workers.

Acknowledgments. We thank: James Lin, Jeffrey Pierce, and Aruna Balakrishnan for advice during this research; the Lotus Activities team—Miguel Estrada, David Brooks, Marty Moore, Bob Stachel, Suzanna Minassian, and Scott Prager—for ongoing support; and Michael Muller and Gregory Abowd for comments on drafts of this paper.

References

[1] Bardram, J., Bunde-Pedersen, J., Soegaard, M.: Support for activity-based computing in a personal computing operating system. In: Proc. of CHI, pp. 211–220 (2006)

[2] González, V.M., Mark, G.: Constant, constant, multi-tasking craziness: managing multiple working spheres. In: Proc. of CHI, pp. 113–120 (2004)

[3] Muller, M.J., Geyer, W., Brownholtz, B., Wilcox, E., Millen, D.R.: One-hundred days in an activity-centric collaboration environment based on shared objects. In: Proc. of CHI, pp. 375–382 (2004)

[4] Christensen, H.B., Bardram, J.: Supporting Human Activities - Exploring Activity-Centered Computing. In: Proc. of UbiComp 2002, pp. 107–116 (2002)

[5] Kaptelinin, V.: UMEA: translating interaction histories into project contexts. In: Proc. of CHI 2003, pp. 353–360 (2003)

[6] Kuuti, K.: The concept of activity as a basic unit of analysis for CSCW research. In: Proc. of ECSCW, pp. 249–264 (1991)

[7] Dourish, P.: The Appropriation of Interactive Technologies: Some Lessons from Placeless Documents. In: Proc. of CSCW, pp. 465–490 (2003)

[8] Voida, S., Mynatt, E.D.: It feels better than filing: everyday work experiences in an activity-based computing system. In: Proc. of CHI 2009, pp. 259–268 (2009)

[9] Bellotti, V., Thornton, J., Chin, A., Schiano, D., Good, N.: TV-ACTA: embedding an activity-centered interface for task management in email. In: Proc. of Conference on Email and Anti-Spam (2007)

[10] Muller, M.J., Minassian, S.O., Geyer, W., Millen, D.R., Brownholtz, E., Wilcox, E.: Studying appropriation in activity-centric collaboration. International Reports on Socio-Informatics 2, 50–58 (2005)
[11] Balakrishnan, A., Matthews, T., Moran, T.P.: Adopting an Activity-Centric System in an Ecology of Workplace Tools. IBM Almaden Research Center (2009) (Working Paper)
[12] Seidman, I.: Interviewing as Qualitative Research: A Guide for Researchers in Education And the Social Sciences. Teachers College Press, New York (1998)
[13] Moran, T.P., Matthews, T., Vega, L., Smith, B., Lin, J., Dill, S.: Ownership & Evolution of Local Process Representations. In: Proc. of INTERACT (to appear, 2009)
[14] Shipman, F.M., McCall, R.: Supporting knowledge-base evolution with incremental formalization. In: Proc. of CHI 1994, pp. 285–291 (1994)
[15] Dourish, P.: Process descriptions as organisational accounting devices: the dual use of workflow technologies. In: Proc. of GROUP, pp. 52–60 (2001)
[16] Schön, D.A.: The Reflective Practitioner. Basic Books (1983)

Sharing Usability Problem Sets within and between Groups

Gudmundur Freyr Jonasson and Ebba Thora Hvannberg

University of Iceland, Hjardarhaga 2-6, 107, Reykjavik, Iceland
gfj@hi.is, ebba@hi.is

Abstract. Merging similar usability problems is often a time consuming step in the process of usability evaluation. We have developed a tool which can aid with merging duplicate problem descriptions. The tool enables evaluators to define constructs which describe a usability problem and store them in a database. The evaluator can search for existing constructs, problem descriptions and problem sets using the query language XQuery. The XQuery search function also enables the evaluator to compare usability problems. The tool is designed to be flexible and the exchange of data between evaluators is simple exploiting the compatibility of XML (Extensible Markup Language) and a well defined data structure. We have performed an analysis on a problem set to check for the comparison quality of the merging function in the tool. This has allowed us to suggest further ways to compare usability problems.

Keywords: Usability, problem, construct definition, merging, exchange, relevance.

1 Introduction

Merging similar usability problems is often a time consuming step in the process of usability evaluation. The complexity lies in the data which comes from different users and also that evaluations of the same system may be done at different places or times and sometimes with different constructs [1]. Usability problem sets are compared not only for practical reasons but also for research by meta-analysis as several examples show [2, 3]. Our goal is to design and implement a tool which enables usability testers to record usability problems into a database using different usability evaluation methods, and to allow evaluators to search the database for similar problems, exchange data sets and to perform a meta-analysis of data sets. By using the tool the analysis following user tests should be less time consuming and the exchange of data between evaluators made easier. The tool uses methods to diagnose what problem descriptions are unique and what descriptions are duplicates by comparing constructs within problem descriptions. If there is a match according to a predefined relevance scale the descriptions are grouped together. The user then makes the final call whether or not the two problem descriptions are duplicates. If the user decides that the descriptions are duplicates then those two problem descriptions can be merged into one description.

The tool is designed to be flexible with the use of XML. Therefore, data exchange between different users is easy, given the good compatibility of XML. The

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 596–599, 2009.

tool runs on internet browsers and connects to an XML database called eXist (http://exist.sourceforge.net/) and uses XQuery to search the database.

2 Requirements and Implementation of the Tool

In this section we will describe three main use cases of the tool: defining a construct to describe a usability problem, searching the problem database and matching usability problems. First, the tool allows the user to flexibly define constructs to describe usability problems. At the onset of recording usability problems from an evaluation in the tool, the user can select a set of constructs already defined in the tool, and add new constructs. To enable exchange of data sets and meta-analysis of data sets, it is necessary to record meta-data for individual constructs. Cockton [4] has suggested a coding dimension format which includes: Name of construct, Definition, Operationalization (explains what the coder should do), Evidence Requirements (describes how to code values/levels), Sources of Bias (describes possible flaws in coding), Motivation and Relevance (explains why this construct matters), Used in (references) and Examples (from existing case studies).

Second, after having recorded a set of usability problems, the user can search for problems with the use of XQueries. Given the power of the XQuery language, users can make very precise searches on the whole database, within usability problems of a particular usability group or usability problems recorded by single authors. Queries can be stored and recalled in the tool, allowing the user to use common queries, such as 'Find all problems of highest severity level'. Thus expert users are provided the power to enter any query but novice users of the tool, can apply predefined queries.

Third, the tool, upon request, notifies the user of similar usability problem descriptions. The similarity is based on relevance, defined by a set of constructs, which can be defined by the user. The tool compares constructs of problems and if they satisfy a given relevance the tool groups those problem descriptions together and displays the results to the user. For example, a user defined relevance-scale TEF can be defined for the variables Trigger, Expected Phase and Failure Qualifier. Trigger describes what evaluators are doing when he/she discovers the problem, Expected Phase indicates in which phase of a software development life cycle the developer is expected to be able to fix a usability problem and Failure Qualifier which is used to capture more meaning about the usability problem. If, say, four usability problem descriptions in a given set have exactly the same values in those three constructs, they are grouped together and the user makes the final call by evaluating if the descriptions are describing the same problem or not. Problems which the user has acknowledged to be the same are marked as such in the database by putting them in the same equivalence group. The equivalent groups, i.e. the mergeable problems, appear as output in alternate colours, grey or white.

3 Quality Analysis of Merging

A quality of a merging tool is rated by how accurately it can identify mergeable problems. More specifically we want to find out whether the merging tool could correctly identify set of problems which a human agrees are mergeable. There are several possibilities: a) The tool suggests a match between problems but a human not b) the tool misses a match, which a human has identified c) the tool and the human agree that there is a match. In this pilot research study, we addressed options a) and c), labeling the former a ‘False match’ and the latter ‘Mergable’. In order to inspect the quality of the merging tool, we examined a problem set resulting from a usability evaluation of a learning management system (OWL). The problem set has 71 problems, which have already been consolidated, and thus the problems in the set have been filtered and merged by a usability expert. Despite this earlier merging, the merging tool suggested seven groups of usability problems. We analyzed the quality of the tool by rating the tool’s suggestions either ‘False match or ‘Mergeable’ in column two of Table 1.

Three merging suggestions of seven were correct. This analysis shows that the tool found opportunities for merging problems, even though they had been consolidated earlier manually. Several problems became evident during the merging exercise. First, some of the constructs, i.e. Context has too many values. This can be remedied by suggesting ontology of values within a variable by creating a few subsets of values, where the values within each subset are equal. Second, some of the constructs are textual and hence a text matching algorithm needs to be implemented. We have not yet analyzed the remainder of the problems in the OWL set to see if any problems were missed by the merging function in the tool, i.e. option b) above.

Table 1. Outcomes of matches using the relevance scale TEF

Matched Group ID	Rating	Matched UPs	Similarities
1	False match	1 and 27	The UPs describe two different use cases.
2	False match	14 and 61	
3	Mergeable	16, 56, 57 and 64	UP 16 and 56 could be merged.
4	Mergeable	20 and 40	Both users are trying to change previously stored text.
5	False match	21 and 59	UP 59 should be put in group 4
6	False match	26 and 66	
7	Mergeable	32 and 45	Both users are having difficulties identifying hyperlinks.

4 Future Work

The development and analysis tool described in this paper marks a first step to reach the goal of providing evaluators with an automatic technique to merge problem sets and allow for their exchange. Although the tool itself is quite usable, immediate next steps will include easier access to search functions through an enhanced user interface and import/export functionality from and to Excel. In this paper, we have focused on one of three use cases, i.e. matching usability problems. As we mentioned above, further work is needed to understand the automatic matching of usability problems, through the recoding of variables which have a large range of values such as the variable Context or free text variables such as Description. It will also be interesting to address textual comparison further with advanced algorithms such as text mining or latent semantic analysis. Other use cases await further studies, such as exploring the tool's usefulness for exchanging problem sets with different construct definitions between evaluators. As future work it will be interesting to explore how relationships between the constructs definitions can be made by expressing that a construct can be derived from another; a construct can be equal to another; a construct can be an aggregation of several other constructs, a construct can determine another etc. Thus, the tool can also have a potential for researchers in performing a meta-analysis of problem sets.

Acknowledgement. This work is in part supported by the Icelandic National Research Fund, RANNÍS. We thank the reviewers for their useful comments.

References

1. Law, E.L.-C., Hvannberg, E.T.: Consolidating usability problems with novice evaluators. In: Proceedings of the 5th Nordic conference on Human-computer interaction, Lund, Sweden, pp. 495–498. ACM Press, New York (2008)
2. Hornbæk, K., Law, E.L.-C.: Meta-analysis of correlations among usability measures, Conference on Human Factors in Computing Systems. In: Proceedings of the SIGCHI conference on Human Factors in computing systems, pp. 617–626. ACM Press, New York (2007)
3. Vermeeren, A.P.O.S., Attema, J., Akar, E., Ridder, H., de, D.A.J., van, E.C., Berkman, A.E., Maguire, M.C.: Usability Problem Reports for Comparative Studies: Consistency and Inspectability. Human-Computer Interaction 23(4), 329–380 (2008)
4. Cockton, G.: MAUSE. COST294-MAUSE workshop on coding constructs definitions and coding problem sets, Salzburg, Austria (June 7, 2007), `http://www.cost294.org`

Obstacles to Option Setting: Initial Results with a Heuristic Walkthrough Method*

Silvia Gabrielli and Anthony Jameson

FBK-irst, Trento, Italy
{sgabrielli,jameson}@fbk.eu

Abstract. This short paper is the first step in a line of research that aims to deepen understanding of the difficulties that users often have with *option setting* interfaces: those parts of a system that allow the user to set parameters that influence the system's behavior and appearance. On the basis of a theoretical distinction of three things that users may fail to understand about a given option, we introduce a simple variant of the heuristic walkthrough method that helps evaluators to uncover likely obstacles. We give a quantitative and qualitative overview of the obstacles found through the application of this heuristic walkthrough to parts of four popular applications.

1 Option Setting Interfaces as the Stepchild of Interface Design

With regard to most aspects of mainstream graphical user interfaces and web-based systems, we have an abundance of design guidelines, along with convenient methods for inspection-based evaluation (e.g., heuristic evaluation and cognitive walkthrough). But one common part of these systems tends to remain in the shadows: those screens full of "options" (or "preferences") that each user can set to make the system work in a way that he or she will find satisfying in the long run. Everyday experience and conventional wisdom suggest that many users find it unduly difficult and/or time-consuming to figure out suitable option settings for themselves; and several studies conducted in the 1990s (e.g., [5, 7]) already illustrated this general trend. These difficulties in turn constitute one of the motivations for interfaces that automatically adapt to their users (see, e.g., [4, 7]). What is still largely missing are constructive efforts to help designers to design better option setting interfaces, anticipating and minimizing the various obstacles that users can encounter.

An exception is found in research on privacy and security options and preferences (see, e.g., [1, 2, 3]), where concerns regarding user lack of support/guidance in option setting have been more actively debated. Our ultimate goal is to extend the relatively high level of understanding that has been achieved in this area to other types of option setting.

* The research described in this paper was conducted in the targeted research unit Prevolution (code PsychMM), which is funded by the Autonomous Province of Trento.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 600–603, 2009.

2 Three Types of Obstacle

First, on a relatively theoretical level, we distinguish three types of obstacle that a user may encounter when deciding how to set a particular option:

1. *Meaning unclear:* The user may simply not understand the description of the option and the possible settings. For example, Mozilla Firefox offers three possible actions that the system can take when Firefox starts: *Show my home page*, *Show a blank page*, and *Show my windows and tabs from last time*. A (novice) user may not know whether the first alternative refers to their personal web home page or to some other "home page" (as is in fact the case).

2. *Consequences unclear:* Even if the user understands the meaning of all of the alternatives, they may not be aware of the consequences of each possible setting – for example, the fact that showing a blank page at startup can be much faster than the other two alternatives, especially when the user is not connected to the internet.

3. *Best choice unclear:* Even if the user understands the meaning and the consequences, it may be hard for them to judge which alternative will be best for them in the long run, especially if the system will be used in various different situations with which the user is not yet familiar. Where the best choice is not easy to see, the option setting interface may offer explicit advice (e.g., "recommended if you often work off-line").

Note that each of these three types of understanding does not necessarily presuppose the preceding type(s) in the list.

3 A Variant of Heuristic Walkthrough and Its Initial Application

To help designers and researchers think more systematically about option setting interfaces, we devised a simple variant of the heuristic walkthrough (HW) method (cf. [6]) that makes it less focused on analyzing task sequences: For each option in a given option setting screen of a given system and for each of the three aspects of clarity just discussed, our questionnaire asks the evaluator for a judgment of the likely clarity of that option for a typical user, on a scale with the values "Not clear", "Partially Clear", and "Clear". The same judgment is then also requested under the assumption that the user has read the associated online help.

Three HCI researchers (on the PhD and post-doctoral levels) were asked to apply the HW method sketched above, thinking of a novice user and looking at two categories of options: *General* and *Privacy* options. They analyzed a total of 79 options in 4 popular applications: Firefox 3.0.7, Skype 3.0.8.180, MS Office Outlook Web Access 2007, and the LinkedIn professional network (for registered users). They were given a worksheet that listed the 79 options and provided the coding scale. The time to complete the HW for all 4 systems was approximately 1 hour for each rater.

4 Results

Agreement among evaluators: For each question about each option, the "range" of the three evaluators' ratings was computed (i.e., the number of steps of the scale separating the highest and lowest ratings; possible values being 0, 1, and 2). The mean

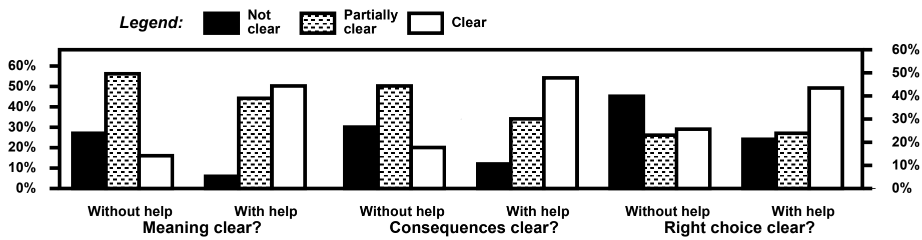


Fig. 1. Frequency distribution of the ratings for each of the three main questions of the Heuristic Walkthrough

ranges for questions concerning meaning, consequences, and overall choice were 0.91, 1.02, and 1.22, respectively. (Since overall no systematic differences were found between the ratings of the options in the General vs. the Privacy categories, these categories will not be distinguished in the following.)

Meaning of options and alternatives: The first group of three bars in Figure 1 shows the relative frequencies of the three possible clarity ratings concerning the meaning of options, assuming that on-line help is not consulted. Only 17% of these ratings were “Clear”. Often, difficulties arise when an option’s description includes technical jargon. For example, in Firefox’s privacy settings, a novice might not understand the term *cookie* in *Accept cookies from third parties*. But as the second group of three bars in Figure 1 shows, clarity is seen to be enhanced considerably by the on-line help, with “Clear” now becoming the most frequent rating. In particular, Firefox’s help includes extensive information about cookies.

Consequences: As can be seen in the third group of three bars in Figure 1, the distribution of clarity ratings concerning the consequences of an option is roughly similar to that for the ratings of meaning clarity. Especially consequences that may occur only in the long term – such as those of allowing one’s *on-line status to be shown on the web* (in Skype) were considered likely not to be appreciated. Here again, consultation of on-line help was judged to have a large impact on the recognition of relevant consequences (cf. the fourth set of three bars in Figure 1), though even with help about half of the options were considered not to be “Clear”.

Choice: Not surprisingly, the evaluators saw the greatest potential for uncertainty when it came to actually making a choice about an option. For example, with the *message tracking* options of Outlook Webmail, it is not hard to understand the possible ways of dealing with requests for automatic responses (*ask me before sending*, *always send,* or *never send*). But judging which of these choices will be best in terms of the overall costs and benefits is trickier. Here again, the on-line help was seen as eliminating a good deal of the uncertainty; but still almost a quarter of the ratings are “Not clear”. For example, the Outlook help about message tracking offers no advice as to the user attributes or situations that tend to be associated with particular choices. The help system writers should not necessarily be faulted: If the considerations that can influence the choice of a given setting are numerous and perhaps highly specific, there may be no alternative to letting users arrive at their decisions on their own, maybe after some trial and error. These cases remind us that the obstacles to option setting do not consist solely in a lack of provided information or in the inherent difficulty of technical concepts: The diversity of people and situations may be the most challenging obstacle for users and designers alike.

5 Conclusions and Next Steps

The initial application of our heuristic walkthrough for option setting interfaces has confirmed that it is helpful to distinguish between three types of difficulty that a user can have in deciding how to set a given option: lack of clarity about meaning, consequences, and overall choice, respectively. Although these types of clarity are interrelated in various ways, they raise different challenges for users and designers. And although all three types of clarity do not need to be present for any given option, they are all worth bearing in mind when a solution is crafted for any given option.

The fact that "help helps" (at least when it is consulted, and according to the assessments of our raters) is encouraging in some ways, but it does suggest that attention should be devoted to the question of how users can be supported in their decisions as to whether it is worthwhile for them to consult the help on a given option.

Natural next steps in this research include (a) comparing the ratings made on our heuristic walkthrough to the responses of actual users; (b) formulation and testing of guidelines for option descriptions (and the associated on-line help); and (c) expansion of the scope of the analysis to include the process of trial and error that users often rely on in addition to the information offered in the interface and the help system.

References

1. Ackerman, M.S., Cranor, L.F., Reagle, J.: Privacy in e-commerce: Examining user scenarios and privacy preferences. In: ACM Conference on Electronic Commerce, pp. 1–8 (1999)
2. Furnell, S.M.: Making security usable: Are things improving? Computers & Security 26(6), 434–443 (2007)
3. Furnell, S.M., Jusoh, A., Katsabas, D.: The challenges of understanding and using security: A survey of end-users. Computers & Security 25(1), 27–35 (2006)
4. Jameson, A.: Adaptive interfaces and agents. In: Sears, A., Jacko, J.A. (eds.) The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications, 2nd edn., pp. 433–458. Erlbaum, Mahwah (2008)
5. Page, S.R., Johnsgard, T.J., Albert, U., Allen, C.D.: User customization of a word processor. In: Proceedings of CHI 1996, pp. 340–346 (1996)
6. Sears, A.: Heuristic walkthroughs: Finding the problems without the noise. International Journal of Human-Computer Interaction 9(3), 213–234 (1997)
7. Trewin, S.: Configuration agents, control and privacy. In: Proceedings of the ACM Conference on Universal Usability, Arlington, VA, pp. 9–16 (2000)

Dimensions of Context Affecting User Experience in Mobile Work

Heli Wigelius and Heli Väätäjä

Tampere University of Technology, Human-Centered Technology
P.O. Box 589, 33101 Tampere, Finland
{heli.wigelius,heli.vaataja}@tut.fi

Abstract. Understanding the contextual factors affecting user experience is essential in designing and evaluating mobile systems for mobile work. The aim of this paper is to explore these contextual factors through three case studies: of safety observation at construction sites, passenger transportation with taxis, and mobile news journalism. For each case study we describe the nature of the mobile work and present our findings on the contextual factors that were found to affect the user experience. Based on the results, we present and discuss five dimensions of mobile work context affecting user experience: 1) social, 2) spatial, 3) temporal, 4) infrastructural, and 5) task context. Compared to earlier frameworks of context for mobile work, the social context as well as the infrastructural context was emphasized in our findings. The presented framework elaborates the dimensions of context affecting user experience of mobile systems and services in mobile work in particular. The framework is also applicable for mobile consumer systems and services.

Keywords: context, mobile work, user experience, mobile systems.

1 Introduction

The importance of understanding context has been highlighted in many studies in the field of Human-Computer Interaction (HCI) [1], [2], [3], [4]. When a new technology, for example a mobile system, becomes a part of traditional work tasks it is essential to ensure that work activities are appropriately supported [5]. Furthermore, several user experience frameworks and definitions mention context as one of the components affecting user experience [6], [7]. [8]. If the context has not been identified and understood, the use of the mobile systems may fail or the user experience may suffer. However, empirical research on user experience has mainly concentrated on consumer products and services, with only a few focusing on the mobile work context [4], [9], [10], [11], [12].

In this paper, we present our findings from three exploratory case studies on the contextual factors affecting user experience in mobile work. We focused on mobile systems that support the work of mobile workers. The case studies were of safety observation at construction sites, passenger transportation, and mobile news journalism. The data was gathered via observations and interviews. Through these case studies, we examine the nature of mobile work and present examples of identified contextual

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 604–617, 2009.

factors affecting the user experience. Based on our findings we then present an elaborated framework for the dimensions of these contextual factors.

The rest of the paper is structured as follows. First, we review earlier research on context from ubiquitous computing, user experience and mobile work. We continue by describing the research methods and by presenting the results regarding the contextual factors found to affect user experience in the cases studied. We then discuss and summarize the results by presenting a framework for context in mobile work and describe the identified dimensions.

2 Related Work

The concept of context has been discussed extensively in the areas of Computer-Supported Cooperative Work (CSCW) [13], ubiquitous and context-aware computing [14], [15], [16], and user experience [17] as well as in mobile work [4], [18]. A number of attempts have been made to define context [14], including a standardized definition for the context of use [19]. As Dey states, context is a word that is understood easily but is hard to define clearly [15]. One of the most recent definitions for context is proposed by Bradley and Dunlop [14] for context-aware computing. They propose a dynamic model of context based on the definitions taken from linguistics, computer science and psychology. According to them, the user's contextual world comprises the task, physical, social, temporal, and cognitive context.

Context is also an essential part of the interaction oriented frameworks of user experience. In these frameworks user experience is defined as a consequence of the user's interaction with the product, system or service, which is affected by characteristics of the user and the system as well as by contextual factors [6], [7], [8]. In the case of mobile browsing, Roto [17] identifies four dimensions of mobile context affecting the user experience, namely physical context, social context, temporal context and task context. Roto [17] describes the physical context as being any circumstances that physically affect the use of the mobile browser, such as weather conditions, light, crowds, and noise. Social context refers to the expectations of other people in regard to the user in the present context, such as in a meeting. Temporal context refers to the time period that the user dedicates to the system, but limited by the restrictions of the context, such as a commuter interrupting their usage as they get on and off a bus. Roto [17] uses the term task context to refer to the user's higher-level goal, where mobile browsing is merely one task in accomplishing the goal. She gives the example where a user is going home by a bus and checking the timetables with the mobile browser, but their high level goal is to get home.

Several definitions have been presented for both mobile context [20], [21] and mobile work context [4], [18], as shown in Table 1. For example, the definition of mobile work context given by Zheng and Yuan [4] consists of mobile context (where and when), mobile workers (who), mobile technologies (how), and mobile tasks (what); a definition also supported by other literature (See Table 1). Furthermore, Turel [22] includes users' motivation in his definition of context, which is mentioned as one of the user characteristics in frameworks for user experience [8], [17].

In our study we follow the approach of the interaction oriented definitions of user experience, where user, system and context are the basic elements affecting the user

Table 1. Elaborations of context in the literature

Authors	Viewpoint	Elaborations of context
Bradley and Dunlop [14]	Context-aware computing	Task context, physical context, social context, temporal context, cognitive context
Forlizzi and Ford [6]	User experience	Social, cultural and organizational factors
Hassenzahl and Tractinsky [7]	User experience	Organizational setting, social setting, meaningfulness of activity, voluntariness of use
Kankainen [8]	User experience	People, places and things surrounding user
Lee et al. [21]	Consumer	Personal (emotion, time, movement), environmental (physical, social)
Roto [17]	User experience	Physical context, social context, temporal context, task context
Turel [22]	Mobile work	Users, time, physical environment (incl. technology and organization), motivations, tasks
Wiberg and Ljungberg [18]	Mobile work	Time, place, tasks
Zheng and Yuan [4]	Mobile work	Mobile context, mobile workers, mobile technologies, mobile tasks

experience. Therefore we separate the characteristics of the user and system from the context. We also follow the currently proposed definition in ISO standardization [23], which defines user experience as "*a person's perceptions and responses that result from the use or anticipated use of a product, system or service*." Contextual factors may significantly contribute to the user's perceptions, preferences, behaviors, accomplishments and even emotional responses to using a mobile system or service.

Table 1 summarizes the previously discussed elaborations of context also from user experience and context-aware computing. We can see that definitions of context in different fields have similar components, especially those related to time and place. However, tasks as a contextual component are more strongly emphasized in the mobile work context. Moreover, social context is included in the user experience literature as well as that of context-aware computing, whereas it is missing from the frameworks of context in mobile work related literature. In the case of mobile work, technologies are also mentioned as a part of the context.

User experience has received considerable attention in consumer products and services research. However, there is a lack of empirical research on user experience when using mobile systems for mobile work. The aim of this paper is to contribute to the body of knowledge on factors affecting user experience in mobile work by focusing specifically on the contextual factors. The findings were categorized to form a framework for general dimensions of context in mobile work. Such frameworks have the benefit of making designers, developers as well as user experience experts aware of, and more easily able to identify and evaluate, the contextual factors affecting user experience.

3 Research Methods

The focus of this paper is to explore contextual factors affecting user experience in mobile work and how they can be categorized. A multiple case study approach was chosen as the research strategy [24] in order to gain more generalizable results. Three cases were studied separately between 2006 and 2008 in Finland, of safety observation at construction sites, passenger transportation, and mobile news journalism (see for example [25], [26]). The cases reported here are a part of two larger research projects.

We used two sources for qualitative data in the safety observation and mobile news journalism case studies, namely observations and semi-structured interviews. By observation we were able to study what kind of contextual factors existed in real usage situations and to gain an understanding of how context affects the user experience. Because in the case of passenger transportation the nature of their work meant the taxi drivers could not be observed, focus groups were arranged instead. Semi-structured interviews were conducted to complete the results of the observations and to gain a deeper understanding of the factors affecting user experience. Notes were written during the observations, while the interviews and focus group discussions were recorded and transcribed.

Each case study was first analyzed separately using content analysis and applying both data triangulation and investigator triangulation. The transcribed observation notes and interviews were first read through to form a general understanding of the data in relation to contextual factors. Inductive reasoning was used in giving descriptive labels and assigning them to lines, sentences or paragraphs [27]. These labels were then grouped to higher level categories. We used cross-case analysis [24] and earlier frameworks of context from ubiquitous computing, user experience, and mobile work related literature to form the main dimensions of context from the categories.

3.1 Case Study A: Safety Observation at Construction Sites

At construction sites, we focused on the safety observation process known as TR-measurement [28]. The process is carried out weekly by the industrial safety delegate and the site supervisor on Finnish construction sites and is aimed at reducing the number of industrial accidents.

Two researchers observed the construction workers for two hours at three different construction sites. Altogether there were six participants aged 25 – 60 years old in three observations. They had used the mobile service from between five to 12 months. At the time of the study, they used the mobile service via their mobile phones (Nokia 6630). After the observations, the participants were interviewed for background information and to go over the observation notes.

3.2 Case Study B: Passenger Transportation

In this case study we focused on the work of taxi drivers, where mobile systems play a crucial role in major work tasks such as accepting incoming requests, searching for locations on the map and communicating with dispatch and with other taxis. Installed in the taxis were mobile systems including General Packet Radio System (GPRS) and a touch screen, along with other devices such as a taximeter and an additional touchpad for operating the system.

Two two-hour long focus groups were arranged with taxi drivers. The first group's five participants (one woman, four men) had an average of 20 years driving experience. The second focus group had four participants (all men) whose taxi driving experience varied from 1.5 to 10 years. The focus groups covered questions related to mobile work tasks, and the benefits and usability of the mobile system, as well as possible problems faced while using it.

3.3 Case Study C: Mobile News Journalism

In this case study, ten graduate students of journalism and nine graduate students of visual journalism published an online blog of university events as part of a university course project. Uploading to the blog was done wirelessly directly from the proximity of the reported events. Students used the mobile phone based system for capturing multimedia items, and creating news stories with text and multimedia. The main parts of the mobile journalism system were a mobile multimedia phone (Nokia N82) for photo and video capture, a wireless Bluetooth keyboard (Nokia SU-8W) for writing text, and a research prototype of a mobile software application developed for mobile news creation and submission. A publishing schedule was given by the producer of the publication, replicating the working process of an editorial staff in a newspaper organization.

Three researchers observed eight participants using the mobile system during the first project day, some 24 hours of observations of the mobile work context. Five participants were interviewed before they used the system so as to gather background information, for example, on their earlier experiences of using mobile phones in journalism. Eleven participants were interviewed after the project day to gather information on their experiences including context related aspects. Additional observations of three journalist-photographer pairs and four interviews were made on a second project day a month later.

4 Results

In this section we present the results of each case study. First, we characterize the mobile work in each case, as well as the work community in general. We then describe the contextual factors that were emphasized in the cases.

4.1 Case Study A: Contextual Findings in Safety Observation

Work at construction sites is traditionally mobile; workers move in and out of the building under construction. The safety observation is also mobile as an industrial safety delegate and a site supervisor walk through the site. The industrial safety delegate makes his safety related findings during the safety observation round and the accompanying supervisor documents them, either traditionally by filling out a specific paper form, or by using a dedicated mobile service developed for this purpose. Photos, complementing the written descriptions of the safety issues, are taken either with the mobile phone or with a separate digital camera. The results of the safety observation round are presented to contractors at a weekly site meeting.

The construction sites of course vary according to the specific building under construction, the area of the site, the number of workers on the site at any time, and the overall progress of the project. The progress of the construction affects how long the observation round takes and the volume of the findings. In other words, when the construction is at an early stage, the observation round takes about an hour, and there are less than 100 findings; at a later stage, the round can take about two hours and may have up to 500 findings.

Based on our findings with construction workers, infrastructural factors such as functionality of the network connections and the capacity of the mobile device were highlighted. If the building has floors below ground level, such as a cellar, the network connection may be limited or even unavailable at times. However, since the mobile service used for the safety observation sent data to the server in real time, unavailable or cut network connections got the service stuck and the data was not uploaded. In the worst case, the user had to reboot the mobile device and all the gathered data was lost. In that case, the user had two choices in continuing the safety observation: either they tried to remember all the findings already made, or they started again. This took extra time and the user became frustrated. In addition, the risk of making errors increases when the user has to remember all the findings.

Another infrastructural issue highlighted was the capacity of the mobile device due to the working situation. When the observation round takes approximately two hours, the mobile phone battery must be fully charged before the round begins. Environmental factors, such as cold and humidity, were also a factor on construction sites, since although winter weather did not complicate a safety observation round conducted with pen and paper, it did complicate the use of the mobile service. For example, entering data with gloves on is difficult. To quote one user: *"The less data is required to be entered when using the mobile service, the better"*. One of the users had cut the thumb off his gloves to better use the mobile service in winter time. Humidity in contrast, complicated observation rounds conducted with pen and paper, but not with a mobile service, since rain makes paper documentation difficult to write and read, whereas the mobile device was easily placed in a pocket.

4.2 Case Study B: Contextual Findings in Passenger Transportation

Taxi drivers' main task is to drive customers from one place to another throughout the day and night, a very customer-orientated service. They accept incoming requests and search for locations on the map with their mobile system. Taxi drivers work in the middle of the traffic with their peak hours being nights and weekends, whereas during the week and the daytime it is quieter. In Finland the taxi owners usually hire one or more drivers. Both the owners and the hired drivers drive the cars, so the users of the mobile system in any given car may be very different. There are, for example, older drivers that have driven taxis for years, and have used older technologies for their work tasks. Younger drivers are usually familiar with computers and new technology and are more at ease using new systems.

We found social factors to be important for taxi drivers. Taxi drivers see each other, for example, when they are waiting for possible customers at a railway station, or when they arrange to have a coffee break together. They also communicate through their mobile system by sending text messages to each other. However, due to the lack of support for more private channels between the members of the work community, some of the participants find this communication disturbing. Moreover, unnecessary messaging was also seen to affect driving safety. In addition to messaging, drivers can also follow where their colleagues are through the mobile system, and in this way they can meet each other when they have a break.

Because taxi drivers' work has long traditions, they also have a strong culture among their work community. In related cultural traditions and respect for customers, one taxi driver gave the example that they are supposed to "*get out of their car, open the door for customers, and talk to them*". According to the interviews, this is more important to them than knowing where a specific street is located. They value their customers and they have also been guided to minimize the use of the mobile system when customers are in the taxi. One driver said: "*Prodding at the touch screen draws too much of the customer's attention*", and customers may think that the driver is not interested in them. Instead of using the touch screen, many drivers use their mobile system via a touchpad located between the two front seats under the driver's hand, which is therefore less visible and less noticeable to the customers.

Taxi drivers are also proud of their professionalism. As one taxi driver said: "*The system cannot guide us by giving, for example, information about the travel time or travelling in general"* (for example, how long it takes to drive the customer to the destination). Accordingly, they do not want the mobile system to guide them too much, because they think they know which way to drive better than the system does.

According to our findings, environmental factors were highlighted in taxi drivers' work. Sunshine, for example, makes the work of taxi drivers difficult at times. As one taxi driver stated: *"If the sun is shining low and you have a light colored shirt, the shirt reflects the sunshine on the screen."* On the other hand, at nights or in the dark the brightness of the screen disturbs the driver if they do not use a screen saver. However, not all taxi drivers were familiar with the screen saver and therefore not all used it. Winter too, makes the taxi drivers' mobile system difficult to use as coldness causes it to boot up very slowly and it takes time to start the work.

Another important contextual factor related to system usage in the taxi drivers work was safety. As one taxi driver said: "*Every additional button press is a safety risk*". If taxi drivers are required to take their eyes off the road when they use their mobile system while driving, the safety risk increases. Moreover, too small a font size also affects safety, because users have to narrow their eyes to read the text in the display.

4.3 Case Study C: Contextual Findings in Mobile News Journalism

The work of news journalists and photographers is highly mobile by nature. In the field, an essential part of the work is to gather the material needed for reporting a newsworthy event. Traditionally, a news journalist uses a pen and a notebook for taking notes and occasionally uses a voice recorder for capturing interviews. On the other hand, a news photographer carries a digital camera and lenses in a camera bag and occasionally a digital camcorder for capturing news clips for the on-line version of the paper. Usually journalists and photographers return to the newsroom to write the story or to select and process the captured photos or videos on desktop computers. However, the work of news journalists and news photographers is changing, due to the increasing speed of news publishing, where Internet publishing may require instant reporting and submission of the material from the news scene. Therefore using mobile multimedia phones and services is of interest in this area also.

The work in a news organization is characterized by a fast pace, unexpected events and ever-changing plans. To manage the work of the editorial staff and publishing, publication schedules are drawn up and work is tied to deadlines to allow the timely printing of the newspaper. The pace of the mobile work varies from slow to hurried, covering phases from idle waiting to the moments of rushed work when the deadline is drawing near or when one needs to hurry in submitting time-critical material to an on-line publication. In on-line and mobile news "*the deadline is now*" as one of the interviewees stated.

In journalism the mobile system, in this case the mobile phone, the wireless keyboard and the mobile journalism application, is primarily a means to achieve a higher-level goal, like rapidly reporting news to the on-line publication. The mobile system, service and the related infrastructure should therefore support the achievement of these goals. Based on their usage experiences participants emphasized the need for reliable and high performance in the submission of the material. Problems encountered in the study, such as upload times that took over an hour or were interrupted, were partly due to insufficient throughput in submission and in the worst cases prevented the participants carrying out their next assignment or capturing newsworthy material. In addition, we found the editorial process as well as the publishing system to affect the user experience of the journalists and photographers. These need to compensate for the limitations of the mobile system, such as the spelling errors due to writing text with the mobile system.

The reactions of external persons in various forms were important for the users and how they felt about using the system. In their field work, both journalists and photographers work among other people, including their audience and occasionally colleagues from other news organizations. Journalists and photographers are in direct contact with their interviewees, or those they capture in photos and video clips. Participants felt that the mobile system was less visible to others in multimedia capture due its everyday like character and therefore they stated that in certain situations they could more easily fit into the crowd. On the other hand, using a mobile phone made some participants feel that they had lost a symbol of their profession: a large and visible camera. Most of the participants also mentioned that due to the familiarity of the mobile phone, their interviewees were more relaxed when they were photographed with the mobile phone than when they are photographed with a camera.

The locations the participants worked at varied from quiet spaces, like the library or the offices of the university staff being interviewed, to noisy hallways and canteens. To work with the material they had gathered both journalists and photographers typically looked for a nearby public space with chairs, sofas or preferably with tables, like cafes. One journalist described this: "*Although you can carry the system to any place, you still need the space and time to work.*" The public spaces chosen were often crowded and noisy with interruptions and disturbances from people talking and passing by. In addition, journalists' working positions while writing were dictated by the environment. Journalists had trouble in placing their needed items, such as mobile phone, foldable wireless keyboard and traditional notebook, into a stable position on their laps if no table was available. The keyboard had no locking mechanism and

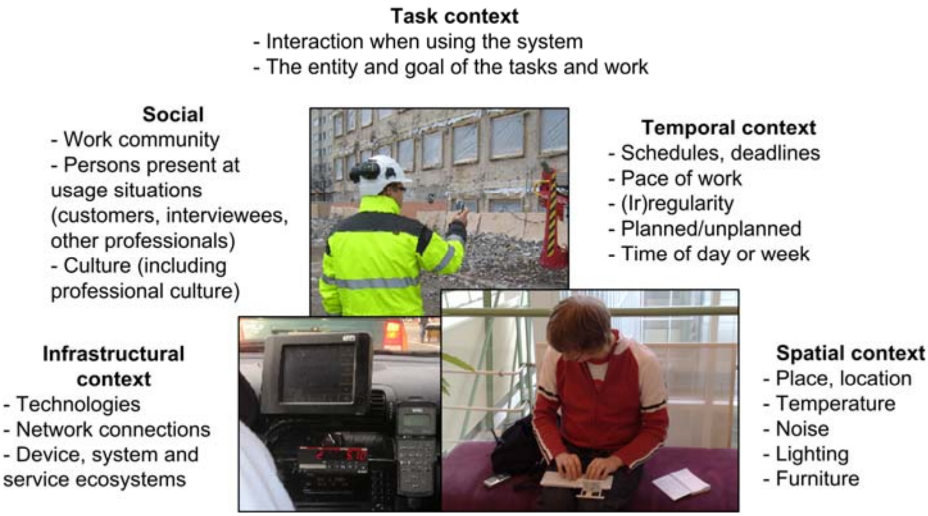


Fig. 1. Contextual factors affecting user experience in mobile work

therefore it started to fold easily while writing. There was also no mechanism for fastening the phone to the keyboard for this type of usage situations (see also Figure 1).

The environmental conditions emphasized by the participants were related to the temperature, lighting and ambient noise when using the mobile system. Photographers commented especially on lighting as an important element for capturing stills and video clips, as if it is too dim or too bright the quality of the captured photos and videos suffers due to the technological limitations of the current mobile phone - or it may not even be possible to capture multimedia material with sufficient quality for publishing. Photographers also emphasized the significance of relatively low ambient noise level for capturing video clips due to the limitations of the microphone. The handling of the mobile system in freezing weather was mentioned to be clumsy with cold fingers and it is not possible to wear gloves when working with it. In addition the battery life was mentioned as being shorter in freezing weather.

5 Discussion

In the previous section, we described our case study findings regarding the contextual factors affecting user experience. In this section, we discuss the results and categorize the findings under five dimensions of context relevant in mobile work. In addition, we discuss design implications arising from the findings.

We categorized our findings of the contextual factors affecting user experience in mobile work under five main dimensions, namely 1) social, 2) spatial, 3) temporal, 4) infrastructural, and 5) task context. These five dimensions are illustrated in Figure 1 with a summary of the findings from the three case studies. The presented framework, with its five dimensions (Fig. 1), elaborates upon the approaches in earlier frameworks of context (see Table 1). Compared to these earlier frameworks, two new dimensions are brought in, namely the social and infrastructural contexts. As discussed in Section 2, social context is highlighted in user experience frameworks [6], [7] and in context-aware computing [14]. Forlizzi et al. [6] as well as Hassenzahl et al. [7] also mention organizational and cultural factors, which we find to not be separate dimensions of context, but rather to be related to and affecting in the other dimensions of context. Infrastructural context has not been part of the earlier frameworks, although there are references to technology as being part of the context. In the following section we discuss the dimensions of the framework in more detail.

Social context. We found that the presence and reactions of external persons affected both the usage and user experience of the mobile systems. These external persons involved were either in direct contact with the user or indirectly present at the work situation. In our cases the persons who were directly involved with the users in usage situations were taxi passengers, interviewees and persons of whom photos and video clips were captured. Moreover, indirectly present were, for example, construction workers, colleagues or bystanders when reporting news in the field situations. For example, taxi drivers limited the usage of their mobile system when customers were present and preferred to use a conveniently placed touchpad, which drew less attention from the customer. In the case of journalism the participants found that their interviewees were more at ease when they were photographed with a mobile multimedia phone. We found also, that the work community and unofficial mobile communication with colleagues was important, especially in the case of taxi drivers. This also supports the earlier studies of the importance of the need for social communication in mobile work [9] [18]. Taxi drivers communicated with their colleagues through the mobile system by text messages, even though the system was not initially designed for social communication. Moreover, the professional and organizational cultures affect the values and norms of the professional community.

Spatial context. This refers not only to the location where the mobile system is used, but also to the environmental conditions and circumstances, such as available furniture, affecting the work situation and therefore the user experience. Environmental issues, such as temperature, noise, and light affect the use of mobile systems. The effects of cold upon the use and performance of the mobile systems were mentioned in all cases. When capturing multimedia material, lighting and sound conditions especially have an effect on the quality of the captured material. This was particularly emphasized by the news photographers, for whom the quality of the captured multimedia material is very important.

Temporal context. Temporal factors are related not only to the absolute time of day, week or year, but also to schedules and deadlines as well as the pace and regularity of the work. In the case of news journalism, the organization sets schedules and strict deadlines for the news reporting. The pace of the work can also vary due to, for example, external reasons not controllable by the user, like an interviewee being late for the arranged meeting. Plans and schedules may suddenly be changed and cause extreme moments of rushing to accomplish the new assignment. For the taxi drivers the time of day or week affects the pace of the work, nights and weekends being busier for them; whereas the safety observation rounds at construction sites are done regularly on a weekly basis.

Infrastructural context. The infrastructural factors, especially the functionality of the network connection was important in safety observation and in mobile news journalism. Unavailability, unreliability or other performance related problems of the network make the users frustrated and unsatisfied. Problems in network connections or throughput of the devices may force the users to redo their work tasks or prevent the users of accomplishing further tasks. Further examples of infrastructural context are carrier related costs and in some countries legal circumstances which regulate the use of the mobile system or service and therefore affect the user experience. In the case of mobile news journalism there exist also other systems and processes, including the editorial system and the editorial process, which affect the user experience of the mobile worker. The entire ecosystem of used mobile devices, systems, and services as well as related other systems and work processes could therefore be seen as part of the infrastructural context.

Task context. The task or set of tasks that users carry out with their mobile systems are often only part of higher level goals they have in their work [17]. In the case of news journalism the main tasks for a journalist that are designed to be accomplished using the mobile system are to capture photos or video clips, write the text for an article, add the multimedia items into it and then submit it to editorial. However, the actual goals in reporting news are related to the content and meaningfulness of the news, photos and videos to the audience as well as to the journalistic quality, like error-freeness, of the published material. This type of goal cannot be accomplished by any system alone, rather the user is in the key role - quoting the words of a participant of a mobile news journalism study - of "*using brains*" as well as the entire work process. Therefore, understanding both the tasks and the higher level goals is important in designing mobile systems that support the mobile work tasks. However, understanding the diversity of higher level goals especially in non-routine, creative work also simultaneously explains why mobile systems sometimes cannot be designed to support all the goals of their users. This is especially the case when both professional goals and personal ambitions are present.

Implications for design. In the light of the presented contextual dimensions and findings from the study, we present implications for design when developing mobile systems for work purposes. First, regarding the social dimension of context, the mobile systems should be acceptable both for the actual users and also for other people around them. Moreover, results indicate that social communication is important for mobile workers, and something that could be supported by providing private channels for communication. Second, the total attention required to operate the mobile system should be minimized so as to not only maintain the safety of the mobile workers but also of others. Third, mobile systems should support working in different temporal situations allowing flexible and reliable use, especially when in a hurry or when several tasks need to be performed simultaneously. Fourth, optimizing the mobile system and service for maximum throughput in data transmission is of essential importance. Fifth, the real usage environments and situations should be studied, to ensure that the developed systems and services are operable in them. For example, especially in case of mobile workers, users often move during the task or data transmission and therefore the used network connection may not be available at all times. Moreover, due to the work in the mobile field conditions, the drain of the mobile system should be minimized so as to extend the device's battery life. Sixth, designers should identify and define the environmental circumstances in which the mobile services are intended to be used. Spatial factors, such as cold and lighting conditions, can be managed, thus allowing the users to control the characteristics of the user interface and to personalize them. Things sometimes taken for granted when writing with mobile systems, like tables or other furniture, are not always available in usage situations.

The dimensions of context presented based on the cross-case analysis of the results are not the only way to categorize the contextual factors affecting user experience. However, we found the five dimensions capture our findings well and the presented dimensions can be used when analyzing and defining the contextual factors in the mobile system and service development. With the help of the framework, for example design heuristics for designing successful mobile services for mobile work and questionnaires for measuring user experiences can be prepared. Although three case studies were used in the analysis, more studies are needed to gain an insight into the similarities and differences between different fields of work as well as defining a more comprehensive list of factors within the dimensions of context. In addition, different cultural and ethical issues may affect the findings and therefore similar studies in different countries would bring interesting new insights regarding the contextual factors. Due to the relatively small number of participants in the cases, the findings can be seen as examples, but they do serve well in this purpose, although especially in the case of the mobile news journalism a further study with professionals working for newspapers is needed. We believe that the reported findings from these three different types of mobile work help in understanding the diversity of contextual factors affecting the user experience in mobile work when using mobile systems for work related tasks. Further studies are needed to verify the applicability of the framework for identifying the contextual factors which affect user experience when designing and evaluating mobile systems and services.

6 Conclusions

To explore the nature of mobile work and the contextual factors affecting user experience in different areas of mobile work, we conducted three case studies with construction workers, taxi drivers, and mobile news journalists and photographers. We described the contextual factors that were emphasized in each case study, with the presented results giving examples of the variety of contextual factors affecting user experience. We categorized the findings under five dimensions of context, which were found to affect user experience in mobile work. These five dimensions are 1) social, 2) spatial, 3) temporal, 4) infrastructural, and 5) and task context; they elaborate earlier frameworks for context, and also emphasize factors related to the social and infrastructural dimensions.

Understanding the contextual factors affecting user experience is important in designing and evaluating mobile systems and services. The five dimensions of context presented here can be utilized as a general framework when identifying more specific contextual factors as well as when evaluating the user experience of mobile systems and services. Further empirical research on context would increase the knowledge on similarities and differences in the contextual factors affecting user experience and on the applicability of the framework and its dimensions also in other fields than mobile work.

Acknowledgments. We thank our colleagues for their contribution in this work. The research on safety observation and passenger transportation were part of the project MOMENTO, funded by Tekes (the Finnish Funding Agency for Technology and Innovation). The research on mobile news journalism was part of the PALTI project (2007-2009), funded by the Ministry of Education, Finland.

References

1. Christopoulou, E.: Context as a Necessity in Mobile Applications. In: Lumsden, J. (ed.) Handbook of Research on User Interface Design and Evaluation for Mobile Technology, IGI Global, pp. 187–204 (2008)
2. Kakihara, M., Sørensen, C., Wiberg, M.: Fluid Interaction in Mobile Work Practices, Tokyo Mobile Roundtable, Tokyo, Japan (May 2002)
3. York, J., Pendharkar, P.C.: Human-computer Interaction Issues for Mobile Computing in a Variable Work Context. Int. J. Human-Computer Studies 60(5-6), 771–797 (2004)
4. Zheng, W., Yuan, Y.: Identifying the Differences Between Stationary Office Support and Mobile Work Support: a Conceptual Framework. Int. J. Mob. Commun. 5(1), 107–122 (2007)
5. Reddy, M.C., McDonald, D.W., Pratt, W., Shabot, M.M.: Technology, Work, and Information Flows: Lessons from the Implementation of a Wireless Alert Pager System. Journal of Biomedical Informatics 38(3), 229–238 (2005)
6. Forlizzi, J., Ford, S.: The Building Blocks of Experience: An Early Framework for Interaction Designers. In: DIS 2000, Brooklyn, New York, pp. 419–423 (2000)
7. Hassenzahl, M., Tractinsky, N.: User experience – a Research Agenda. Behaviour and Information Technology 25(2), 91–97 (2006)
8. Kankainen, A.: UCPCD: User-centered Product Concept Design. In: Proc. DUX 2003, pp. 1–13. ACM Press, New York (2003)
9. Nilsson, M., Hertzum, M.: Negotiated Rhythms of Mobile Work: Time, Place, and Work Schedules. In: 2005 international ACM SIGGROUP Conference on Supporting Group Work. GROUP 2005, pp. 148–157. ACM, New York (2005)
10. Pascoe, J., Ryan, N., Morse, D.: Using While Moving: HCI Issues in Fieldwork Environments. ACM Transactions on Computer-Human Interaction 7(3), 417–437 (2000)
11. Sawyer, S., Tapia, A.: The Sociotechnical Nature of Mobile Computing Work: Evidence from a Study of Policing in the United States. International Journal of Technology and Human Interaction 1(3), 1–14 (2005)
12. Hickey, S., Rosa, C., Isomursu, M.: Evaluating the Use of an Audio-video Mobile Phone for Web Magazine Reporters. In: Int. Conference on Mobile Technology, Applications, and Systems 2007, pp. 511–514. ACM Press, New York (2007)
13. Chalmers, M.: A Historical View of Context. Computer Supported Cooperative Work 13, 223–247 (2004)
14. Bradley, N.A., Dunlop, M.D.: Toward a Multidisciplinary Model of Context to Support Context-Aware Computing. Human-Computer Interaction 20, 403–436 (2005)
15. Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing 5, 4–7 (2001)

16. Dourish, P.: What We Talk about When We Talk about Context. Personal Ubiquitous Computing 8, 19–30 (2004)
17. Roto, V.: Web Browsing on Mobile Phones – Characteristics of User Experience. PhD thesis, Helsinki University of Technology, Finland (2006)
18. Wiberg, M., Ljungberg, F.: Exploring the Vision of "anytime, anywhere" in the Context of Mobile Work. In: Malhotra, Y. (ed.) Knowledge Management and Virtual Organizations, pp. 112–128. Idea Group Publishing (1999)
19. ISO 13407:1999. Human-centred Design Processes for Interactive Systems
20. Kim, H., Kim, J., Lee, Y., Chae, M., Choi, Y.: An Empirical Study of the Use Contexts and Usability Problems in Mobile Internet. In: 35th Annual Hawaii International Conference on System Sciences (2002)
21. Lee, I., Kim, J., Kim, J.: Use Contexts for the Mobile Internet: A Longitudinal Study Monitoring Actual Use of Mobile Internet Services. Int. J. Human-Computer Interaction 18(3), 269–292 (2005)
22. Turel, O.: Contextual Effects on the Usability Dimensions of Mobile Value-added Services: a Conceptual Framework. Int. J. Mobile Communications 4(3), 309–332 (2006)
23. ISO DIS 9241-210:2008. Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems (formerly known as 13407). International Standardization Organization (ISO). Switzerland. Enquiry stage 40.20 (2008)
24. Yin, R.K.: Case Study Research: Design and Methods, 3rd edn. Sage Publications, Inc., Thousand Oaks (2003)
25. Väätäjä, H., Männistö, A., Vainio, T., Jokela, T.: Understanding User Experience to Support Learning for Mobile Journalist's Work. In: Guy, R. (ed.) The Evolution of Mobile Teaching and Learning, pp. 177–210. Informing Science Press, USA (2009)
26. Wigelius, H., Aula, A., Markova, M.: Modeling Tool for Designing Usable Mobile Services. In: Tiainen, T., Isomäki, H., Korpela, M., Mursu, A., Paakki, M.-K., Pekkola, S. (eds.) Proc. of 30th Information Systems Research Seminar in Scandinavia, IRIS30, Tampere, Finland, August 11-14 (2007); Dep. of Comp. Sciences, University of Tampere, Finland, Series of Publications: D - Net Publications D-2007-9, 1330 pages (September 2007), `http://www.cs.uta.fi/reports/dsarja/`
27. Miles, M.B., Huberman, A.M.: Qualitative Data Analysis: An Expanded Sourcebook, 2nd edn. Sage Publications, USA (1994)
28. Laitinen, H., Marjamäki, M., Päivärinta, K.: The Validity of the TR Safety Observation Method on Building Construction. Accident Analysis and Prevention 51(5), 463–472 (1999)

When Joy Matters: The Importance of Hedonic Stimulation in Collocated Collaboration with Large-Displays

Jasminko Novak and Susanne Schmidt

Dept. of Informatics, University of Zurich, Binzmühlestrasse 14, 8050 Zurich, Switzerland
{novak,schmidt}@ifi.uzh.ch

Abstract. Hedonic aspects are increasingly considered as an important factor in user acceptance of information systems, especially for activities with high self-fulfilling value for the users. In this paper we report on the results of an experiment investigating the hedonic qualities of an interactive large-display workspace for collocated collaboration in sales-oriented travel advisory. The results show a higher hedonic stimulation quality of a touch-based large-display travel advisory workspace than that of a traditional workspace with catalogues. Together with the feedback of both customers and travel agents this suggests the adequacy of using touch-based large-displays with visual workspaces for supporting the hedonic stimulation of user experience in collocated collaboration settings. The relation of high perception of hedonic quality to positive emotional attitudes towards the use of a large-display workspace indicates that even in utilitarian activities (e.g. reaching sales goals for travel agents) hedonic aspects can play an important role. This calls for reconsidering the traditional divide of hedonic vs. utilitarian systems in current literature, to a more balanced view towards systems which provide both utilitarian and hedonic sources of value to the user.

Keywords: Collaboration, large displays, hedonic information systems, user experience.

1 Introduction

Hedonic aspects of system use (joy of use) are increasingly considered as an important factor in user acceptance of information systems, especially for activities with high self-fulfilling value [5, 16, 10]. Thereby, the investigation of hedonic aspects is often biased towards systems supporting user activities which are hedonic by nature (e.g. entertainment, games [12]). This is reflected in a distinction between hedonic and utilitarian systems, which differentiates between the satisfaction of self-fulfilling user needs (hedonic) and activities in which functional, pragmatic goals are in the foreground. While such a distinction may be suitable in certain occasions, here we discuss the role of hedonic aspects in activities which are utilitarian by nature.

In particular, we report on the results of an experiment investigating the role of hedonic aspects in designing interactive large-display visual workspaces for collocated collaboration in the domain of travel advisory – an issue barely examined in existing

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 618–629, 2009.

work. Face-to-face advisory is a special class of collocated collaboration in which an expert advisor interacts with the customer in creating a personalized offering of a product or service (e.g. travel planning, financial services) [24]. In contrast to online sales which target high-volume transactions with relatively simple and well-understood needs, face-to-face consultations target the sales of complex products, where customer needs are difficult to formulate and translate to a tailored solution [24]. While utilitarian concerns motivate the activity (finding a suitable vacation at suitable cost), emotional aspects of the product (vacation) and user experience of the process play an important role.

The use of large displays for group interaction has been studied in two main areas: 1) supporting group awareness and collaboration in the workplace [3, 20, 28] and 2) supporting social interaction and media sharing in (semi)public settings [17]. Examples include coordination of small groups in the medical domain [32], the use of large-displays for large-scale collaboration in NASA control rooms [15, 29] and in classrooms [8]. Earlier work also considered requirements and guidelines for single display groupware in collocated collaboration [27], but little work specifically examined the role and support for hedonic aspects, "the joy of use". The existence of two different components (hedonic and utilitarian) in user attitudes towards the use of information systems has been evidenced already in [1], but few empirical studies successfully verified this in specific domains. Huang [14] found that web-sites must fulfill both information-related and hedonic, often entertainment-related needs. Studies of hedonic and utilitarian aspects of user attitude in business management software [25] and online retail shopping behavior [2] point to similar findings.

2 Hedonic vs. Utilitarian Information Systems

Task-related qualities like perceived usefulness and perceived ease of use have been widely studied and recognized as crucial concepts explaining the user acceptance of information systems [5, 6]. Perceived usefulness reflects the user's belief that "using a specific application system will increase his or her job performance within an organizational context" [11] i.e. that the provided functionality will fulfill the needs for accomplishing given tasks. Perceived ease of use reflects the user's belief of an effortless usage of functionality. Related system characteristics are thus task-related qualities [10, 9] and provide the user with instrumental value i.e. they are of a utilitarian nature [12].

While a number of studies have confirmed this approach, the importance of perceived enjoyment has been highlighted in studies where it has played a greater role in system acceptance than perceived usefulness and ease of use [19, 30]. Perceived enjoyment refers to "the extent to which the activity of using the computer is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated" [5, p. 1113]. Therefore, enjoyment can be assumed to be a factor for intrinsic motivation. This notion has also been described with the term hedonic, referring to the concept of hedonism as a doctrine in which pleasure is the primary goal of existence [18]. As a result, information systems tend to be divided along two main lines: between approaches emphasizing the primacy of functional (task-related) aspects and those focusing on the user experience of the process.

The former are also referred to as utilitarian systems which aim at "providing instrumental value to the user" [12], while the latter are denoted as hedonic systems where value for the user comes from his self-fulfillment in using the system [12]. Utilitarian systems address tasks and activities where user motivation in using the system is driven by the expectation of a reward or benefit external to his interaction with the system [12]. On the other hand, hedonic systems serve activities in which users are intrinsically motivated by benefits stemming from the interaction with the system as such, rather than task-related outcomes of their interactions [12]. As a result, the primary design objective of utilitarian systems is productive use whereas for hedonic systems it is prolonged use of the system [12].

However well supported by different studies, such distinction seems overly sharp. On one hand, a natural goal of each system design is to encourage prolonged user acceptance. On the other hand, studies such as [10] suggest that task-unrelated qualities of software systems are an essential element in explaining system acceptance. Stimulating overall enjoyment and "joy of use" can encourage users not only to develop a positive attitude towards system use but to entice a lasting usage. Especially in the field of HCI, the importance of designing for "joy of use" rather than only for functional usefulness and ease-of-use as an important driver of user satisfaction and system acceptance has been increasingly recognized in research and practice [16, 10]. In this view, successful systems are conceived not only in terms of user efficiency in fulfilling their task, but also in terms of user enjoyment in the process as their integral part [10]. In everyday life, this has been vividly demonstrated with the success of digital consumer devices (e.g. iPod, iPhone). In spite of diverging views emphasising the primacy of efficiency and control [13], the importance of hedonic aspects has been empirically demonstrated for the design of systems mediating the satisfaction of intrinsically hedonic user needs (e.g. finding movies to watch) [12]. Though most of these findings apply to activities of a hedonic nature, they suggest that even in utilitarian activities, it may be important to consider hedonic aspects as an essential design objective. The next sections discuss this issue in collocated travel advisory.

3 Emotional Collaboration in Face-to-Face Travel Advisory

The elicitation of customer needs and their mapping to product features is the main pragmatic challenge addressed by sales-oriented travel advisory in a face-to-face setting [24]. User preferences are "sticky": difficult to elicit and to be described, evolving with the perceived possibilities from the solution space. On one hand, effective solution of this problem requires overcoming the intrinsic information asymmetry in which only the agent has access to all relevant solution space information (travel database, Internet) (Fig. 1).

On the other hand, sales advisory is not merely about pragmatic problem-solving: it is also a highly emotionally colored process. Customer criteria can rapidly change based on impressions of presented alternatives and especially so in travel advisory: a terrific photograph, a compelling video or a funny anecdote can swiftly shift customer desires and create emotional bonds to destinations or to the agent. The problem-solving process is intertwined with an emotional dimension of the consulting experience: desires are stirred, moods awaken and in this situation a decision process is

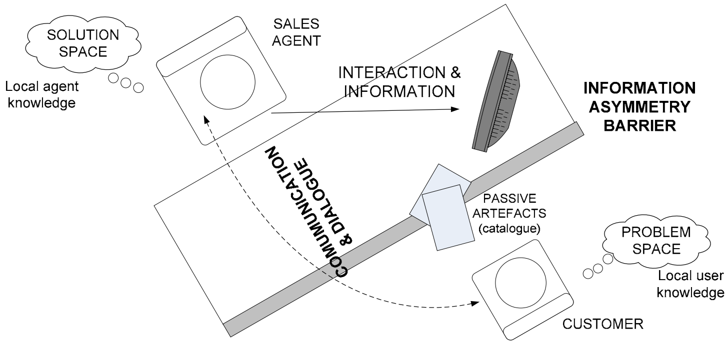


Fig. 1. Traditional travel advisory setting [21]

carried out. We refer to this as “emotional collaboration” [23]. This suggests a high importance of hedonic quality for travel advisory systems: though seeking professional travel advisory is a highly utilitarian activity (effectively finding suitable vacation offers in a desired price range), underlying user motivations for searching travel products and undertaking a vacation are intrinsically hedonic (e.g. enjoyment, relaxation, adventure). Accordingly, making the process of devising a personal itinerary an enjoyable experience – rather than a time consuming activity of shifting through loads of impersonal information - is likely to achieve higher customer satisfaction.

Achieving high customer satisfaction has an even greater importance for offline travel agencies: they are under great pressure to provide added-value services which distinguish them from purely online competition that offers easy access to a vast range of travel offerings, literally at the users’ fingertips [22]. At the same time, pragmatic aspects such as the identification and creation of offers matching personal preferences under suitable economic conditions as well as the trustworthiness of the information and of the travel agent also play an important role in the decision process [4]. Thus, we argue that designing systems to successfully support face-to-face travel consultancy requires considering both the pragmatic aspects of overcoming the obstacles of effective customer-agent collaboration [21] and the hedonic aspects of overall user experience.

4 Design Principles and Hedonic Aspects of a Prototype System

Our concept for addressing these issues is based on two main design goals: 1) reducing information asymmetry through shared visualization and 2) heightening emotional user experience (“joy of use”). Both of these are addressed by introducing a touch-sensitive large-display (a Smartboard[1]) with an interactive, visual travel advisory system. The first is expected to improve the problem-solving process (eliciting customer needs and constructing appropriate solutions) and increase the trustworthiness of the setting while the latter is aimed at facilitating the process by increasing emotional user stimulation and satisfaction. Our concrete solution is depicted in Figure 1.

[1] ®Smart Technologies, http://smarttech.de

Fig. 2. Cooperative travel advisory system based on a touch-sensitive large display [23]

As the user-centered design process and the system functionalities have been described in [23, 26], we only briefly outline the main aspects.

The physical arrangement is such that the customer and the sales agent stand in front of a large, touch-sensitive display. Large displays lend themselves readily to providing a shared visual workspace that can be inspected jointly by both participants. In contrast to multiple small screens with a single mouse interaction [24] this allows natural interaction for both participants (dragging and selecting with bare hands) and natural coordination of access to the shared screen (pointing, gesture). Above all, their visual qualities are ideal for amplifying emotional impact, while touch-interaction provides not only a natural interaction method but also stimulates sensory experience.

The travel advisory application provides a visual overview of the travel products on offer, contextualized on a geographic map (Google Maps). This travel space can be searched and browsed by different criteria and includes multimedia information both from official travel agency database and user-generated content from Internet portals and communities. Details of the developed travel advisory application and how it supports the pragmatic functionalities (e.g. shared view on information resources, proactive user exploration, user-generated content) are described in [23, 26]. To further user's emotional engagement the system takes advantage of multimedia (photos, videos) on the large screen. Due to the physical arrangement (customer and agent standing in front of the screen) and the screen size the user is engulfed in the visual experience. The closeness to the display invites active involvement. As reflected in user comments, this stimulates hedonic experience by capturing attention, providing exciting novel functionalities and easy access to otherwise unavailable content [11].

5 Experimental Evaluation: Pragmatic vs. Hedonic Qualities

The results of a first evaluation of this design concept in a proof-of-concept prototype suggested that functional qualities of the system and the user experience are superior to the traditional setting [23, 26]. They exhibited a high user preference for the consultancy with the large-display workspace (10 of 12 users). The informal feedback pointed out user excitement about the visual qualities of the large-display and the experience of touch-based interaction [23]. User observation also suggested emotional experience as a possible factor of user satisfaction and system acceptance. Accordingly, we undertook a second experiment in order to investigate hedonic qualities in more detail. To this end we employed the AttrakDiff2 questionnaire for elicitation of perceived hedonic quality [11] accompanied by the UTAUT questionnaire for elicitation of user attitude to using the system [31].

5.1 Investigating the Role of Hedonic Quality

In order to determine the hedonic qualities of an interactive system the AttrakDiff2 instrument [11] differentiates between pragmatic and hedonic qualities as independent constructs determining the overall perception of the attractiveness of a system [9, 10]. Pragmatic quality (PQ) refers to the perceived quality of manipulation (i.e. effectiveness and efficiency of use). Hedonic quality (HQ) is described in terms of hedonic stimulation (HQ-S) and hedonic identity (HQ-I). The stimulation quality refers to the extent to which the system stimulates the innate human need for personal development (e.g. new skills and knowledge). By offering exciting functionalities, content or interaction styles, a system can heighten the user's attention, overcome motivational barriers or ease a problem-solving process [11] – thus supporting effective task completion and further usage. The identity aspect refers to addressing a personal need of expressing oneself and being perceived by others in a certain way [11]. As people commonly express themselves through personal objects (e.g. clothes, jewelry, mobiles), the functionalities, design or visual appearance of a system can relate to a user's need for communicating a certain identity. In the operationalization users are asked to rate the system on 7 semantic differentials (Table 1) for each construct with a 7-point Likert-scale.

Table 1. Semantic differentials of hedonic stimulation (HQ-S) and pragmatic quality (PQ) [11]

Indicator for	positive part	negative part
HQ-S	challenging	harmless
HQ-S	exciting	dull
HQ-S	creative	uninspired
HQ-S	inventive	conventional
HQ-S	novel	usual
HQ-S	innovative	conservative
HQ-S	courageous	careful
PQ	practical	unpractical
PQ	manageable	unmanageable
PQ	predictable	unpredictable
PQ	clearly	confusing
PQ	directly	awkward
PQ	humanly	technical
PQ	simple	complicated

In line with the prior discussion we stated the following hypotheses[2]:

H1.1: The hedonic stimulation quality of the touch-sensitive large-display workspace perceived by the customers is higher than that of the traditional travel consultancy workspace.

H1.2: The hedonic stimulation quality of the touch-sensitive large-display workspace perceived by the travel agents is higher than that of the traditional travel consultancy workspace.

H2.1: The pragmatic quality of the touch-sensitive large-display workspace perceived by the customers is higher than that of the traditional travel consultancy workspace.

H2.2: The pragmatic quality of the touch-sensitive large-display workspace perceived by the travel agents is higher than that of the traditional travel consultancy workspace.

5.2 Experiment Design

The experiment involved 22 customers divided in two groups in a between-subjects design: one group of 11 participants performed a vacation planning task (planning an activity holiday) in the traditional consultancy setting while the other group of 11 participants performed the same task with the large-display workspace. We have chosen a between-subjects design to exclude any possible carry-over effects and user strain noticed in the first test. The sessions were conducted by 4 travel agents in the traditional and 5 agents in the large-display setting.

The system used in the traditional setting was comprised out of a standard PC with the travel agency database (incl. photos and videos), Internet access and print catalogues (Figure 1). The experiment took part in a real-world travel agency specialized in individual and student travel. The participants were randomly recruited from the agency's clients and through postings on a university forum (each received a 150 CHF travel discount voucher). They were between 21-37 years of age with high proficiency in computer use (80%) and both groups having the same composition regarding participants' sex (36% female, 64% male).Travel agents received a 30mins hands-on training with the system. Customers received no training. The sessions were limited to 30mins (duration of a typical advisory session). All received the AttrakDiff2 questionnaire [11] based on a 5-point Likert-scale[3].

5.3 Results

The results are depicted in Figure 3. We considered both hypotheses for customers and agents. The data was tested with a two-sided t-test for samples with differing variances. The results indicate a statistically significant ($p<0.05$, $df=17$, $t=3.61$) higher hedonic stimulation (HQ-S) for the large-display workspace (HQ-Savg = 4 vs. 3.17) regarding customers. This confirms the hypothesis H1.1. The difference in

[2] We don't consider hedonic identity (HQ-I) since the system doesn't provide means for the user to communicate a personal identity to other people. For HQ-I constructs see [10].

[3] The 5-point Likert-scale was employed in order to keep a uniform scale design across all questionnaires in the experiment (others addressed issues not subject of this paper).

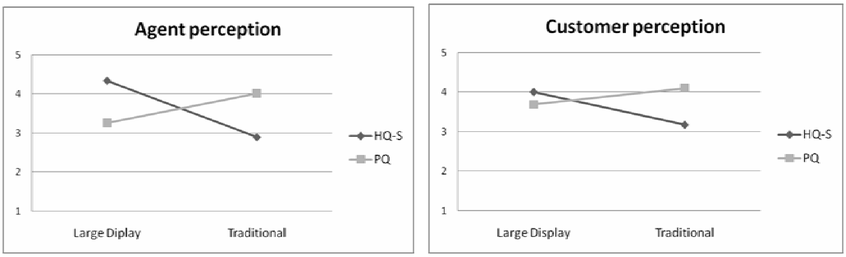


Fig. 3. Pragmatic and hedonic qualities of the large-display vs. traditional workspace

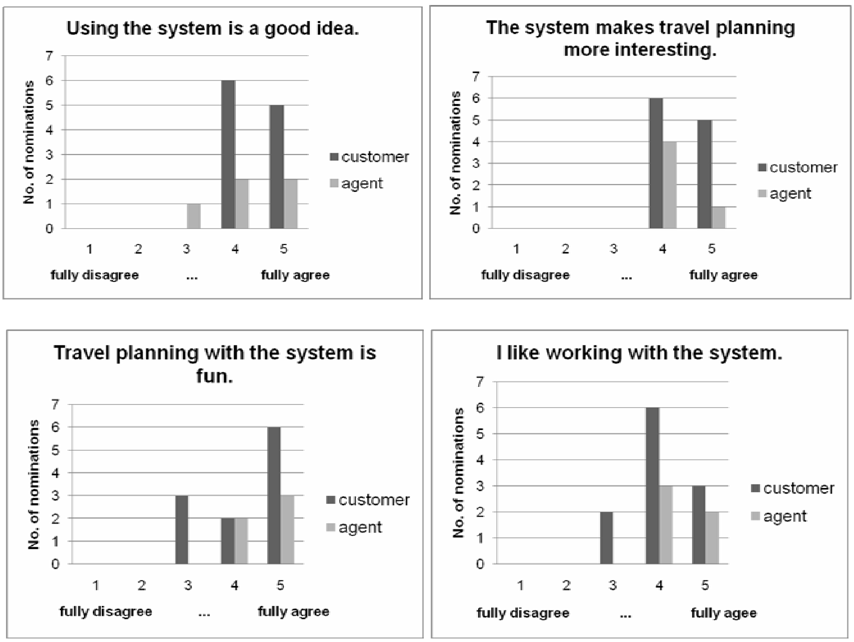


Fig. 4. Results of the UTAUT questionnaire regarding attitude towards using the system

pragmatic quality (PQ) is statistically significant regarding customers ($p<0.05$, df=20, t=-3.21) but lower for the large-display workspace (PQavg =3.67 vs. 4.1). This refutes H2.1.

Agents also rate a higher hedonic stimulation (HQ-Savg=4.34 vs. 2.89) for the large-display workspace. The difference is statistically significant ($p<0.05$, df=5, t=8.05). This confirms H1.2. The lower perception of pragmatic quality for the large-display workspace by the travel agents (PQavg=3.25 vs. 4.01) is in line with the lower pragmatic quality perceptions of the customers, but not statistically significant. Thus H2.2 could not be validated in either sense.

The results of the UTAUT questionnaire (Figure 4) indicate the positive attitude towards using the system both for customers and travel agents. All customers and all but one agent considered using the system to be a good idea, while all of them perceived the system making travel planning more interesting. All agents and the majority of customers (72.7%) liked working with the system and considered travel planning with the system to be fun (all agents, 81.8% customers).

5.4 Discussion and Limitations

Given no other differences between the two travel advisory workspaces, such results suggest that the use of a touch-sensitive large-display with a visual workspace leads to a higher hedonic stimulation. The lower pragmatic quality of the large-display workspace compared to the traditional setting additionally supports this interpretation. As the elicited pragmatic qualities (e.g. manageable, practical, simple, see Table 1) relate to the perceived ease of using the system in accomplishing the given task, the overall positive user attitudes towards the system (like working with, makes travel planning more interesting etc., Figure 4), can be related to its task-unrelated aspects i.e. hedonic quality. Since effective support of hedonic aspects is an important requirement of this problem class (as argued in introductory sections) this implies that using touch-sensitive large-displays is a good design choice for interactive visual workspaces in travel advisory. Coupled with informal user feedback highlighting the visual qualities and touch-based interaction, these findings also suggest that touch-based large-displays with a highly visual interactive application seem to exhibit intrinsic qualities of hedonic stimulation. This also makes them an interesting choice for other application domains in which hedonic stimulation and emotional collaboration play an important role.

The results of the travel agent perceptions point to another interesting observation. Though their motivation for using travel advisory support systems (e.g. booking systems, catalogues, internet) is highly utilitarian by nature (reaching their individual sales goals effectively and efficiently), their responses indicate a high perception of hedonic quality and positive emotional attitudes towards the large-display system (makes fun to use, like working with). This indicates that even in utilitarian activities (e.g. travel sales from the agents' point of view) hedonic aspects of the supporting system can play an important role in user attitude. This suggests that the existing distinction in literature between utilitarian and hedonic information systems [12], each providing a different source of value to the user (self-fulfilling vs. instrumental) should be reconsidered towards a more balanced view including a class of systems combining both sources of value.

A likely cause for lower perception of pragmatic quality of the large-display workspace (contradicting a previous trial [23]) may be the lack of proficiency of agents in using the system and in adapting it to their advisory practice. In fact, one limitation of the validity of the results is that though explicitly asked to rate the large-display workspace the participants may not have entirely differentiated between this system and their overall impression of the advisory process.

In other words, the high ease-of-use and the ease of learning to use the system with the large display (reported in [23]), do not automatically translate into effective advisory process in the new, technologically-enhanced setting. In contrast, the advising process in the traditional setting is well established with extensive experience of the agents. This is likely to lead to a higher user perception of the advisory process effectiveness. This suggests the need for a new advisory process, better taking advantage of pragmatic functionalities offered by the system. Such a new process may emerge from agent self-learning through more extensive use of the system in advisory practice or it could be developed and trained in advance. Here, we can only speculate that this may increase the perception of pragmatic quality. But we can exclude a possible influence of this factor on hedonic stimulation since hedonic quality of the traditional setting in which advisors are more proficient was significantly lower than that of the large-display system. Hence, advisor training could only further improve this difference to the advantage of the new system.

6 Conclusions

In this paper we have discussed the importance of hedonic qualities for interactive systems supporting collocated collaboration in utilitarian activities such as sales-oriented travel advisory. The results of our experiment show a higher hedonic stimulation quality of a touch-based large-display cooperative travel consultancy workspace than that of a traditional advisory setting. Coupled with qualitative user feedback highlighting the visual qualities and touch-based interaction, this suggests intrinsic hedonic stimulation qualities for large-display visual workspaces. This makes them a suitable design choice for supporting collocated collaboration in expert-customer sales consultancies as well as an interesting medium for other application domains with high importance of hedonic factors. Finally, the results indicate that even in utilitarian activities (e.g. travel sales from the agents' point of view) hedonic aspects can play an important role in user attitude towards system use. This calls for considering a "new" class of hybrid systems which combine pragmatic and hedonic sources of value for the user.

References

1. Batra, R., Ahtola, O.T.: Measuring the hedonic and utilitarian sources of consumer attitudes. Marketing Letters 2(2), 159–170 (1991)
2. Childers, T.L., Carr, C.L., Peck, J., Carson, S.: Hedonic and utilitarian motivations for online retail shopping behavior. Journal of Retailing 77(4), 511–535 (2001)
3. Churchill, E.F., Nelson, L., Denoue, L.: Multimedia Fliers: Information Sharing With Digital Community Bulletin Boards. In: Proc. of Communities and Technologies 2003 (2003)
4. Coulter, K.S.: The Effects of Travel Agent Characteristics on the Development of Trust: A Contingency View. Journal of Travel&Tourism Marketing 11(4), 67–85 (2002)
5. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. Journal of Applied Social Psychology 22(14), 1111–1132 (1992)
6. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 13(3), 319–340 (1989)

7. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User Acceptance of Computer Technology: A Comparison of two Theoretical Models. Management Science 35(8), 982–1003 (1989)
8. Ganoe, C.H., Somervell, J.P., Neale, D.C., Isenhour, P.L., Carroll, J.M., Rosson, M.B., McCrickard, D.S.: Classroom BRIDGE: using collaborative public and desktop timelines to support activity awareness. In: Proc. of the ACM, UIST 2003, New York, pp. 21–30 (2003)
9. Hassenzahl, M., et al.: Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal. In: Proc. CHI 2000 Conf. Human Factors in Computing, pp. 201–208. ACM Press, Addison-Wesley, New York (2000)
10. Hassenzahl, M., Beu, A., Burmester, M.: Engineering Joy. IEEE Software 18(1), 70–76 (2001)
11. Hassenzahl, M., Burmester, M., Koller, F.: Attrakdiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Proc. Mensch & Computer, pp. 187–196 (2003)
12. van der Heijden, H.: User acceptance of hedonic information systems. MIS Quarterly 28(4), 695–704 (2004)
13. Hollnagel, E.: Keep Cool: The Value of Affective Computer Interfaces in a Rational World. In: Proc. HCI Int'l 1999, vol. 2, pp. 676–680. Lawrence Erlbaum, Mahwah (1999)
14. Huang, M.-H.: Designing website attributes to induce experiential encounters. Computers in Human Behavior 19(4), 425–442 (2003)
15. Huang, E.M., Mynatt, E.D., Trimble, J.P.: When design just isn't enough: the unanticipated challenges of the real world for large collaborative displays. Personal Ubiquitous Computing 11(7) (2007)
16. Igbaria, M., Schiffman, S.J., Wieckowski, T.J.: The Respective Roles of Perceived Usefulness and Perceived Fun in the Acceptance of Microcomputer Technology. Behaviour & Information Technology 13(6), 349–361 (1994)
17. Izadi, S., et al.: The iterative design and study of a large display for shared and sociable spaces. In: Proc. of DUX 2005. ACM Press, New York (2005)
18. Merriam-Webster. Merriam-Webster's Collegiate Dictionary. Merriam-Webster Inc., Springfield (2003)
19. Moon, J.-W., Kim, Y.-G.: Extending the TAM for a World-Wide-Web Context. lnformation and Management 38(4), 217–230 (2001)
20. Mynatt, E.D., et al.: Flatland: New dimensions in office whiteboards. In: Proc. CHI 1999. ACM Press, New York (1999)
21. Novak, J.M.: Yours...Ours? Designing for Principal-Agent Collaboration in Interactive Value Creation. In: Proc. of Wirtschaftsinformatik 2009, Vienna (2009)
22. Novak, J., Schwabe, G.: Designing for reintermediation in the brick-and-mortar world: Towards the travel agency of the future. In: Electronic Markets, vol. 19(1). Springer, Heidelberg (2009)
23. Novak, J., Aggeler, M., Schwabe, G.: Designing Large-Display Workspaces for Cooperative Travel Consultancy. In: Proc. of CHI 2008, Florence, Italy. ACM Press, New York (2008)
24. Rodden, T., et al.: Designing novel interactional workspaces to support face to face consultations. In: Proc. CHI 2003, pp. 57–64 (2003)
25. Schrepp, M., Held, T., Laugwitz, B.: The influence of hedonic quality on the attractiveness of user interfaces of business management software. Interacting with Computers 18(5), 1055–1069 (2006)
26. Schwabe, G., Novak, J., Aggeler, M.: Designing the Tourist Agency of the Future. In: 21st Bled eConference on eCollaboration, Slovenia (2008)

27. Stewart, J., Bederson, B.B., Druin, A.: Single display groupware: a model for co-present collaboartion. In: Proc. of CHI 1999. ACM Press, New York (1999)
28. Streitz, N.A., Geißler, J., Holmer, T.: Roomware for cooperative buildings: Integrated design of architectural spaces and information spaces. In: Streitz, N.A., Konomi, S., Burkhardt, H.-J. (eds.) CoBuild 1998. LNCS, vol. 1370, p. 4. Springer, Heidelberg (1998)
29. Tollinger, I., et al.: Collaborative knowledge management supporting mars mission scientists. In: Proc. CSCW 2004. ACM Press, New York (2004)
30. Venkatesh, V.: Creation of Favorable User Perceptions: Exploring the Role of Intrinsic Motivation. MIS Quarterly 23(2), 239–260 (1999)
31. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User Acceptance of Information Technology: Toward A Unified View. MIS Quarterly 27(3), 425–478 (2003)
32. Wilson, S., Galliers, J., Fone, J.: Not all sharing is equal: the impact of a large display on small group collaborative work. In: Proc. CSCW 2006. ACM Press, New York (2006)

The 'Joy-of-Use'-Button: Recording Pleasant Moments While Using a PC

Robert Schleicher[1] and Sandra Trösterer[2]

[1] Deutsche Telekom Laboratories
[2] Chair of Human-Machine Systems, TU Berlin,
Straße des 17. Juni 135, 10623 Berlin, Germany
robert.schleicher@tu-berlin.de, sandra.troesterer@mms.tu-berlin.de

Abstract. This paper describes the conceptualization and implementation of a free software tool to record positive moments while using a PC. The application runs in the background and is only visible as an icon in the system tray. Once called, the user is prompted to quantify the Joy-of-Use experience with a five-point rating which is saved together with a screenshot of the currently active window. The user can also mark areas in the screenshot that were relevant to the positive experience. In the advanced mode, the event can additionally be characterized by selecting typical aspects of such moments that were determined empirically, or by giving an own description in a free text field. A running version for Microsoft Windows XP and its source code are available online for other researchers and practitioners who wish to collect their own data on Joy-of-Use.

Keywords: emotion, fun, usability evaluation tool, Microsoft Windows XP.

1 Introduction

Current approaches in Human-Computer-Interaction claim that research should comprise more than just assessing the effective and efficient solution of the task at hand, insisting on an holistic view on the overall *user experience* beyond the predominant ergonomic criteria of classical usability research [1]. To describe what else makes using an application or electronic device enjoyable, different terms like *emotional design* [2] or *hedonic quality* [3] have been coined. Here, phrases like *joy of use* [4] illustrate that attention is paid to factors that actively promote pleasant interaction instead of simply singling out potential annoyances – similar to a development in emotion research to also examine the nature of positive emotions after predominantly enquiring into the origins of affects like fear or anger/aggression [5]. However, in both areas it sometimes appears almost more difficult to induce and capture positive moments than unpleasant ones, especially if one tries to move away from pure laboratory settings. Two interdisciplinary student projects at the Technical University of Berlin tried to investigate this issue and come up with a tool to record joy-of-use experiences. Their results are reported here.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 630–633, 2009.

2 How to Measure Joy-of-Use?

The main purpose of the initial project was two-fold: first to obtain a broad collection of moments of pleasant human-computer-interaction and next to develop own ideas on how to measure these events in practice. To fulfill the first goal, all 12 participants (7 male, 5 female students of human factors or technical graduate studies like engineering, computer science etc. with an age range from 20 to 30 years) were asked to keep a diary for one month on when and where they enjoyed using a computer or an electronic device. The description did not have to be extensive, however, it should be clear what caused the pleasant feeling specifically and why. As well 'pleasant moments' should include all events from persistent enthusiasm to brief satisfaction – anything that was somehow consciously perceived as positive.

Altogether 157 events were collected and classified with respect to typical aspects and situations in which they occurred. Surprisingly, the most frequent category was labeled *time saving* with events like 'using a shortcut instead of having to click through the menu tree', which is more or less equivalent to the classic usability criterion of efficiency. Other categories were for example *infotainment* (e.g. the 'hybrid view'-option in google maps) or *availability/multifunctionality* (e.g. dual use of mp3 player as external data storage). The complete list of common properties can be seen in figure 1.

Next, the most popular measurement instrument and tools to assess user satisfaction or hedonic quality of a software were determined by reviewing available literature. Besides verbal questionnaires like the SUMI (Software Usability Measurement Inventory [6]) and AttrakDiff [7], a pictogram-based rating scale, the Self-Assessment-Manikin (SAM, [8]) appeared to be the most commonly used methods. However, all instruments are usually filled in after a long interaction phase and it was in dispute to what extent they would have been able to capture the brief moments of enjoyment that were reported in the self-observations. For these transient and less intense events an unobtrusive and quick rating option seemed best-suited. Thus some kind of button that could be pressed when desired was suggested. Additional requirements were that any tool should not demand too much of the user's resources (i.e. disrupt ongoing activity) and have a clear emphasis on positive aspects of human-computer-interaction to prevent it from being turned into an electronic complaint form.

3 The 'Joy-of-Use'-Button

The concept of a continuously available, but unobtrusive button was realized as a program that runs in the background and is visible only in the system tray (see Figure 1). For practical considerations it was decided to implement this application in C# and choose Microsoft Windows XP as the target operating system. A stable running version and its source code is available online[1].

If the user wishes to record a Joy-of-Use (JoU) experience, she calls the application with a mouse click on the icon or a self-defined keyboard shortcut and the rating

[1] http://www.tu-berlin.de/index.php?id=53724

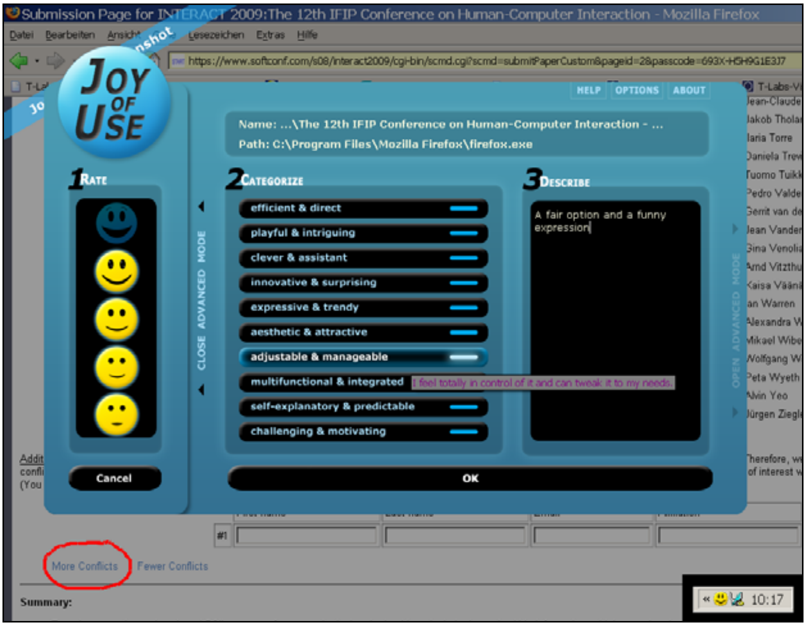


Fig. 1. the Joy-of-Use-Button interface in advanced mode. In the simple mode, only the smiley scale is present (1). In the advanced mode, common aspects of Joy-of-Use epxeriences can be selected (2), whereby each item is briefly described in a tooltip-statement which is visible for the item 'adjustable & manageable' here. Additional remarks can be entered in the free text field (3). In the background a screenshot of the currently open application, where the user marked the 'More conflicts' link in the left lower corner to specify what specifically caused the Joy-of-Use moment. In the right lower corner the icon as it appears in the system tray.

interface pops up. There are two versions of the interface: In the simple mode, she just has to give a rating on the five step smiley-scale, the dialogue window disappears and she can continue working. Her rating as well as a the path of the currently active application is saved as ASCII data, together with a screenshot of the current active window where the user can also mark areas of interest while the JoU-Button is active. In the advanced mode, she can additionally select up to ten properties that had been previously identified as characteristic of JoU-moments or make further remarks in a free text field. Figure 1 shows the functionalities of the advanced mode.

4 Summary and Conclusion

This project explored Joy-of-Use in humane-machine interactions and how to adequately record it in an applied environment. The collection of self-experienced Joy-of-Use moments as well as their categorization are surely neither exhaustive nor final and need to be validated in further studies. However, the intention was to get an idea about general characteristics of these kinds of events. One central finding was that many moments of enjoyable interaction are not that intense and are of short duration, or at least, their conscious reflection 'I like this' is only present for a brief moment. To capture these moments, a measurement device should be available on the spot and operable without requiring many resources that have to be deduced from the task one is enjoying at the moment. For a similar reason we also refrained from implementing a paced or automatic pop-up prompt for a rating – it would possibly interrupt the flow and as such alter the phenomenon it is trying to measure.

At the present stage, the main purpose of this application is to allow other researchers and practitioners to compare our observations with their own experiences. If it turns out that the present verbal labels in the advanced mode are not satisfying, they can be easily changed in the source code that is also available on the homepage. All recorded data are stored locally and only accessible to the current user. Those who are willing to make their recordings available for further analysis can send them to joyofuse@qu.tlabs.tu-berlin.de . All data will be used for scientific purposes only and processed anonymously.

Aknowledgements. We would like to thank the participants of both student projects for their collaboration.

References

1. Hassenzahl, M., Tractinsky, N.: User experience - a research agenda. Behaviour & Information Technology 25, 91–97 (2006)
2. Norman, D.A.: Emotional design: why we love (or hate) everyday things. Basic Books, New York City (2004)
3. Hassenzahl, M.: The Thing and I: Understanding the Relationship Between User and Product. In: Blythe, M.A., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.) Funology. From Usability to Enjoyment, pp. 31–42. Kluwer Academic Publishers, Dordrecht (2004)
4. Nielssen, J.: User empowerment and the fun factor. In: Blythe, M.A., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.) Funology. From Usability to Enjoyment, pp. 103–105. Kluwer Academic Publishers, Dordrecht (2004)
5. Lemonick, M.D.: The biology of joy. Time. 165, A12-4, A17 (2005)
6. Kirakowski, J., Corbett, M.: SUMI: The Software Usability Measurement Inventory. British Journal of Educational Technology 24, 210–212 (1993)
7. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Szwillus, G., Ziegler, J. (eds.) Mensch & Computer 2003. Interaktion in Bewegung, pp. 187–196. Teubner, Stuttgart (2003)
8. Bradley, M.M., Lang, P.J.: Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. Journal of Behavior Therapy and Experimental Psychiatry 25, 49–59 (1994)

Introducing a Pairwise Comparison Scale for UX Evaluations with Preschoolers

Bieke Zaman

CUO/IBBT, KULeuven, Parkstraat 45 bus 3605, 3000 Leuven, Belgium
Bieke.Zaman@soc.kuleuven.be

Abstract. This paper describes the development and validation of a pairwise comparison scale for user experience (UX) evaluations with preschoolers. More particularly, the dimensionality, reliability and validity of the scale are discussed. The results of three experiments among almost 170 preschoolers show that user experience cannot be measured quantitatively as a multi-dimensional construct. In contrast, preschoolers' UX should be measured directly as a one-dimensional higher order construct. This one-dimensional scale encompassing five general items proved to be internally consistent and valid providing evidence of a solid theory-based instrument to measure UX with preschoolers.

Keywords: UX evaluation, pairwise comparison scale, preschoolers.

1 Introduction

Making decisions on the development and launch of new technologies has become very difficult. Nowadays, complex product characteristics matter as unique selling points to distinguish from competitors. Companies have to come up with innovative ideas that generate a noticeable user experience that totally fits today's users' unfulfilled, and often unspecified dreams and wishes. These user experiences are often hard to reveal, and even harder to measure. Nonetheless, the launch of a new or improved digital product totally depends on whether it leads to a better user experience than the previous versions or competitor's products. The importance of good benchmarking measurement tools for decision makers can thus no longer be neglected.

Standardized questioning based on *closed-ended scales* is necessary in order to benchmark and compare user experiences across products. For each product or user tested, the conditions –items and underlying scale- should be similar. This way, closed-ended scales allow measuring quantitative differences in product preferences based on user experiences.

Because of the ungraspable or intangible character of user experiences, we cannot analyze user experience as an independent measure. Instead, comparable or relative measures are needed. *Relative scaling* reveals a motivated domination of some objects over others with respect to a common attribute [1]. In our case, one technology might be preferred over another with respect to the user experience it brings to the user. Each relative UX score depends on the score on the same attribute UX of another object. This contrasts physical scaling which scores concrete attributes such as length, temperature or weight independently [1]. The strength of asking respondents to make

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 634–637, 2009.

relative comparisons and judgments, rather than absolute judgments has first been theorized by Thurstone in 1927 [2].

In 2004, Hanna et al. [3] suggested the potential of *pairwise comparisons* for the evaluation of digital technologies with children. As far as we know, no research project ever took that challenge. In general, quantitative methods that rely on relative and closed-ended scaling for UX evaluations with young children rarely exist. Most methods for preschoolers are qualitative, and the existing quantitative methods are often not appropriate to work with this age group who cannot yet read or write. In this paper, however, we will suggest a pairwise comparison scale for UX evaluations that is adapted to work with preschoolers.

2 The UX Scale for Preschoolers

In the next paragraphs, we will report on the results of three experiments in which a UX scale for preschoolers was developed and tested. The UX scale is part of the *This-or-That method*, a mixed-method approach to benchmark children's user experiences with technologies [4]. Chronologically, the This-or-That method equates a within-subject experiment consisting of four phases in which children are invited individually to judge preferences on user experiences of two technologies. During the first phase, children can explore both technologies. Secondly, a quantitative survey questionnaire (based on the UX scale) is verbally administered to let children judge preferences between both conditions. The answers are then validated qualitatively through a short probing interview. Finally, their preferences are also validated through a behavioural choice: at the end of the experiment, we allowed children to play one of the two conditions again as 'a reward for good participation' ('free play option').

2.1 The Dimensionality of the UX Scale for Preschoolers

In [5], we described how literature review on the one hand, and empirical qualitative data on the other hand revealed a preliminary five component classification of the UX construct for preschoolers: 1) challenge & control, 2) fantasy, 3) creative and constructive expressions, 4) social experiences, 5) body and senses. These categories have been repeatedly reported as explanatory factors for what children positively experience and thus like in technology. We then hypothesized that these categories are correlated and together measure the underlying construct of user experience. This classification should make from UX a construct that can be measured multi-dimensionally.

Three out of five components were measured by five specific questions each and tested in an experiment with 36 preschoolers (Ngirls:17, Nboys:19 - Mage:65 months, SDage:8 months). In the experiment, we added five more general questions to control whether UX can also directly be measured on a one-dimensional level. Because of the low *internal consistency* of the three specific subscales related to the three specific UX components (Chronbach's alphas with values lower than .6), we decided to further test the one-dimensional UX (sub)scale with preschoolers. The Chronbach's alpha for the one-dimensional UX subscale indicated a good reliability: .882 (M=7.91, SD=2.050). The high internal consistency of this last overall UX scale was confirmed in a second experiment among 113 preschoolers (*N*boys:56, *N*girls:57, *M*age:58,39 months or 5 years, *SD*age:14 months) resulting in a Chronbach's alpha of .797 (*M* =6,147, *SD* =1.569).

Results of *principal component analysis* confirm that UX cannot be measured multi-dimensionally with preschoolers. We had to refute our hypothesis that for each specific component of our UX construct the corresponding variables would correlate.

We can conclude that although the multi-dimensionality of UX can be measured in a qualitative way [7], its multi-dimensionality cannot be measured quantitatively with such a young age group. UX should instead be measured directly as a one-dimensional higher order construct. This one-dimensional scale measures UX via the following five general items: 1) *"Show me which product was most fun?"* 2) *"Show me which product would you like to receive as a gift?"* 3) *"Show me which product would you like to take home?"* 4) *"Show me which product would you like to play again?"* 5) *"Show me which product you found a little bit stupid?"*

2.2 Convergent Validity, Criterion Validity and Scale Validity

The *convergent validity* of the general UX scale was assessed by comparing its scores to a free play option. In both experiments, we found a significant correlation between the first free play moment and the scale (r=.570, N=33 for the first experiment, and r=.581, N=111 for the second experiment, both Kendall's tau at the p<.01 level).

As for the scale's *criterion validity*, many research papers do report on the relationship between usability and hedonic qualities [e.g. 6,7]. Our results are in line with previous studies. We indeed found a significant correlation between the results on our UX scale and the usability of the product. More particularly, the results of our first experiment showed significant correlations between the UX scale and both the completion time with both interfaces as well as the successful completion of the game. More details on these correlations are discussed in [8]. In sum, the correlations between the UX scale and usability in our first experiment suggest a good criterion validity of our scale.

A third experiment was set up with 18 children (*M*age: 4.3 years, *N*boys:9, *N*girls:9) to test the accuracy of preschoolers' responses (*scale validity*). In this third experiment, children were asked to judge two identical games, so that there was no difference between both conditions, except from the actual UX. The researcher did not mention anything about the fact that they had to play and judge one game twice.

The results of the third experiment show that children only had difficulties or doubts to state preferences as an answer to the first question; after that first question they could answer more easily. To our surprise, the overall product preference expressed via the UX scale corresponded well with the free play option (Kendall's tau r=.659, N=13, p<.01). The significant correlation between children's answers and their behavioural preference refutes the idea that children were inventing answers to please the adult. In contrast, the results assume that children discerned real differences in user experiences between the alternatives and answered consistently. The qualitative interview conducted after the questionnaire, confirmed that children were judging the experiences they had with the game rather than the game characteristics in isolation. Most of their answers reflected a more positive UX towards the second condition than the first, which was most likely due to the fact that they were more experienced in the second condition and therefore could reach higher goals as well.

3 Conclusion

In this paper, we aimed at filling the gap on quantitative methods that allow for product benchmarking and UX evaluations with preschoolers. More particularly, we found that pairwise comparison techniques lead to reliable answers from these young children who cannot yet read or write. The paper introduced a valid five-item pairwise comparison scale that allows for relative measurement of UX and comparisons across two products in terms of preferences. The scale is perfectly suitable for prototyping to decide on which alternative designers should continue. In such a design context, the quantitative phase should be complemented with qualitative research to get insight into the reasons for UX preference in order to recommend design improvements. In follow-up papers, we will document more on this mixed-method This-or-That approach -including the UX scale and qualitative interview techniques- and how/why it is exactly so appropriate to use with preschoolers. We will also go more deeply into the possibilities that our quantitative and pairwise comparison scale of UX provide.

References

1. Saaty, T.L.: Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors-The Analytic Hierarchy/Network Process. RACSAM 102(2), 251–318 (2008)
2. Thurstone, L.L.: A law of comparative judgement. Psychological Review 101(2), 266–270 (1994)
3. Hanna, L., Neapolitan, D., Risden, K.: Evaluating computer game concepts with children. In: IDC 2004. ACM Press, Maryland (2004)
4. Zaman, B., Abeele, V.: How to Measure the Likeability of Tangible Interaction with Preschoolers. In: Proc. CHI Nederland, pp. 57–59. Infotec Nederland BV Woerden (2007)
5. Zaman, B., Abeele, V.: Towards a Likeability Framework that meets Child-Computer Interaction & Communication Sciences. In: Proc. of IDC, pp. 1–8 (2007)
6. Hassenzahl, M., Tractinsky, N.: User experience – a research agenda. B&IT 25(2), 91–97 (2006)
7. Tractinsky, N.: Aesthetics and apparent usability: Empirically assessing cultural and methodological issues. In: Proc. of CHI, pp. 115–122 (1997)
8. Abeele, V., Zaman, B., Vanden Abeele, M.: The Unlikeability of a Cuddle Toy Interface: An experimental study of preschoolers' likeability and usability of a 3D game played with a cuddle toy versus a keyboard. In: Markopoulos, P., de Ruyter, B., IJsselsteijn, W.A., Rowland, D. (eds.) Fun and Games 2008. LNCS, vol. 5294, pp. 118–131. Springer, Heidelberg (2008)

The Effect of Brand on the Evaluation of Websites

Antonella De Angeli, Jan Hartmann, and Alistair Sutcliffe

Manchester Business School - University of Manchester
Booth Street West – Manchester, M15 6PB, UK

Abstract. The effect of brand on consumer attitudes towards real and virtual goods is largely documented in consumer psychology and marketing. There is an obvious link between the design of a website and its brand. Yet, this effect has attracted little attention from the HCI community. This paper presents empirical evidence showing that brand attitude influences the evaluation of websites. The effect was reliable across different measures: people holding better attitudes were more positive in the evaluation of aesthetics, pleasure and usability. A sample of students (N=145) with a background in HCI was tested, suggesting that brand may influence the output of expert evaluators. The study provides support to the proposition of UX as a contextual-dependent response to the interaction with computing systems and has important implications for the design and evaluation of websites which are discussed in the conclusion.

1 Introduction

There is a pervasive move in HCI research and practice to consider *hedonic* aspects of user experiences alongside *pragmatic* qualities in the evaluation and design of interactive systems [1-3]. Responses such as emotions, aesthetic judgments and fun are among the most prominent aspects of experience investigated [4-6]. The User Experience (UX) is a complex response to the interaction with computing systems [1]. This response is a consequence of individual predispositions of the user (e.g., attitudes, motivations and needs), characteristics of the interactive system (e.g., purpose, functionality and usability) and contextual dependencies (e.g., task and environment). Thus, UX is elicited by an array of variables which combine to determine overall quality judgments [7].

The effect of brand on consumer attitudes towards real and virtual goods is well documented in consumer psychology and marketing. There are some evident links between a website design and its brand, yet, with the exception of studies of trust [8] and the emerging field of captology [9], brand has attracted little attention from the HCI community. In marketing research, the website is considered a powerful medium in forming and sustaining a positive attitude towards a brand and its products [10]. Empirical research has demonstrated that attitude toward a web site is a good predictor of consumer brand choice [11]. In this paper, we deal with the opposite research question, namely *Is attitude towards brand a good predictor of the user-experience with a website?*

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 638–651, 2009.

2 Related Work

Research on the UX evolved from classical studies of usability as technology moved out of the working context to pervade people's everyday life. The discourse initiated in the late '90s as researchers started considering how factors beyond usability might be integrated in design and evaluation. These considerations were rooted in the recognition of the importance of hedonic qualities [2], such as pleasure [12], fun [5], aesthetics [6], and emotion [4] in HCI.

2.1 UX Research

UX research shares some principles which differentiate it from traditional usability studies [1]. Usability research equated quality with absence of problems. UX instead focuses on *positive* outcomes of interaction and actively tries to add extra value to HCI by designing for emotion, fun and personal fulfilment. Hence, compared to usability, UX research takes a more holistic approach to the definition of quality. This definition subsumes a multi-dimensional model including both *pragmatic* qualities (traditional usability desiderata) and non-task related aspects, or *hedonic* qualities, such as beauty, challenge, stimulation and self-expression [2].

Another unique theme of UX is its emphasis on the subjective nature of experience. Traditional usability research emphasized the objectivity of evaluation methods and measures by favouring, for example, user observation over user opinion. Performance variables, such as error rate and time to complete a task, were deemed to be relatively stable among a sample of users with comparable level of computing knowledge, motor and cognitive skills. On the contrary, UX explicitly recognises variations across judgments and focuses on understanding the factors which modulate people's opinions. This theoretical and methodological difference is due to the different research object of the two frameworks. Usability focuses on pragmatic attributes which are "inextricably tied to internally generated or externally given behavioural goals and resulting tasks" [2]. Hedonic attributes, on the other hand, relate to the user's inner self which is, by definition, unique.

Experiences are contextually situated and unfold over time [4, 7, 13, 14]. They are the results of the encounters between an active user (with their own unique attitudes, expectations and feelings) and a system (with a set of technical and design characteristics). Individual sources of variations are particularly influent at initial encounter with a system. At this stage, the affect system judges the interface at a visceral level by assigning positive or negative valence to it [4]. As the interaction progress, cognitive elaborations are performed assigning meanings to interface characteristics through interpretation and rationalisation. The affect and the cognitive system are interrelated, so that some affective states are driven by cognition, while cognition can be influenced by affect. Emotions affect cognitive processing by inducing confirmation bias. Positive emotions could cause "mental processes to be more creative more tolerant of minor difficulties" [4] thus potentially favouring evaluation of a system. Negative emotions, generated by a bad first impression, may lead to negative judgement even in the presence of disconfirmatory evidence.

2.2 Brand

A brand is a name, word or symbol which identify the goods or services a vendor and differentiate them from those of the competitors [15]. The value ascribed to brands by consumers is referred to as consumer-based brand equity, i.e., the added value of a product in comparison to an otherwise identical product without the brand's name [15, 16]. This value lies in the minds of the consumers. It is fuelled by what consumers have experienced, learned, and felt about the brand over time, rather than by objective facts about the brand. Hence, consumer brand equity is subjective and related to the personality, history and preferences of individual customers.

Strong brand equity contributes towards the key components in achieving competitive advantage both in real-life [15, 16] and on-line [17]. In the digital economy, a considerable research effort has been devoted to understanding how brand equity in the real market extends to the digital word [10]. Key features investigated in this line of research are trust and the importance of website design in forming and sustaining a positive attitude towards a brand and its products. Results suggest that branding can facilitate consumers' acceptance of electronic commerce.

3 Method

The study investigated the role of brand on the evaluation of the UK-portal for Nike and Reebok. Following the contextual conceptualization of UX, we hypothesized that *brand attitude will affect quality judgment of website properties.* In particular, we expected that people holding positive attitudes towards the brand of the website would be more positive in their evaluation of the website design. The effect should be *stronger with respect to hedonic attributes*, which are more susceptible to individual preferences, *than to pragmatic attributes* which should be more objective.

3.1 Participants

A total of 184 second-year students at the University of Manchester participated in this study as part of their course-work for an advanced HCI course. All students had successfully passed an introductory HCI course in their first year. The sample was mainly composed of males (86%) between the age of 19 and 25 (mean 20.5); 105 students evaluated the Nike website, 79 the Reebok website.

3.2 Procedure

Prior to the evaluation, a pre-test questionnaire elicited demographic data and a measure of attitude towards the brand of the website the participants had to evaluate (Nike or Reebok). Participants were then given written instructions on the evaluation procedure and an experimental booklet they had to fill in reporting usability problems, as well as their opinions and impressions on the website. Students were given 4 weeks to perform the evaluation. They were instructed to familiarise themselves with the website, visiting as many sections as possible. To control for a minimum range of common browsing across all participants, they were given a set of information retrieval tasks and had to report their answers in the evaluation booklet. Tasks were designed to cover a cross-section of each website sharing some similar features. They included the online shop, the women's section, and the most interactive section of each website. The live websites were evaluated in May 2006. They did not change significantly within the evaluation period.

3.3 Design

The study adopted a between-subjects design contrasting brand attitude (3) and website (2). Brand attitude was measured by a pre-test questionnaire. Participants were clustered in three categories: negative, neutral and positive attitude. Participants were also randomly assigned to evaluate either the Nike or the Reebok website.

3.4 Instruments

Brand attitude was measured by 3 items ("XX is good value for money"; "XX is a well respected brand"; "People like me wear XX"). Reponses were modulated on a Likert scale, from 1 = strongly disagree to 5 = strongly agree. The Perceived Visual Aesthetic scale, PVA was used to provide a measure of *classical* and *expressive* aesthetics [6]. Classical aesthetics refers to traditional aesthetic notions emphasising orderly and clear design. It includes items such as "pleasant", "clear", "clean", "symmetrical" and "aesthetic design". Expressive aesthetics is characterised by qualities that capture the user's perception of creativity and originality of the site's design. Relevant items in this dimension are "creative", "fascinating", "original", "sophisticated design", and "use of special effects". The scale also includes a dimension of perceived *usability* and *service quality*.

In addition, participants filled in the Interface Quality Scale IQS, a 13 item Likert-scale developed by the authors to obtain a measure of *usability*, *content* and *pleasure* in interaction. The statements addressing usability were based on the ISO definition. The statements addressing content quality were based on the elaboration of selected items from the Bernier Instructional Design Scale scale [18]. The statements addressing pleasure were developed by the authors and pilot tested in previous research. The wording of the items is reported in Table 1. Both the PVA and IQS scale were modulated on 7 points (1 = strongly disagree; 7 = strongly agree).

Overall preferences were assessed directly by a 7-point semantic differential item asking users to rank the goodness of the website (good-bad) and 4 Likert-scale items inquiring about future behavioural intentions. Participants were asked to indicate whether "in the next future" they intended to (*a*) "buy from the website"; (*b*) "recommend the website to my friend"; (*c*) "revisit the website" and; (*d*) "use the website as inspiration for my own design".

3.5 Websites

The Nike and the Reebok websites were professionally designed and localised for the UK market. They were composed of a variety of sub-sections that showcased product lines of the two companies and included e-commerce features. The Reebok website was designed in a more traditional style with some animated and interactive elements only in selected sections. The Nike website, on the other hand, featured several animations and some level of interactivity across the entire website.

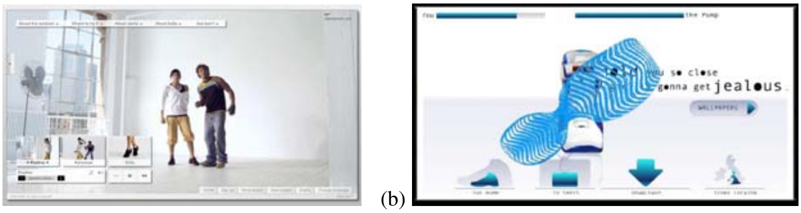
(a) (b)


Fig. 1. The Nike's Women's section (a) and the Reebok's Pump shoe section (b)

The Homepage of the Nike website displayed the main website sections. Tasks concentrated on the "Nike Id – Customise and Buy", "Football", and "Women's" sections. The Nike Id section was an online store which allowed the user to design their own piece of footwear or clothing (e.g., changing the colour of shoes or adding initials on shirts) and to buy the item. A range of standard, non-customised items was also available but it was featured less prominently. The Women's section showed a performer doing dance moves as part of a workout routine (Figure 1a). The user could view different clips and browse information on the merchandise featured in the workout. The Football section featured user-generated videos of amateur and semi-professional players doing football tricks. Users were invited to browse and view the videos, vote for their favourite actor in a league-system, or upload their own videos, although generating content was not mandatory.

The UK portal of the Reebok website consisted of various sections that featured distinct product-lines. Tasks required visiting the online store, "Jamelia's" women's fashion section, and the "Pump shoe" section. The online store featured a traditional HTML design. Jamelia's section displayed videos of the musician and information on where to buy the merchandise featured in them. The pump shoe section was highly interactive (Figure 1b) and included a game where the user was to fight the shoe by moving the mouse and clicking to invoke fight moves. Users could also watch videos, download web art, or locate a store.

4 Results

The reliability analysis on the three pre-test items measuring brand attitudes returned a Cronbach's alpha of .68. This value was considered satisfactory for such a short scale as all corrected item-total correlations were above .48. The *brand-attitude index* was thus computed averaging individual scores on the three items. The index was normally distributed, ranging from 1.33 to 5 (mean = 2.75, std dev = .79). Participants were divided into 3 equal-size groups: positive attitude (highest thirdile of the distribution), neutral (middle thirdile), and negative (lowest thirdile). This ordinal index was used as between-subjects factor in the Anovas. The original interval index was used as predictor in the regression analysis.

A *behavioural intention index* was computed averaging scores on the 4 items enquiring about future behaviour (Cronbach's alpha = .85). *Mean scores* for all dimensions elicited by the PVA (classical and expressive aesthetics, usability, and service

Table 1. Factor loadings for the IQS scale

	Factor		
	Content	Usability	Pleasure
The level of detail of the content is good	**.816**	-.171	.140
The right amount of content is provided	**.709**	.076	-.023
The content is relevant	**.559**	.099	-.103
The content is of good quality	**.502**	.119	.124
The website is easy to use	.023	**.912**	-.079
I feel in control when I am using this website	-.027	**.807**	.033
The website requires little effort to use	-.030	**.763**	.058
Using the website is effective	.221	**.638**	.091
I feel pleasure interacting with the website	-.165	.038	**.923**
The website is pleasurable to look at	.070	-.055	**.737**
The website has design features I like	.145	-.020	**.692**
The website evokes positive feelings	.020	.097	**.641**
The website is trustworthy	.208	.159	.193

quality) were also computed (Cronbach's alpha > .80). Scores on the IQS items were entered into an exploratory factor analysis for each site separately and using the combined data set. The factors were extracted applying the Maximum Likelihood method with Oblimin rotation, as it is acknowledged that different dimensions of the user experience tend to correlate [3, 6, 13]. The three analyses reported consistent factorial configurations. Loading of variables on factors for the entire sample are reported in Table 1. To simplify interpretation, variables are ordered and grouped by size of loading and loadings above .50 are reported in bold. Inspection of the scree-test plot and comparison of the goodness of fit of different configurations allowed the identification of three main factors which could clearly be interpreted as *usability*, *content* and *pleasure*. They accounted for 65% of the variance (goodness of fit: $\chi^2_{(42)} = 75.7$, $p < .01$).

To summarise, 7 variables were tested in the study: classical and expressive aesthetics, usability and service quality (collected by the PVA scale), and usability, pleasure and content (collected by the IQS). Correlations between mean scores on each dimension are reported in the lower part of Table 2. With the exception of the relationship between expressive aesthetics and the two usability indexes, all other variables significantly correlated with each other. To account for these correlations, *factor scores* were computed for each dimension using the Anderson-Rubin approach [19]. Factor scores are weighted combination of scores on a series of factors which provide an estimate of the scores participants would have received on each factor if they would have been measured directly. The Anderson-Rubin approach produces scores which are uncorrelated with each other even if the original scores are correlated. By taking away the shared part of variance between factors, factor scores provided less biased measures of the 7 UX dimensions. The correlations between factor

Table 2. Correlation between mean scores (lower part) and factore scores (upper part) on UX dimensions (* = p < .05; ** = p <.01)

	1	2	3	4	5	6	7
1. Content IQS		.000	.000	.085	.274 (**)	.279 (**)	.078
2. Usability IQS	.494 (**)		.000	.191 (*)	.184 (**)	.720 (**)	.636 (**)
3. Pleasure IQS	.421 (**)	.365 (**)		.299 (**)	.611 (**)	.103	.075
4. Classical Aesthetics PVA	.427 (**)	.483 (**)	.614 (**)		.000	.000	.000
5. Expressive Aesthetics PVA	.302 (**)	.050	.618 (**)	.393 (**)		.000	.000
6. Usability PVA	.520 (**)	.798 (**)	.372 (**)	.604 (**)	.066		.000
7. Service Quality PVA	.499 (**)	.636 (**)	.334 (**)	.427 (**)	.251 (**)	.651 (**)	

scores are reported in the upper part of Table 2. The significant correlations between similar dimensions measured by different scales provide support to validity of the measures.

Each mean and factor scores were entered as dependent variables into a univariate Anova with Website (2) and Brand-Attitude (3) as between-subjects factors. When the two analyses returned similar results, only those obtained on factor scores are reported as they are more conservative. Otherwise, the difference is acknowledged. For ease of interpretation, however, all descriptive statistics are computed on mean scores. Partial eta squared statistics (partial η^2) were used as estimate of effect sizes. Partial η^2 are computed considering the variance attributable to the effect of interest plus error [19]. As a general guideline, η^2= .01 are considered small, η^2= .06 medium, η^2= .14 large [20]. Post Hoc tests for the three levels of the brand attitude factor were computed using the Least Significance Difference algorithm.

4.1 Usability

Significant results returned from the two Anovas using usability factor scores as dependent variables are summarized in Table 3. Both analyses showed significant, medium-sized effects for website. On average, the Reebok website was evaluated as more usable than the Nike one. The effect of brand attitudes was also significant in both analyses, although the effect size obtained on the PVA scale was stronger than that obtained on the IQS scale.

Post-hoc analyses revealed that the effect of brand attitude was mainly due to participants who held negative attitudes towards the brand, as they were more critical in the evaluation (Figure 2). No significant differences emerged between people with neutral and positive attitudes, although the linear trend was still evident.

Table 3. Anovas results for usability factor scores

Dependent variable	Source	df	F	Sig.	η_p^2
Usability PVA	Brand Attitude	2,142	7.39	.001	.09
	Website	1,142	13.52	.001	.09.
Usability IQS	Brand Attitude	2,172	2.72	.07	.03
	Website	1,172	14.24	.001	.07

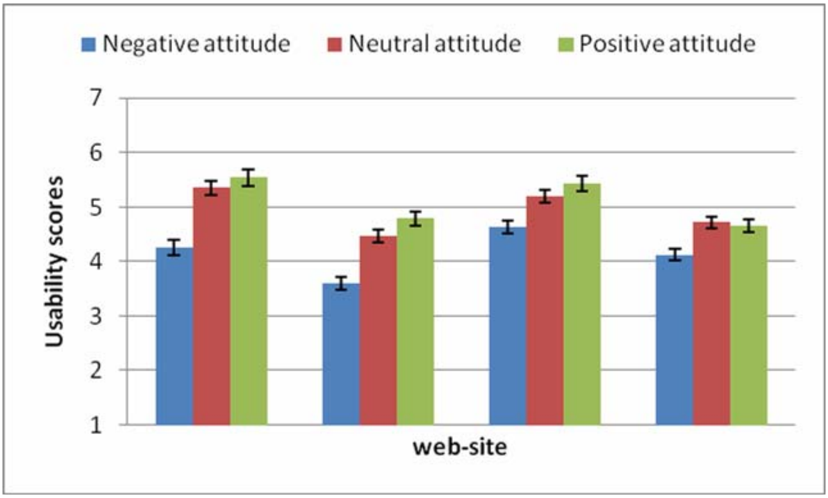


Fig. 2. Mean and standard error scores on usability as a function of experimental conditions

4.2 Aesthetics

The Anova using factor scores on classical aesthetics as dependent variable returned only a significant, medium-sized effect for brand predisposition ($F_{(2,142)} = 4.57$, $p < .05$, partial $\eta^2 = .06$). A linear trend was evident. Participants holding a negative attitude towards the brand evaluated the web site as less attractive (mean = 4.50; std error = .12), than participants holding a neutral attitude (mean = 4.98; std error = .12), or participants holding a positive attitude (mean = 5.31; std error = .14). Post-hoc tests showed significant differences between participants holding negative attitudes and the other two groups.

The Anova with factor scores on expressive aesthetics as the dependent variable indicated a significant effect for web-site ($F_{(12,142)} = 31.7$, $p < .001$, partial $\eta^2 = .18$) and a significant interaction web-site * attitude ($F_{(2,142)} = 4.12$, $p < .05$, partial $\eta^2 = .06$). Nike was clearly preferred over Reebok on this aesthetic dimension. Participants holding better attitudes towards Nike were also more positive in the evaluation of expressive aesthetics; whereas the evaluation of Reebok was unaffected by participants' attitudes towards the brand.

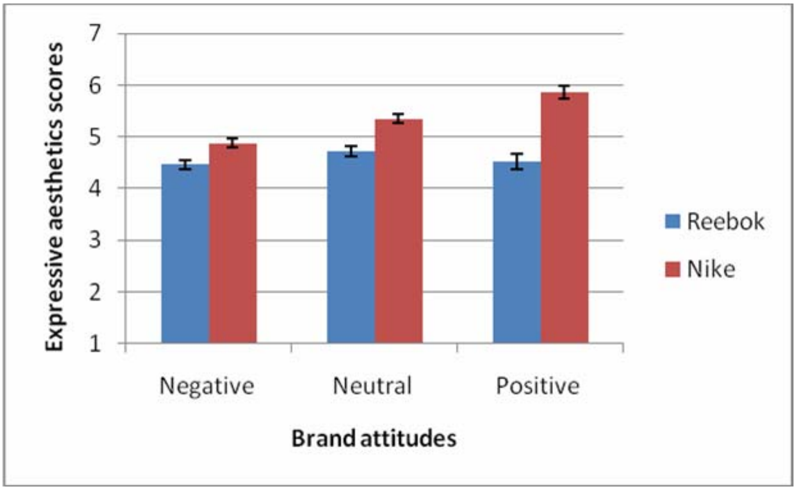


Fig. 3. Mean and standard error scores on expressive aesthetics as a function of experimental conditions

4.3 Pleasure

A significant main effect for website ($F_{(2,172)}$ = 19.27, p < .001, partial η^2 = .12) and for brand affinity ($F_{(2,172)}$ = 6.87, p < .01, partial η^2 = .07) were returned by the Anova on pleasure factor scores. Overall, participants tended to prefer the Nike website (mean difference = 0.55) and their evaluation was modulated according to their attitudes towards the brand: the most positive the attitude, the most positive the evaluation. In particular, post-hoc tests revealed that people holding negative attitudes were significantly different in their evaluation from people holding neutral or positive evaluations. No significant difference emerged between the two latter groups.

4.4 Content

The Anova using factor scores on content as dependent variable returned a marginally significant effect for brand attitude ($F_{(2,172)}$ = 2.61, p =.07, partial η^2 = .03). The mean trend is consistent with previous analyses showing a linear improvement in the evaluation as a function of attitudes. However, post-hoc test indicated that the effect is significant only for the comparison between negative and positive attitudes.

The effect of brand attitude cleanly appeared when the mean scores were analysed ($F_{(2,172)}$ = 7.11, p < .01, partial η^2 = .08). All differences between consecutive levels of attitudes were significant. The difference in the two analyses was probably due to the strong correlation between content and usability (Table 2).

4.5 Service Quality

No significant results were returned from the factor scores on service quality. On the contrary, the Anova on mean scores indicated a large significant effect for brand attitude ($F_{(2,156)}$ = 10.99, p <.001, partial η^2 = .12) and a small effect for website

($F_{(2,172)}$ = 4.89, p < .05, partial η^2 = .03). These effects indicated that people holding negative attitudes were more critical in evaluating the quality of the website than people with a neutral or a positive attitude. On average, the service quality of Reebok was scored slightly better that that of Nike (mean difference = .33).

4.6 Overall Evaluation

The overall evaluation of website goodness was affected by brand attitudes ($F_{(2,178)}$ = 6.89, p <.001, partial η^2 = .08). People holding negative attitudes towards the brand evaluated the website more negatively than people holding neutral or positive attitudes.

4.7 Behavioural Intentions

A stepwise regression analysis was run to investigate the relative influence of different UX dimensions in predicting participants' behavioural intentions with the website. Using this technique, significant variables are entered in the regression model according to their relative contribution to it. The entire sample was used for the analysis. Factor scores on classical and expressive aesthetics, usability (measured by the PVA scale), content and brand attitudes (interval scale index) were entered as predictors. Four significant models were returned, p < .001 (Table 4). Brand attitude (BA), expressive aesthetics (EA), usability (U) and classical aesthetics (CA) together accounted for 46% of the variance in behavioural intention with the website (model 4).

Brand attitude was the most important predictor of behavioural intentions, accounting for 22% of its variance. The second variable entered in the model was expressive aesthetics which added another 13% of variance. Usability accounted for 7% and classical aesthetics for 5% of the variance.

Table 4. Significant models returned from the stepwise regression analysis on behavioural intentions

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1. (Constant), BA	.473	.224	.218	1.48946
2. (Constant), BA, EA	.589	.347	.338	1.37073
3. (Constant), BA, EA, U	.646	.417	.405	1.29918
4. (Constant), BA, EA, U, CA	.686	.471	.456	1.24270

5 Conclusion

The objective of this study was to assess the effect of brand attitudes on the evaluation of websites. For this purpose we have tested a large sample of evaluators asking them to rate the quality of two corporate websites (Nike and Reeebok) with an array of UX

Table 5. Summary of significant effects from the Anovas on factor scores

UX dimension	Brand Attitudes	Website	Interaction
ability	Negative < Neutral = Positive	R > N	n.s.
Classical Aesthetics	Negative < Neutral = Positive	R = N.	n.s.
Expressive Aesthetics	Negative < Neutral = Positive only for Nike	R < N	Yes
Pleasure	Negative < Neutral = Positive	R = N.	n.s.
Content	Negative < Positive	R = N	n.s.
Service quality	n.s.	R = N	n.s.
Overall goodness	Negative < Neutral = Positive	R = N	n.s.

questionnaires. Results suggested that the evaluation was influenced by participants' previous knowledge and expectations of the website brand. A summary of significant effects returned from the analyses of variance conducted on the factor scores on each individual UX dimension is reported in Table 5.

Attitude towards the brand proved to be an important moderator of judgments, affecting not only the perception of hedonic qualities (classical aesthetics, expressive aesthetics and pleasure) but also the perception of some pragmatic qualities (usability and in part content), which are generally considered to be more objective. A linear trend appeared in all dimensions: evaluation improved alongside attitudes. The post-hoc analyses suggested however that, with the exception of content, the difference in all dimensions was due to participants holding negative attitudes who were more critical in their evaluation than both participants holding neutral or positive attitudes. This is consistent with the 'negativity bias' widely documented in psychological research implying that people's reaction to negative stimuli evoke stronger and more rapid physiological, cognitive, emotional, and social responses than neutral or positive stimuli [21].

With the only exception of expressive aesthetics, which was affected by the interaction between brand attitudes and website, the effect of brand was reliable across the websites. Overall, Nike was evaluated better on the expressive aesthetics dimension, whereas Reebok was favoured on usability and content. No significant differences emerged in the overall evaluation of the website goodness. The trade-off between expressive aesthetics and usability was shown in more controlled experimental settings suggesting that engaging design tend to add a level of complexity to the detriment of ease of use [3, 7].

5.1 Brand and UX

The effect of brand on the user evaluation of web-sites supports the prevailing conceptualization of UX stating that experiences are intrinsically subjective and dependent on the evaluator personality, expectations, and goals [2, 7, 14, 15]. No conclusive evidence can be derived from this study to understand how exactly brand affects the evaluation of websites. Theoretically, this effect may be elicited at different processing levels [4]. At the initial encounter with the website, brand may influence the output of the affect system, acting as a mental determinant which induces a visceral reaction to it. This reaction can subsequently bias the output of the cognitive system with a general halo effect. The importance of first impressions on the evaluation of websites has been demonstrated in previous UX studies. Participants reliably decided which design they liked and which ones they did not like within 500ms exposure [22] and these impressions were not transient [23]. Furthermore a large corpus of psychological research has provided evidence that even subliminal exposure to a brand may affect consumer mood and choice [24].

However, brand can also influence cognitive processing directly as an attitudinal variable. In this case brand can be used as a criterion signalling trust, identification and reliability. We believe that the specific nature of our sample target (university students with a background in HCI and previous experience in evaluation) should make them less prone to evaluation bias, and may suggest a direct effect of brand on cognitive elaboration of judgements.

5.2 Limitations and Future Research

The study reported in this paper has several limitations which may have affected its reliability. An obvious limitation regards the nature of the sample. More research is needed to understand how these findings extend to different user categories, including usability professionals and final users. Similarly, the study should be replied using a larger set of web-sites as stimuli and more robust measures of attitudes towards brand. Despite of these limitations, the large sample tested in the study and the strong reliability of the effect which persists across different measures and operational definitions of UX dimensions, is an encouraging indication of external validity.

The choice of favouring an ecological setting over a laboratory-based experiment implied that results may be affected by confounding factors. For example, we had no information about the browsing behaviour of the evaluators, or about the environmental setting where the evaluation was carried out. Furthermore, as the evaluation addressed two live websites we did not have direct control on their content. These problems were, at least partially, counteracted by taking into consideration a number of explicit and implicit information provided by the evaluators in their reports. We checked the quality of the reports and discarded those students who appeared not to have put enough effort into the assignment. Furthermore, participants were motivated in performing the evaluation to the best of their capabilities, as the assignment accounted for 20% of their mark at the module.

The study presented in this paper has important implications for UX practice and research. On a practical point of view, it warns evaluators and designers of the possibility of a systematic bias in the evaluation affecting not only hedonic qualities but also usability. On a research perspective, this study calls for a re-examination of current UX models assuming that experiences are composed of distinct and orthogonal attributes. Tractinsky & Zmiri [25], for instance, highlighted three attributes: usability, aesthetics, and symbolism which were considered to be independent of each other. In their model, aesthetics is regarded as the sensorial experience generated by the look and feel of the interface. Symbolism refers to meanings and associations which are elicited during cognitive elaborations. Our study provides support to the importance of symbolism in UX but it questions its conceptual independence. Both UX research and practice must seriously consider the operational definition of quality dimensions, as we have demonstrated that mean scores, traditionally used as indications of user evaluations, are prone to over-estimation of statistical effects.

Acknowledgement

We are grateful to Mark Hassenzahl for his suggestions on the statistical analyses reported in this paper and the many stimulating discussions on validity in HCI research.

References

1. Hassenzahl, M., Tractinsky, N.: User experience - A research agenda. Behaviour & Information Technology 25, 91–97 (2006)
2. Hassenzahl, M.: The Interplay of Beauty, Goodness, and Usability in Interactive Products. Human-Computer Interaction 19, 319–349 (2004)
3. De Angeli, A., Sutcliffe, A., Hartmann, J.: Interaction, usability and aesthetics: what influences users' preferences? In: 6th conference on Designing Interactive systems DIS2006, pp. 271–280. ACM, State College (2006)
4. Norman, D.A.: Emotional Design: Why We Love (or Hate) Everyday Things. Basic Books, New York (2004)
5. Blythe, M.A., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.): Funology: From Usability to Enjoyment. Kluwer Academic Publishers, Dordrecht (2005)
6. Lavie, T., Tractinsky, N.: Assessing dimensions of perceived visual aesthetics of web sites. International Journal of Human-Computer Studies 60, 269–298 (2004)
7. Hartmann, J., Sutcliffe, A., De Angeli, A.: Towards a theory of user judgment of aesthetics and user interface quality, vol. 15, pp. 1–30. ACM Press, New York (2008)
8. Briggs, P., Burford, B., De Angeli, A., Lynch, P.: Trust in online advice. Soc. Sci. Comput. Rev. 20, 321–332 (2002)
9. Fogg, B.J.: Persuasive technology. Morgan Kaufmann Publishers, San Francisco (2003)
10. Chaffey, D., Ellis-Chadwick, F., Johnston, K., Mayer, R.: Internet Marketing - Strategy, Implementation and Practice. Prentice Hall, Essex (2006)
11. Byung-Kwan, L., Ji-Young, H., Wei-Na, L.: How Attitude Toward the Web Site Influences Consumer Brand Choice and Confidence While Shopping Online 9 (2004)
12. Jordan, P.W.: Designing Pleasurable Products: An Introduction to the New Human Factors. Taylor & Francis Ltd., London (2000)
13. Hartmann, J., Sutcliffe, A., De Angeli, A.: Investigating attractiveness in web user interfaces. In: CHI 2007, pp. 387–396. ACM Press, New York (2007)
14. Hartmann, J., De Angeli, A., Sutcliffe, A.: Framing the User Experience: Information Biases on Website Quality Judgement. In: CHI 2008, pp. 855–864. ACM, Florence (2008)
15. Aaker, D.A.: Building strong brand. Free Press, New York (1996)
16. Keller, K.L.: Conceptualizing, Measuring, Managing Customer-Based Brand Equity. Journal of Marketing 57, 1–22 (1993)
17. Ward, M.R., Lee, M.J.: Internet shopping, consumer search and product branding. Journal of Product & Brand Management 9, 6 (2000)

18. Bernier, M.J.: Establishing the psychometric properties of a scale for evaluating quality in printed educational materials. Patient Education and Counseling 29, 283–299 (1996)
19. Tabacknick, B.G., Fidell, L.S.: Using multivariate statistics. Pearson International Edition (2007)
20. Pallant, J.: SPSS Survival Guide. Allen & Unwin (2007)
21. Taylor, S.E.: Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. Psychological Bulletin 110, 67–85 (1991)
22. Lindgaard, G., Fernandes, G., Dudek, C., Browñ, J.: Attention web designers: You have 50 milliseconds to make a good first impression! Behaviour & Information Technology 25, 115–126 (2006)
23. Tractinsky, N., Cokhavi, A., Kirschenbaum, M., Sharfi, T.: Evaluating the consistency of immediate aesthetic perceptions of web pages, vol. 64, pp. 1071–1083. Academic Press, Inc., London (2006)
24. Theus, K.T.: Subliminal advertising and the psychology of processing unconscious stimuli: A review of research. Psychology and Marketing 11, 271–290 (1994)
25. Tractinsky, N., Zmiri, D.: Exploring attributes of skins as potential antecedents of emotion in HCI. In: Fishwick, P. (ed.) Aesthetic Computing, pp. 405–422. MIT Press, Cambridge (2005)

Does Branding Need Web Usability? A Value-Oriented Empirical Study

Davide Bolchini[1], Franca Garzotto[2], and Fabio Sorce[2]

[1] Indiana University, School of Informatics (IUPUI) – 535 W. Michigan St.
Indianapolis, IN 46202-3103 (U.S.A.)
dbolchin@iupui.edu
[2] HOC Lab – Department of Electronics and Information, Politecnico di Milano
Via Ponzio 34/5, 20133 Milano (Italy)
{franca.garzotto,fabio.sorce}@polimi.it

Abstract. Does usability of a web-based communication artifact affect brand, i.e., the set of beliefs, emotions, attitudes, or qualities that people mentally associate to the entity behind that artifact? Intuitively, the answer is "yes": usability is a fundamental aspect of the quality of the experience with a website, and a "good" experience with a "product" or its reifications tends to translate into "good" brand perception. To date, however, the existence of a connection between web usability and brand perception is shown through anecdotic arguments, and is not supported by published systematic research. This paper discusses a study that empirically investigates this correlation in a more rigorous, analytical, and replicable way. Our main contribution is twofold: on the one hand, we provide empirical evidence to the heuristic principle that web usability influences branding, and we do that through four between subjects controlled experiments that involved 120 subjects. On the other hand, we inform the study with a systematic value-oriented approach to the user experience, and thus provide a conceptual framework that can be reused in other experimental settings, either for replicating our study, or for designing similar studies focusing on the correlation of web branding vs. design factors other than usability.

Keywords: brand, usability, value-centered approach, web application, communication goal, empirical study, inspection.

1 Introduction

For companies, educational or cultural institutions, charities, governmental bodies, politicians, artists, and many other subjects, the web has become one of main channels not only to inform and offer services, but also to build a relationship with stakeholders, and influence their attitudes and behavior [30][35]. In other words, the web is more and more a component of an ubiquitous and pervasive network of communication, interaction and information that aims at influencing the minds of people and ultimately, creating or strengthening a *brand* in the global society.

It is well known in brand design that a "good" direct experience with a "product" or one of its reifications translates to a "good" brand perception [18][34][37][41].

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 652–665, 2009.

Since usability is a fundamental aspect of the quality of the experience with a website [1][2][29], we may well accept the thesis that *usability* somehow affects *brand perception*. Still, the existence of such correlation is usually discussed through anecdotic evidence, intuition, common sense, or design experience [16][23][28][38][26], but lacks support by publicly available systematic research.

In the context of a wider research action aimed at understanding the role of branding in the web design process [3][4][43], this paper attempts to investigate the relationship between *brand perception* and *web usability* in a more rigorous way. We present a study that involved 120 participants and was designed as *four* controlled between subjects experiments, each one considering a different website of a large company or institution, and its modified version in which we altered the usability for better or for worse. Our results show a statistically significant difference in brand perception when users are exposed to websites sharing contents and functionality but having different degrees of usability, thus providing some empirical evidence that usability *does* affect the strengthening (or weakening) of brands.

Our work also offers a novel contribution from a *methodological* perspective, and this consists of:

(a) adopting a structured *value-oriented evaluation approach* for operationalizing and measuring brand from a user-experience perspective (i.e., not from a marketing or economical perspective [17][42][36][31]);

(b) defining a *conceptual evaluation framework* that can be *reused* in other experimental settings, either for replicating our study or for designing similar studies focusing on the correlation of web branding *vs.* design factors other than traditional usability (e.g., related to aesthetics or attractiveness [9][10][21]22][24][25][27][32]).

The rest of the paper is organized as follows. After a survey of the related state of the art, we introduce the general evaluation approach and describe in details our empirical study. A discussion of our findings and some general lessons learned concludes the paper.

2 Background and Related Work

The notions of brand and its many declinations (e.g., branding and brand experience) are quite broad, and are defined in many different ways in the current literature on design, marketing, and advertising [16][18][31][36][41][42]. For the purpose of this paper, *branding* is the intentional process aimed at creating or promoting a brand, while *brand* defines the core messages designed around "who we are", "what we believe", "why you should trust us". The term brand refers therefore to the expected cognitive and emotional associations that people make with an "entity" (being it a product, a service, a company, an institution, a person or, at a broader level, a country or a culture). We can also say that a brand is a *promise of values* [41] that the entity can keep to all its stakeholders - customers, trades, stockholders, employees, fans, or supporters. As a person gets in touch with or lives by any concrete reification of a brand, a *brand experience* takes place. It consists of all the perceivable elements (communication artifacts, physical products, people, services, or events) that give to the customer the experiential, comprehensive feeling of the brand. A brand experience is what creates not only a functional, but also an *emotional* relationship with a product or service *and* the brand underneath it [26][28].

A number of studies pinpoint that when individuals have an *active* or *interactive* experience with a brand reification, e.g., shopping in a branded store, buying a service or searching for information in a branded web site, this has a greater affect on brand perception than an *indirect*, *passive* exposition to the brand such as in traditional advertising, because people tend to assign more weight to emotions emerging directly from their proactive experience [34][17][32].

Based on this general idea, several authors suggest that the quality of a web experience affects users' brand perception [2][16][18]. In particular, *usability* is held to play an importance in positively or negatively influencing the user's attitude towards the entity behind the website, as it is a fundamental factor in determining the quality of the user experience [2][16]. This principle is also illustrated in online articles and practitioner's guidelines by J. Spool [37][38], whose usability studies claim to show that a more successful user experience is created when website help users achieve their goals, which in turn translates into an improved brand perception. This study, however, like others often quoted in the practitioners' design literature, seems to be proprietary knowledge, whose data, methods and instruments are not widely shared with the research community. Little o no data are published on this topic. In most of the works that have examined how to design to maximize the effectiveness of online brand communication, and how usability may facilitate or hinder this process, arguments are at most anecdotic, and show little or no empirical data or systematic experiments.

Interestingly, one of the few published data sets on usability vs. branding, empirically shows an influence in the opposite direction, i.e., brand perception affects the *perceived* usability, and experimentally illustrates that (in mobile applications) the users' affection to a brand accounts for an improved level of perceived usability, despite the absence of significant differences in task performance [33].

Some empirical work explores the correlation between web *usability and other quality factors of the user experience*. The study reported in [9], for example, investigates the relationships between content, presentation, usability and memory, and their relative importance to the user's preferences. Moreover, attractiveness and aesthetic design are key factors in persuasive computing [13]. A number of studies have shown correlations between the perceived aesthetic quality of a system's user interface and overall user satisfaction [22][24][25][27], leading to claims that aesthetic design can have a stronger influence on users' attitude towards the system than traditional usability [39][40]. As to vertical domains, there are studies that examined the design factors that affect consumer's trust in e-commerce websites [15][38], or influence people's perception of the website credibility [11][12].

None of these studies explicitly investigated the effects of usability on the users' perception of the *entity behind the web system*, thus they do not consider brand perception. Still, most of the above mentioned works have an important role in the research area of branding in relationship to user experience design, and in particular can pave the ground to theoretically framing the results of our study from a wider perspective.

Methodologically, the study presented in this paper systematically adopts a *value-oriented* approach, and is the logical follow-up of our previous works in web branding, communication requirements analysis and evaluation [3][4] [43]. In [3], we introduced the concept of communication goals for web applications, framed it in the context of goal-based requirements engineering, and modelled its impact on the various design dimensions. We have started systematically addressing the problem of elaborating requirements to inform *how* brand values are communicated through design. Then, thanks to extensive project experience and work with design teams, this initial study led to a refinement of our approach, reported in [4][43], pinpointing more precisely the relationships between communication requirements, branding, and design aspects.

All the above mentioned methodological results are centered around the concept of "value" - a notion that has been investigated in different communities including HCI, e-commerce, design, requirements engineering, and web engineering. *Values-Sensitive Design* [14], for example, stresses the importance of embedding value-aware requirements throughout any technology development, particularly emphasizing values with moral significance. *Value-Centered Design* [7][8] generalizes the notion of value by extending it to whatever a person finds worthwhile. In web engineering, *Value Based Design* (VBD) [19] provides a more pragmatic and systematic approach with respect to HCI design works, looking at the notion of "value as worth" from a strictly business perspective, in terms of the economic benefit that is induced by a web system.

For the purpose of the research reported in this paper, the notion of *value* is regarded not only as a driver for requirements and design of branded web applications, but it also provides the key to operationalize and measure brand perception into more precise and analytical elements, as explained in the next section.

3 Experimental Design

3.1 Scope

The scope of our research action are *content intensive branded web sites,* i.e., web applications that provide a large quantity of articulate content and services and at the same time need to sustain or to create a positive brand image. We purposefully do not address web sites that are mainly a marketing tool, and, like wall paper street advertisement or TV commercials, are almost exclusively designed to convey brand messages. These are typically websites with little or no informative content, but a lot of animations, gaming activities, or streaming media (videos): their main goal is not to inform and provide services, but to attract user's attention and provoke emotions.

Measuring brand in relationship with usability requires operationalizing the concepts of usability and brand.

For the precise scope of our study, we are interested to the "classical", *functional* view of web usability, which considers the operational ease of use and the degree of effective and efficient support to users' information and operational goals, according to the known ISO-9241 definition. We well acknowledge that current research explores aspects of the quality of experience (e.g. aesthetics, attractiveness, engagement, and affective dimensions) that goes much beyond a traditional notion of usability [10][1][29][32]. Still, the impact of these user experience dimensions on brand is outside the scope of this paper, although the methodological elements we propose provide a conceptual framework that can be also adopted for research on this issue.

It is well known in HCI how to measure “traditional” usability. In contrast, how to translate the general and vague concept of brand into lower level measurable factors from a user-experience perspective (i.e., not from perspective of the economic value of a brand) is still a largely unexplored issue.

As we define brand as a promise of values, we conceptualize *brand* in terms of *Brand Values*. A brand value can be a moral, ethical, social, or cultural *belief* which an entity is committed to, or a *quality of the entity* that is worthwhile for people, either at individual or collective level. This quality is not necessarily functional, but can be something that gives rise to positive, emotional or affective, effects. Brand values represents therefore those traits of an entity’s personality that the entity holds as salient to communicate to its target via the its products or communication artifacts (e.g., its web site) and that are expected to be perceived by the users. We then measure brand in terms of *Brand Values Perception,* i.e., the degree of people’s conscious awareness of and trust on the values promised by the entity or institution associated to the brand.

3.2 Hypothesis and Study Variables

In order to investigate the effect of usability on brand, we need to study whether the exposure to web sites having different levels of usability determine a *variation* of the users’ perception of brand values.

The study was designed as *four between subjects controlled experiments* (each one involving a different website), in which the independent variable is the *usability* of a website, manually manipulated with focused design interventions, and the dependent variable is the *difference in brand value perception* induced by the use of the web site.

In a nutshell (see Figure 1), the experimental design was set up in such a way to create, for each different web site, two experimental conditions: the original web site and a modified web site where we intervened on the design by improving or worsening the measured degree of usability, without altering content and functionality. Participants were randomly assigned to an experimental condition, and brand values perception in each group was assessed *before* and *after* a supervised use to the system.

The underlying hypothesis of the study is that *there is a statistically significant difference in the variation of brand values perception induced by the exposure to the two different experimental conditions* (an original web site and its version with modified level of usability). This hypothesis can be formalized as follows:

$$U(\mathrm{W}) > U(\mu(\mathrm{W})) \rightarrow B(\mathrm{W}) > B(\mu(\mathrm{W}))$$
$$U(\mathrm{W}) < U(\mu(\mathrm{W})) \rightarrow B(\mathrm{W}) < B(\mu(\mathrm{W}))$$
$$p(B(\mathrm{W}), \mathrm{B}(\mu(\mathrm{W})) < \alpha$$

Whereas: W is a web site, $\mu(\mathrm{W})$ is its *modified* version; $U(\mathrm{W})$ and $U(\mu(\mathrm{W}))$ are the usability respectively of the web site and its modified version; $B(\mathrm{W})$ and $B(\mu(\mathrm{W}))$ are the *variations of brand values perception* induced by the exposure respectively to the original web site and its modified version; p is the significance level on the increment or decrement of brand value perception B w.r.t. to a statistically critical value α.

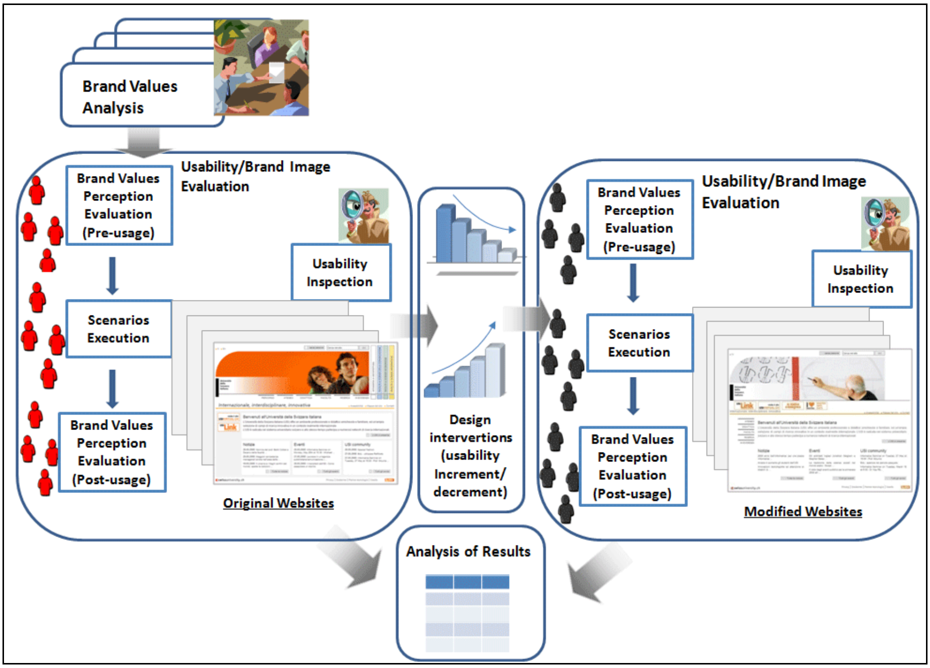


Fig. 1. Synopsis of the experimental flow used in the study

In the following subsection, we illustrate in detail the steps and the instruments of the experimental method put in place and replicated on four different websites, in order to ensure repeatability for future studies.

3.3 Instruments

Four websites (see Table 1) were selected randomly among a pool of pre-selected 100 websites - from different domains - meeting the criteria of being information-intensive (offering an articulate and rich set of content and services) and brand-intensive (designed as the intended reification on the web of the brand image of a consolidated company or institution).

For each website, two subject groups (each one composed by randomly selected subjects corresponding to the profile of potential user of the website) were recruited for the study and randomly assigned to either the original version of the website or the version with better or worse usability. *Only persons who never used the website under study were screened.*

Table 1 also shows the distribution of the subject groups across the four websites. Starting from the initial research on the first website (with 5 users), we progressively increased the number of participants in our studies, to provide more substantial evidence for the initial findings and corroborate the study framework. Overall, from March to June 2008 we involved a total of 120 users, distributed in 8 experimental groups.

Table 1. The web sites considered in our study

Website	URL	Sector	Code	Sample size (original version)	Sample size (modified version)
University of Lugano (Switzerland)	www.unisi.ch	Higher education	USI	5	5
Ryanair Italy	www.ryanair.it	travel	RYANAIR	15	15
Disney Italy	www.disney.it	entertainment	DISNEY	25	25
Municipal IV Circle of Verbania (Italy)	www.quartocircolo.vb.it	cultural center/ education	VERBANIA	15	15

3.4 Eliciting Brand Values and Salient Scenarios

We carried on a detailed analysis of the brand values of the entities behind each considered web site to determine the key messages the website is supposed to communicate to its various target audiences. The core question underlying this investigation is to gather traits or attributes describing the brand from the perspective of the company or institution and in relationship to the characteristics of their "customers": how would an entity want to be perceived by its target audience?

We followed the conceptual guidance provided by the value-driven requirements elicitation method illustrated in our previous work [3][4]. According to this method, brand value elicitation is a combination of business analysis, techno-organizational analysis, and user analysis. In our study, this activity was carried out through complementary value-based requirements elicitation techniques, namely: (1) background research and value extraction from explicitly declared brand image sources (institutional mission statements, press releases and investor relations public domain information); (2) semi-structured interviews (via phone or face-to-face) to both institutional stakeholders of a web site (e.g., the university Dean and Students Recruitment Manager in USI) and the "customers" of the entity underneath it (e.g., USI students).

The qualitative data collected with these instruments were consolidated and distilled in a list of (ten, on average) brand values for each website that were further validated with institutional stakeholders. Syntactically, a brand value takes the form of single statements defining a trait of the brand, as shown in Table 2.

Table 2. Excerpt from Brand Values (BV) List of RYANAIR

Ryanair	BV1	BV2	BV3	BV4	BV5	BV6	BV7	BV8
	Economic	Easy	Trustworthy	Safe	Credible	Growing	Competitive	Flexible

During this preliminary phase, we also identified an articulated *profile of the main target users* for each web site, including their information and operational needs, the motivations for using the application, and the context in which it is intended to be used. We exploited this knowledge to define a set of *activity scenarios* to be assigned to the users during the experiment. These scenarios capture important, high-priority web tasks for the user, well aligned with the overall mission of the application (e.g. planning a trip on an airline website, rather than looking for the curriculum in a given area on a university web site). Scenarios were custom for each system under study, and were designed to have the users explore and be exposed to the most salient content, service and interaction features offered, so that they could have a realistic experience of the web site as it would be in a spontaneous, "normal" condition.

3.5 Evaluating Usability

The level of usability for the original and modified websites was obtained by applying a *usability inspection* process based on *heuristics evaluation.* In particular, we adopted the MILE+ method [5], which is the latest of a set of conceptual tools for systematic usability inspection we originally developed for hypermedia and content-rich interactive applications [6], and leverages common practices in web usability engineering.

The inspection protocol of MILE+ offers a built-in library of heuristics, coupled with a set of operational guidelines that identify the inspection tasks that must be undertaken by usability experts. These heuristics decompose the general concept of usability into more measurable attributes, each one addressing a different web design dimensions: *navigation* (heuristics addressing the usability of the information architecture and navigation mechanisms), *content* (heuristics addressing the general quality of the information offered to the user), *layout* (heuristics addressing the semiotics of the interface and its graphical look), and *technology/performance* (heuristics addressing usability issues caused by technological defects).

Each website (both for the original and modified version) has been assigned a team of 3 inspectors, who have individually carried out a systematic heuristics evaluation based on MILE+. Each inspector gave a 1-5 score (level of compliance) to each of the MILE+ heuristics which were found to be applicable to the website under analysis. Then, as in conventional usability inspection, they came together to share and discuss the usability problems found and their severity, consolidate the results and finally converge their scores to an agreed level. At this point, the average heuristics score was calculated and was used as a final single measure for the usability of the website. The same evaluation procedure (using the same team of inspectors) was followed for the modified website. This evaluation enabled to have an overall consistent measure for the usability of website for the original version and for the modified version.

3.6 Treatment: Altering Usability with Design Interventions

We *improved* the usability of the 2 web sites (RYANAIR, VERBANIA) characterized by a low level of usability, and we intervened on the 2 websites (USI, DISNEY) that had a good or very good level of usability, to *worsen* them. In both cases, the actual content and services remained invariant. Examples are shown in Figure 2. The type of design changes on the original version was kept identical across each couple of applications under analysis. In particular, usability worsening involved the following changes: (a) *Navigation usability*: from 4 to 8 links relevant to the task scenarios at hand were removed or misplaced in the information architecture; (b) *Layout usability*: the layout template of the main pages relevant to task scenarios was twisted to lose perfect visual balance, and infringe expected prominence of the main menus; from 4 to 8 link labels relevant to the task scenarios at hand were made more obscure or more

Fig. 2. Homepages of 2 of the original and modified websites. Whereas for Ryanair the changes are visible from the screenshot, for USI the changes included moving the main horizontal main navigation as a left-side bar, moving and increasing the number of mini banners from the left side to the top horizontal bar, and removing clear spacers between the content boxes of the homepage content.

vague in meaning; (d) *Content usability*: content readability was worsened by removing spacing between the paragraphs and font size in text-intensive pages relevant to the task scenarios.

In case of design interventions aimed at improving the degree of usability, similar types of design interventions described above were applied, but in the opposite direction: giving better prominence to key links, clarifying labels, providing more visual balance and layout order, and improving the readability of the text.

3.7 Evaluating Brand Values Perception

To evaluate how the intensity of brand values perception is affected by the exposure to a web site, or its altered version, we proceeded as summarized below:

1. *Assessing pre-usage brand values perception*: a questionnaire (paper-based) was administrated to the users at the beginning of the evaluation (before using

the website). We ask them to express their level of agreement with the brand values statements associated to the brand (resulted from the elicitation activity described in 3.4). The aim of this assessment is to establish a starting level of brand values awareness. Since all users recruited never used the website under study, brand values perception captured at this stage derives from previous experience with any reifications of the brand but the website, e.g., traditional advertising, word of mouth, or from positive or negative preconceptions. The set of questions were articulated around a precise investigation focus: *how much does Brand X inspire you Value Y*? This was applied to each brand value elicited and the measurement was collected through a likert scale, indicating perception intensity (0: not perceived, 5: strongly perceived).

2. *Performing salient scenarios on the website:* after the pre-usage questionnaire, each subject was asked to use the website by following 3 assigned scenarios defined in the preliminary phase of user analysis. The average duration of a session of use was approximately 40 minutes.

3. *Assessing post-usage brand values perception*: right after the execution of the scenarios, the same questionnaire used in 1) was administrated to the users, with the aim of collecting their actual perception of the brand values *after* the experience with the website, i.e., the short-term "habit change" in brand awareness due to the exposition to the website.

4 Results

The preliminary result is that there is a statistically significant difference ($p<0.01$) between the brand values perception *before* and *after* the use of the websites. This was calculated by running a pair t-test for the results of the pre- and post-usage questionnaires for all the experimental conditions. This shows that the exposition to the website indeed modifies a short-term perception of the brand image.

As to the core of our study, our hypothesis states that there is a significant difference in brand values perception as the *usability* of the website varies. A paired t-test on the brand values perception of the experimental and the control groups also shows that the actual difference in perception is statistically significant ($p<0,01$, see Table 3).

As highlighted also in Figure 3, our data indicate that subjects who used a website with more (resp., less) usability problems tend to express a consistently less (resp. more) positive perception of the brand values, with respect to the subjects who used a

Table 3. Usability and Brand Values Perception variation in the 4 experiments

	Usability		**Variation in Brand Values Perception**					
	original	*modified*	original	*modified*	variation	sample size	*T*	*Signif. level*
USI	3,93	*2,70*	0,64	*-1,480*	2.12	5	6,0204	0,0003
RYANAIR	2,86	*3,51*	-0,396	*0,204*	0.6	15	3,5507	0,0014
DISNEY	4,07	*2,80*	-0,111	*-0,878*	0.767	25	4,4904	0,00004
VERBANIA	2,43	*3,49*	-0,57	*0,514*	1.084	15	3,6988	0,0009

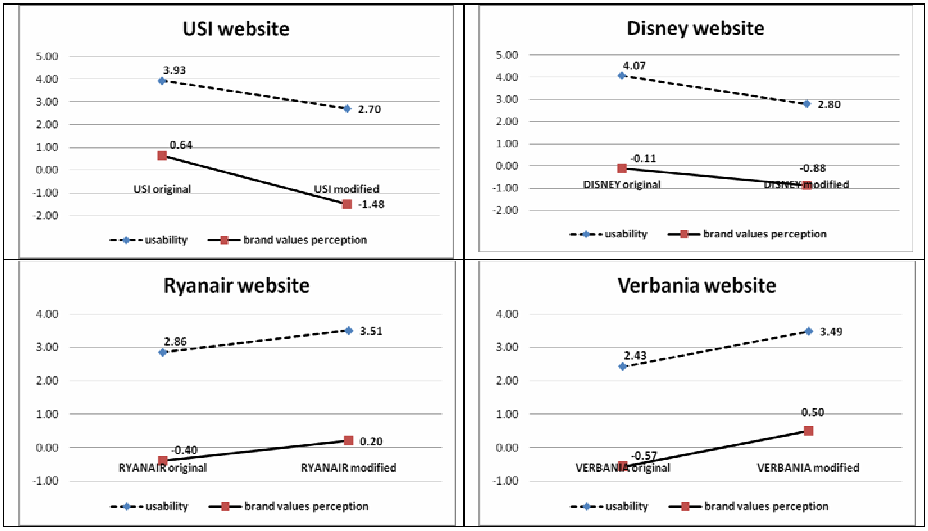


Fig. 3. Results synopsis for the four website experiments. Data points represent the 2 versions of the websites: *original* and *modified* with respect to the usability level.

version of the same website featuring fewer (resp. more) usability breakdowns. In other words, the results from all four controlled experiments seem to *confirm our research hypothesis*: as the usability of a website improves, the level of brand value perception raises, while when usability is worsened, the brand values perception decreases significantly.

5 Discussion and Conclusions

Even though the results of this research were partially predictable, to our knowledge this is the first public study that provides some empirical evidence on the correlation between brand perception and web usability. Our research indicates that, in the four experiments, web usability significantly affects brand image, and this influence tends to suggest a positive correlation: the better usability, the better would be the perception of the brand image; the worse is the usability, the more negative would be the users' perception of the brand. Our work considered a specific category of applications, i.e., content intensive branded web sites. It is an open issue for research whether our methodological hypothesis also holds in other categories of branded web applications, which have limited amount of information or services and are almost exclusively designed to convey brand messages by engaging the user emotionally, e.g., thorough impressive visual effects, animations, or videos. We may expect that in web sites of this kind the degree of usability can have a less significant influence on brand perception than other factors related to aesthetics, attractiveness, and emotional design. [10][1][29][32].

Our research has its shortcomings, most notably the relatively limited sample size in each experimental condition. Still, the fact that we have tested our hypothesis in four controlled experiments designed according to the same methodological framework, partially compensates this drawback and strengthens the validity of our results. As also pinpointed by Greenberg & Buxton in their recent paper [20], "... rigorous science also demands replication, since the scientific method advocates for repeated attempts of refuting the hypothesis. …The more the test tries to refute the hypothesis, the more powerful it is. ...If the hypothesis holds in spite of attempts to refute it, there is more validity in its claims".

This study has also contributed methodologically, in several aspects. While using controlled experiments, scenario-based evaluation, and questionnaire-based data collection are obviously not new, the idea of focusing the overall process on the concept of brand value is original. Existing value-driven approaches normally address two main issues: how to design technology that accounts for human values [14] or creates value for the user "in the world" [8], and how to evaluate how a system meets these requirements [7]. In contrast, we shift the value-oriented perspective to the consideration of the users' perception of the *qualities of the entity behind a system*, i.e., the *brand* values. Users may or may not share these values, nor find them worthwhile; still, in a branding strategy it is important to evaluate whether or not (and at which degree) these brand values are perceived. In other words, differently from other value-oriented approaches, our focus is on how a system communicates these values, how the users become aware of and associate them to the image of that entity in their mind, and how *usability* plays a role in this articulate communication process.

In addition, the general evaluation method that has been adopted in all our four controlled experiments is founded on some novel aspects. Firstly, we operazionalize more rigorously the concept of brand, as usually intended by people in marketing or design, in order to make it measurable. Secondly, we systematically combine the evaluation of brand values perception and usability with a method for systematic elicitation of brand values [3][4][43]. The integration of these methodological components provides a comprehensive conceptual framework that can be reused in other experimental settings, either for replicating our study, or for designing similar studies focusing on the correlation of web branding vs. design factors other than traditional usability. Future researchers can employ or build upon it, saving time and effort.

We have showed that in content intensive web sites, usability *as a whole* has an impact on the positive or negative perception of brand values. Still, the individual contribution of each different attributes of usability (e.g., usability of information architecture, content, or layout), and of other aspects of interface design (e.g., attractiveness or aesthetics) [9][10][21] to the effective communication or miscommunication of the brand values is to be further explored. What is the role that each of these elements plays to influence the effectiveness of brand values communication? Our study provides initial results and methodological guidelines that future research can refine, either to analytically distill the usability dimensions that bear most of the responsibility in this process, or to prove the actual impact of other design factors. As a final note, it is important to acknowledge that we have considered the influence of usability on the *short-term* "perception" of brand values, as induced by the web exposure. Studying *sustained*, long-term brand values perception would be a promising follow-up for this line of work.

References

1. Alben, L.: Quality of Experience. Interactions 13(5), 12–15 (1996)
2. Baty, S.: Brand Experience in User Experience Design, http://www.uxmatters.com
3. Bolchini, D., Garzotto, F., Paolini, P.: Branding and Communication Goals for Content-Intensive Interactive Applications. In: Proceedings of 15th IEEE International Conference on Requirements Engineering, New Dehli (India), pp. 173–182. IEEE Press, Los Alamitos (2007)
4. Bolchini, D., Garzotto, F., Paolini, P.: Value-Driven Design for "Infosuasive" Web Applications. In: Proceedings of 17th International World Wide Web Conference – WWW 2008, Beijing, China (2008)
5. Bolchini, D., Garzotto, F.: Quality and Potential for Adoption of Web Usability Evaluation Methods: An Empirical Study on MILE+. Journal of Web Engineering 7(4), 299–317 (2008)
6. Costabile, F., Garzotto, F., Matera, M., Paolini, P.: The SUE Inspection: A Systematic and Effective Method for Usability Evaluation of Hypermedia. IEEE Trans. on Systems, Man, and Cybernetics, Part A 32(1), 93–103 (2002)
7. Cockton, G.: A Development Framework for Value-Centered Design. In: Proc. CHI 2003, pp. 1292–1295. ACM Press, New York (2003)
8. Cockton, G.: From quality in use to value in the world. In: CHI 2004 Extended Abstracts. ACM Press, New York (2004)
9. De Angeli, A., Sutcliffe, A., Hartmann, J.: Interaction, usability and aesthetics: what influences users' preferences? In: Proc. 6th Conf. on Designing Interactive Systems, pp. 271–280. ACM Press, New York (2006)
10. De Angeli, A., Lynch, P., Johnson, G.I.: Pleasure versus efficiency in user interfaces: Towards an involvement framework. In: Green, W.S., Jordan, P.W. (eds.) Pleasure with products: Beyond usability, pp. 97–111. Taylor & Francis, Abington (2002)
11. Fogg, J.B., Soohoo, C., Danielson, D.R., Marable, L., Stanford, J., Tauber, E.R.: How do users evaluate the credibility of Web sites? A study with over 2,500 participants. In: Proc. Designing For User Experiences 2003, pp. 1–15. ACM Press, New York (2003)
12. Fogg, B.J., et al.: What Makes Web Sites Credible? A Report on a Large Quantitative Study. In: Proc. CHI 2001, pp. 61–68. ACM Press, New York (2001)
13. Fogg, B.J.: Persuasive Technology. Morgan Kaufmann, San Francisco (2003)
14. Friedman, B., Kahn, P.H.: Human values, ethics, and design. In: Jacko, J.A., Sears, A. (eds.) The human computer interaction handbook. Lawrence Erlbaum Associates, Mahwah (2003)
15. Friedman, B., Kahn, P.H., Howe Jr., D.C.: Trust online. Comm. of the ACM 43(12), 34–40 (2000)
16. Gerstman, R. (ed.): Branding@thedigitalage. Palgrave (2001)
17. Gilbert, A.C., Iacobucci, D.: Marketing Research: Methodological Foundations, 9th edn. Thomson South-Western Publishers (2005)
18. Gobé, M.: Emotional Branding. Allworth Press (2001)
19. Gordijn, J., Yu, E., Raadt, B.: E-Service Design Using i* and e3value Modeling. IEEE Software 23(5), 26–33 (2006)
20. Greenberg, S., Buxton, B.: Usability Evaluation Considered Harmful (Some of the Time). In: Proc. CHI 2008, pp. 111–120. ACM Press, New York (2008)
21. Hartmann, J., Sutcliffe, A., De Angeli, A.: Investigating attractiveness in web user interfaces. In: Proc. CHI 2007, pp. 387–396. ACM Press, New York (2007)

22. Hassenzahl, M.: The interplay of beauty, goodness and usability in interactive products. Human-Computer Interaction 19(4), 319–349 (2004)
23. Jokela, T.: When good things happen to bad products: where are the benefits of usability in the consumer appliance market? Interactions 11(6), 28–35 (2004)
24. Keinonen, T.: One-dimensional usability - Influence of usability on consumers' product preference. University of Art and Design Helsinki, - doctoral dissertation (1998)
25. Kim, J., Lee, J., Choi, D.: Designing Emotionally Evocative Homepages: An Empirical Study of the Quantitative Relations Between Design Factors and Emotional Dimensions. Int. J. of Human-Computer Studies 59(6), 899–940 (2003)
26. Knemeyer, D.: Brand Experience and the Web. Digital Web Magazine, http://www.digital-web.com/articles/brand_experience_and_the_web/
27. Lavie, T., Tractinsky, N.: Assessing dimensions of perceived visual aesthetics of websites. International Journal of Human-Computer Studies 60(3), 269–298 (2004)
28. Marcus, A.: Branding 101. Interactions 11(5), 14–21 (2004)
29. McNamara, N., Kirakowski, J.: Defining Usability: Quality of Use or Quality of Experience? In: Proc. 2005 IEEE Int. Professional Communication Conf., pp. 200–2004. IEEE Press, Los Alamitos (2005)
30. Muller, B., Chandon, J.L.: The Impact of Visiting a Brand Website on Brand Personality. Electronic Markets 13(3), 210–221 (2003)
31. Nolan, K.: Europe Brand Perception: Nokia Dominates for Loyalty and Aspiration. Strategic Analytics Report Overview, http://www.strategyanalytics.com/
32. Norman, D.A.: Emotional Design. Basic Books (2004)
33. Park, S., Harada, A., Igarashi, H.: Influences of Product Preferences on Product Usability. In: Proc. CHI 2006, pp. 87–92. ACM Press, New York (2006)
34. Rondeau, D.: For Mobile Applications, Branding is Experience. Comm. of the ACM 48(7), 61–66 (2005)
35. Sealey, P.: How E-commerce Will Trump Brand Management. Harvard Business Review 68(4) (1999)
36. Sinha, I.: Cost Transparency: The Net's Real Threat to Prices and Brands. Harvard Business Review 78(2), 43–50 (2000)
37. Spool, J.: Determining How Design Affects Branding, http://www.uie.com/articles/design_and_branding/
38. Spool, J.M.: Branding and usability, http://www.uie.com/articles/branding_usability/
39. Tractinsky, N.: Aesthetics and apparent usability: Empirically assessing cultural and methodological issues. In: CHI 1997 Electronic Publications. ACM, New York (1997), http://sigchi.org/chi97/proceedings/paper/nt.htm
40. Tractinsky, N., Shoval-Katz, A., Ikar, D.: What is beautiful is usable. Interacting with Computers 13(2), 127–145 (2000)
41. Wheeler, A.: Brand Identity: A Complete Guide to Creating, Building, and Maintaining Strong Brands. Wiley, Chichester (2006)
42. Young & Rubricam Consulting. Brand Asset Evaluator Survey, http://www.12manage.com/methods_brand_asset_valuator.html
43. Bolchini, D., Garzotto, F.: Design Requirements for Communication-Intensive Interactive Applications. In: Lyttinen, C., Loucopoulos, P., Mylopoulos, J., Robinson, W. (eds.) Design Requirements Engineering: A Ten-Year Perspective, February 2009. Lecture Notes Series in Business Information Processing, pp. 408–431. Springer, Heidelberg (2009)

What Needs Tell Us about User Experience

Annika Wiklund-Engblom[1], Marc Hassenzahl[2,1], Anette Bengs[1], and Susanne Sperring[1]

[1] MediaCity, Åbo Akademi University, Box 311, 65101 Vaasa, Finland
[2] Folkwang University, Campus Universitat Duisburg-Essen, Universitatstrasse 12, 45141 Essen, Germany
{Annika.Wiklund-Engblom,Anette.Bengs,Susanne.Sperring}@abo.fi, Marc.Hassenzahl@Folkwang-Hochschule.de

Abstract. The present study explores the sources and consequences of fulfilling six fundamental human needs, namely *Autonomy*, *Relatedness*, *Competence, Stimulation, Influence*, and *Security,* through using interactive products and media. Each need refers to a distinct set of issues, such as according product attributes (e.g., "flexibility") and experiential consequences (e.g., "freedom of choice"). Besides the need-specific content, which helps to characterize and differentiate user experiences, the study reveals the close relation between needs and according product attributes as their mirror images.

Keywords: User Experience, Fundamental Needs, Technology and Media Interaction.

1 Introduction

User experience (UX) as a concept recently gained momentum in the field of Human-Computer Interaction (HCI) [e.g., 5]. Despite the lack of an official definition and some debates about the actual scope of UX [e.g., 7], there is a common understanding that UX is *holistic* – it emphasizes the totality of emotion, motivation, and action in a given physical and social context – and that it is *subjective* – focusing on the "felt experiences" [8] rather than product attributes.

Battarbee [1] distinguishes *person-centred* frameworks, focusing on human needs and/or the relationship people have with products from more *product-centred,* focusing on qualities of the design, i.e. product attributes. Although, many different definitions and models of UX exist, there seems to be some agreement that UX complements a traditional, usability-oriented approach by going beyond the technical or task-related into human emotions and needs, that is, from the product-centred to the person-centred. However, this does not imply that usability is not important anymore. It rather highlights that UX as a consequence of both, product-centred aspects, such as functionality and aesthetics, as well as person-centred aspects, such as personal motivation and needs [10].

Personal motivations and needs received relatively little attention in HCI, but current research on UX seeks to close that gap. A recent definition, for example, describes UX as "a momentary, primarily evaluative feeling (good-bad) while

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 666–669, 2009.

interacting with a product or service. […] Good UX is the consequence of fulfilling the human needs for autonomy, competency, stimulation (self-oriented), relatedness, and popularity (others-oriented) through interacting with the product or service. To do so, products and services need a balance between pragmatic (e.g., usability) and hedonic quality (e.g., novelty)" [3]. Here, underlying, general psychological needs "beyond the instrumental" [5] play a major role as a source of positive emotions (e.g., pleasure, enjoyment) and well-being and as motivators of action [see also 4 and 2].

One research instrument from psychology that specifically targets fundamental needs is a Needs Schedule developed by Sheldon, Elliot, Kim, and Kasser [9]. This questionnaire "measures" ten general fundamental needs. These were used on several occasions to let people describe their experiences with technology and media [3; unpublished studies]. It summarizes and builds upon a variety of prominent theories, such as Ryan and Deci's Self-Determination Theory, Maslow's theory of personality, Epstein's cognitive-experiential self-theory, and Derber's lay theory of human needs.

The present study explores the sources and consequences of the fulfilment of six out of the ten Sheldon's *et al.* fundamental needs, namely *Autonomy*, *Relatedness*, *Competence, Stimulation, Influence*, and *Security*. Instead of using a questionnaire, however, participants answered open-ended questions (in writing) on the experience of the fulfilment of each need through using an interactive product and media. Ten people participated in the study. Each wrote about their needs fulfilment through a freely chosen product and media (i.e., 20 descriptions, two from each participant).

2 Results

We analyzed the data by the method of meaning condensation [6] of the participant's statements to identify sources of needs fulfilment and its experiential consequences. Table 1 shows an overview of the results organized by respective need.

The experience of *autonomy*, for example, was related to the opportunity to "be real", i.e. being enabled by a product to be oneself, without caring much for the demands of other social roles, to "express own opinions" and to define one's "integrity". Both are the consequence of the specific functionality of the product or media mentioned by the participant and revolve around the theme of self-identity. More generally, *autonomy* was connected to the "flexibility" of the product, i.e., the freedom of using or not using it anywhere and anytime. Flexibility can be thought of as more general product attribute, summarizing functionality and other aspects of the products (e.g., mobility, usability). A last source for *autonomy* was the freedom using whatever product one prefers. The value lies in the free choice per se, which cannot be linked directly to specific functionality or other product attributes. *Security* – to give a further example – was implied by "reliability" and "familiarity" of the product (and situations) and "social support" enabled through the functionality of the product (see Table 1 for an overview of the needs related experiences and interpretations).

Table 1. Needs Interpretations Based on Participants' Technology and Media Experiences

Needs	Interpretation	Participants' Experience Descriptions
Autonomy	Flexibility	*"I can listen to it anywhere." [mp3 player]*
	Integrity	*"My phone – integrity" [mobile phone]*
	Freedom of Choice	*"It was a big step to switch from a PC to a Mac." [MacBook]*
	Express Own Opinions	*"I, of course, have my own opinion." [blog]*
	Be Real	*"Being able to be myself without having to care about my other social roles." [Facebook]*
Related-ness	Communication	*"Definitely, by keeping in touch with a large group of friends." [Facebook]*
	Information	*"Yes, when there are local news in the paper." [newspaper]*
	Sharing	*"Not the music, but sharing it with others." [music]*
	Brand/Identity	*"I love my Mac = We the Mac-users against PC." [MacBook]*
Competence	Usability	*"Everything works naturally and well ..." [MacBook]*
	Novelty	*"... at first when I found new functions, but not anymore." [mp3 player]*
	Mastery	*"It gives me a feeling of competence when I find a new radio channel." [iTunes radio]*
	Knowledge	*"I feel up-to-date about what is happening locally and in the world." [newspaper]*
Stimulation	Social	*"... to have access to various chats simultaneously." [Fring]*
	Cognitive	*"Yes, to express oneself in cleaver ways and communicate intelligently with friends provides stimulation." [Facebook]*
	Emotional	*"It stimulates on an emotional level." [music]*
	Existential	*"It confirms my existence." [newspaper]*
	Creativity	*"It gives me new ideas." [newspaper]*
Influence	Connectedness	*"To be able to reach others and be reached increases the feeling of influence." [laptop]*
	Communication	*"By using it to communicate, influencing others." [mobile phone]*
	Awareness of Options	*"I can influence – if I have the energy to do so." [newspaper]*
Security	Reliability	*"I can trust that it always works." [coffee maker]*
	Satisfaction	*"It's a way of winding down at home in the sofa." [TV]*
	Social Support	*"I can call someone. Help is close by if it's needed." [mobile phone]*
	Familiarity	*"Ordinary things in an ordinary environment." [newspaper]*

3 Discussion

This study was a first step in illuminating the issues revolving around the fulfilment of a set of universal human needs through interactive products and media. Besides the insights into the specific topics related to specific needs, such as reliability to *security* or "being real" to *autonomy*, at least two notable, more general aspects became apparent. First, issues are quite different depending on the particular need. Thus, designing for a security experience may invoke a completely different set of requirements compared to a stimulation, competence or autonomy experience. In all cases, needs fulfilment is a source of general positive emotions but the presumably soothing, relaxing experience of security will be quite different from the strong, bold, and self-conscious experiences of autonomy – not only phenomenological, but also in the according functional requirements and the product attributes looked for.

Second, asked about the experience of needs fulfilment through products, participants spontaneously produce *person(self)-centred* aspects, such as emotional states, feelings, and values, as well as *product-centred* aspects, such as functionality or products attributes. On one hand, this highlights that people are aware and can articulate the way in which a product or media fulfils their needs. On the other hand, it clarifies the actual nature of product attributes, such as "usable," "reliable," or "flexible." As the name implies, a product attribute is the consequence of an attribution process, which relates an experience to the product and establishes it as the cause of this experience. In other words, reliability is one consequence of security experiences perceived as being caused by the respective product. This implies fewer differences between a product-centred and a person-centred approach as expected [e.g., 1], because attributes are the consequence of needs fulfilment-related experiences – their mirror images.

Universal needs are limited in number, but may still be a promising way to characterize and distinguish different types of experiences. Other than experience-frameworks such as McCarthy and Wright's [8], needs address the *content* of experiences not only their structure. Where McCarthy and Wright, for example, stop with the observation that experiences have an emotional-motivational and a meaning-making thread, the needs perspective clarifies where the emotion, motivation, and meaning comes from. By that, it becomes much easier to address experiences in the context of product and media design.

References

1. Battarbee, K.: Co-Experience: Understanding User Experiences in Social Interaction, Helsinki, Finland. Publication series of the University of Art and Design (2004)
2. Cockton, G.: From Quality in Use to Value In The World. In: CHI 2004 Conf. on Human Factors in Computing Systems. Extended abstracts, pp. 1287–1290. ACM, New York (2004)
3. Hassenzahl, M.: User Experience (UX): Towards an Experiential Perspective on Product Quality [Keynote]. In: The 20th IHM 2008 Conf. (2008)
4. Gaver, W.W., Martin, H.: Alternatives. Exploring Information Appliances through Conceptual Design Proposals. In: CHI 2000 Conf. on Human Factors in Computing, pp. 209–216. ACM, Addison-Wesley, NY (2000)
5. Hassenzahl, M., Tractinsky, N.: User Experience - A Research Agenda [Editorial]. Behavior and Information Technology 25, 91–97 (2006)
6. Kvale, S.: Interviews: An Introduction to Qualitative Research Interviewing. SAGE Publications, Thousand Oaks (1996)
7. Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., Korte, J.: Understanding, Scoping And Defining User eXperience: A survey approach. In: CHI 2009 Conf. on Human Factors in Computing Systems, pp. 719–728. ACM, New York (2009)
8. McCarthy, J., Wright, P.C.: Technology as Experience. MIT Press, Cambridge (2004)
9. Sheldon, K.M., Elliot, A.J., Kim, Y., Kasser, T.: What Is Satisfying About Satisfying Needs? Testing 10 Candidate Psychological Needs. J. of Personality 80, 325–339 (2001)
10. Vyas, D., van der Veer, G.C.: Rich Evaluations of Entertainment Experience: Bridging the Interpretational Gap. In: The 13th ECCE 2006, vol. 250, pp. 137–144 (2006)

From Paper to PDA: Design and Evaluation of a Clinical Ward Instruction on a Mobile Device

Anne Marie Kanstrup[1] and Jan Stage[2]

[1] Aalborg University, Department of Communication, Kroghstræde 1
[2] Aalborg University, Department of Computer Science, Selma Lagerlöfs Vej 300
DK-9220 Aalborg East, Denmark
kanstrup@hum.aau.dk, jans@cs.aau.dk

Abstract. Mobile devices with small screens and minimal facilities for interaction are increasingly being used in complex human activities for accessing and processing information, while the user is moving. This paper presents a case study of the design and evaluation of a mobile system, which involved transformation of complex text and tables to digital format on a PDA. The application domain was an emergency medical ward, and the user group was junior registrars. We designed a PDA-based system for accessing information, focusing on the ward instruction, implemented a prototype and evaluated it for usability and utility. The evaluation results indicate significant problems in the interaction with the system as well as the extent to which the system is useful for junior registrars in their daily work.

Keywords: Interaction design, mobile system, usability evaluation, interpretive evaluation, case study.

1 Introduction

Mobile devices with a range of purposeful applications have changed many jobs that involve working in new settings and on the move. The challenge for user interaction design is to create applications that are useful in the mobile situation despite severely reduced interaction facilities.

Many early applications on mobile devices involved user interaction with only simple data, e.g. communication facilities (mobile telephone, SMS), and games. This applies also to recent applications like music and video players and cameras. Yet other applications for mobile devices involve more complex data in long lists or large tables. Early examples include calendars, address books and web browsers. Today, applications like email, calendars, contact lists and finance systems as well as systems for sales, business, medicine, travel and education involve complex data in list or table structures. Recent systems for instruction, direction, training, visualization, localization, social networking, remote monitoring, and accessing and browsing information on the web also involve user interaction with very complex data.

Complexity of information is a key difficulty for such applications. This complexity originates either from the sheer amount of information or from the intricate and dense nature of the information [24]. Complexity originating from the amount of

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 670–683, 2009.

information is particularly challenging for the design of systems for mobile devices. Handheld mobile devices have very limited facilities for user interaction; most notably the small size of the screen, but also the replacement of standard input devices like keyboard and mouse with compressed and minimized keyboards, a few control buttons and a stylus. A number of studies have examined the consequences of these differences [1, 8, 10, 29].

This paper contributes to research on interaction design and evaluation of complex information systems for mobile devices. We report from a case study of the design and evaluation of a clinical ward instruction on a mobile device for use in an emergency medical ward. The health care domain is interesting, both because there is increasing focus on the possibilities of deploying mobile systems for use by physicians and nurses [11, 28], and because it is so difficult to develop successful systems.

The paper is organized as follows: In the following section, we present existing research on complex information processing on mobile devices, including experiences from the healthcare domain. Next, our case is described with emphasis on the design challenge. Then we present the transformation of paper instructions into digital form for a PDA. This is followed by results from a conventional usability evaluation and a qualitative interview-based evaluation of utility. Experiences acquired through the study are discussed, and the conclusion sums up and suggests areas for future work.

2 Related Work

There are several studies of interaction with text and tables on mobile devices and the effect of physical movement on text reading and comprehension. Some of these have shown that in various use situations physical movement may reduce performance in the interaction with a mobile handheld device [4, 21, 23, 25, 27, 33]. This occurs for two reasons: (1) the user moves between different locations and thereby changes context of use [16,20]; and (2) the user moves physically which reduces the possibility for precise interaction [27]; this second reason has been examined in a number of studies, documenting that performance is reduced and task load increased when the user is mobile, e.g. [21, 23, 25, 26, 33].

The difficulties of interacting with complex information in tables on mobile devices have been examined in a number of studies, e.g. on table lookup on a small screen [34] and visualization of lists with patient data on a small screen [21]. However, the majority of these studies have been more design-oriented, e.g. [29, 35]. Surprisingly, none of these studies have inquired into the effects of physical movement on user performance.

A key challenge when complex information has to be presented on small screens is to find optimal levels of information hierarchies. Results from a comparison of mobile web browsing with desktop browsing indicate poor performance on a mobile browser, especially when trying to locate content in long narrow pages, because it requires extensive scrolling [29]. A solution to mobile browsing fatigue has been suggested [1]. Another study compared interfaces using zoom and overview with classic zoom interfaces in the context of user navigation of large information spaces on mobile devices [5]. Similar results about overview were found in a study of the usability of four different information hierarchies on mobile devices [12]. These studies improve our understanding of the challenge faced when working with complex information for mobile devices, and they demonstrate that performance is reduced when using desktop designs and scroll interaction for the mobile device, whereas performance is increased when the design provides overview of the complex information.

Some studies have dealt specifically with the use of mobile devices in healthcare. A survey of 3600 users of a handheld clinical reference application reported that current adopters used these tools frequently and found them useful for improving patient care and learning [28]. Similarly, the affordance of handheld computers to clinical practice has been pointed out [1, 2]. However, a literature review on the use of handheld devices in medicine, emphasizes a lack of evidence-based information about the use of PDAs in medicine [11]. A review on adoption of PDAs in healthcare identifies major barriers to adoption, including usability problems, security concerns, and lack of technical and organizational support. Better designed PDA hardware and software applications, more institutional support, seamless integration of PDA technology with hospital information systems, and satisfactory security measures are necessary to increase acceptance and use [22]. Finally, a study identified ward instructions as a central and complex information need [8].

3 Case: A Ward Instruction in an Emergency Medical Ward

The case study reported in this paper was carried out at an emergency medical ward at a large regional hospital in Denmark. The ward that participated in the project is staffed with 16 junior registrars. Due to rotation, 20 registrars in total participated in the different activities of the project.

The target group of users was junior registrars, i.e. physicians who have just ended medical studies in university and face one year of clinical training including six month at a medical ward. Junior registrars need to operationalize their theoretical knowledge into clinical decision making in concrete situations [17, 18]. To support this acquisition, information, typically in paper format, is carried in the pockets of their white coats. Figure 1(a and b) shows an average white coat of a junior registrar at the ward with the average weight of 2,6 kilos. The pocket content is displayed in (b) with reference tables on medicine, phone numbers, diagnosis codes etc., ward instruction, personal notebooks, text books, keys, pens and medical equipment.

Inspired by the junior registrars' request for and recommendation of mobile technology to support their clinical training [14], a research project was carried out with the aim of identifying relevant mobile applications [18]. The project was organized as a participatory design process with initial observations, interviews and a workshop with physicians at the ward. These activities identified the clinical ward instruction as the most central, obligatory and used information material for junior registrars and clinical staff at the ward in general [17]. This focus was supported by related research on handheld devices to support junior registrars in Danish hospital training [8].

The clinical ward instruction is written by the chief physicians and includes descriptions of archetypes of diseases at the ward, list of symptoms, checklists, and instructions for monitoring and treatment. It represents high level expert knowledge, and the instructions are procedures that have been decided upon, and they must be followed by staff.

(a) (b) (c)

Fig. 1. (a) The white coat of a junior registrar. (b) The content of the white coat pockets (in total2.6 kg). (c) The clinical ward instruction made by chief physicians at the ward.

Figure 3(c) shows a ward instruction that has been used for four months by a junior registrar. It is size B5, and the text includes coherent text chunks up to 25 normal pages, in addition to a set of tables, each being typically the size of a full page. Personal notes are added in the margin with the purpose of keeping track, highlighting, and saving advice and experience.

The aim of digitizing the complex information of the ward instruction in a mobile device was complemented with observations of and interviews with junior registrars emphasizing the following meaning of mobility at the medical ward:

- Mobile means walking – the physicians walk while looking up information and interaction techniques for on-the-move-interaction like e.g. one hand interaction, were emphasized.
- Mobile means in-between, on the way from one task to the next. Maximum time for looking up information or making notes was defined to app. 5-10 minutes. As explained by one of the physicians: “if you have more than 10 minutes, you find a PC”. Consequently, navigation supporting easy to find was emphasized.
- Work is filled with interruptions. The physicians are constantly interrupted and as a consequence working hard to keep track via personal notebooks on their patient. Emphasis was on not designing yet another interrupting artifact but rather a personal device supporting tracking of work activities.

4 Interaction Design

The key decisions in the development of the prototype were the design of the basic form of interaction with the digital ward instruction and the transformation of the paper ward instructions into the digital version. These decisions were made in a participatory process that involved the physicians at the ward.

The transformation of the paper version to the digital version on the PDA faced the following challenges: create a useful structure for a database, develop a software editor for producing elements within this structure and design the detailed interaction with the ward instruction on the mobile device.

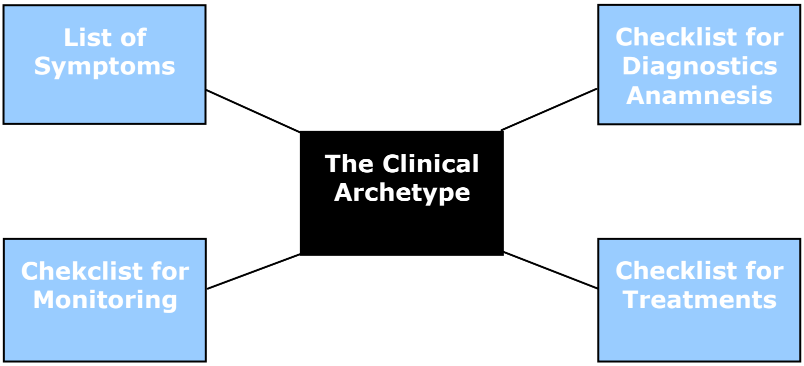


Fig. 2. The basic structure of the digital ward instruction

It was decided to structure the digital ward instruction around archetypes. This decision was based on empirical analysis and a theoretical perspective on medical knowledge [686]. An archetype is a patient with a well-known disease, e.g. the febrile patient. The reasons for this fundamental decision were that archetypes are related to the use situation of the application, and archetypes support the transformation of theoretical knowledge to concrete decision making in clinic faced by the junior registrars: to fit a patient into a correct archetype and perform the necessary individual patient-specific adjustments compared to an abstract, average patient described in the archetype.

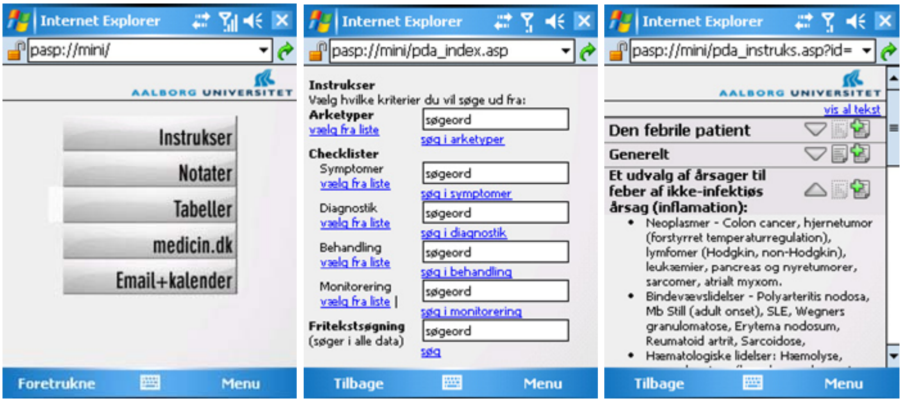
The transformation of the existing paper archetypes into digital archetypes was carried out on the basis of a modified general work-flow model with the following four "milestones", see Fig. 2:

- List of symptoms and signs that contains specific ideas for further and alternative information relevant for a more precise positioning of the patient within the range of the archetype diagnoses.
- Checklist for diagnostics, objective findings and test that contains relevant ideas for specific hospital procedures.
- Checklist for treatment that contains plans for treating the patient.
- Checklist for monitoring that also contains alternative actions.

The software used as editor for producing text on the digital archetypes was designed in Microsoft Word. This was reasonable since the text for the existing ward instructions on paper, see Fig. 1(c) was in Word, and because this software was known to the archetype-producers (the chief physicians). They were allowed to mark and prioritize the order in which they entered the text for later hyperlinking by means of a parser constructed in the project.

The transformation of the long texts for each archetype (some were up to 25 normal pages) has been done by designing the navigation for the text according to general design guidelines and empirical knowledge from the case. Focus was on:

- Interaction using both screen and buttons
- A frontpage with shortcuts to key functions



(a) (b) (c)

Fig. 3. Screen-dumps from the digital ward instruction. (a) The frontpage with key functions. (b) Search functions. (c) An example of an archetype organized in headlines.

- Direct access through search functions
- Providing overview of the main information

Interaction design was not only based on the small screen of the PDA but also on the hardware as interface. A shortcut to the ward instruction was programmed and works with activation of a button at the front of the PDA.

There were shortcuts to key functions in the digital ward instruction via a frontpage that gives access to (1) instructions organized in archetypes, (2) personal notes, (3) tables with specific information and (4) a link to a national website with medical handbooks, drug catalogues etc. Usually, the junior registrars are carrying heavy books with the information mentioned in (4). This frontpage is illustrated in Fig. 3(a). A click on the logo, present at the top of the screen on most pages, takes the user back to the frontpage.

The interaction design also includes a page with search functions that provide direct access to specific information. It is possible to make direct free text search or to search in lists of symptoms, diagnostics, treatments, and monitoring, see Fig. 3(b).

The main information in the system is structured after the archetypes. For each archetype, a number of headlines was defined. They gave the overall structure on the information about the archetype. This rather long text (up to 25 standard Word pages) would be extremely long when transformed to the small screen. The structure and the headlines were defined in co-operation with the archetype-writers/chief physicians. The text under each headline can then be unfolded by clicking a drop-down icon, see Fig. 3(c). The headlines also give 'at a glance' information on whether there are any notes to the text by using a transparent icon of a document when no notes are available and a clear icon when notes has been inserted. The same icon with the '+' is the icon used when inserting notes to the text, see Fig. 3(c).

The paper instructions included a set of tables with structured information. The transformation of these tables to the PDA was a particular challenge, because they

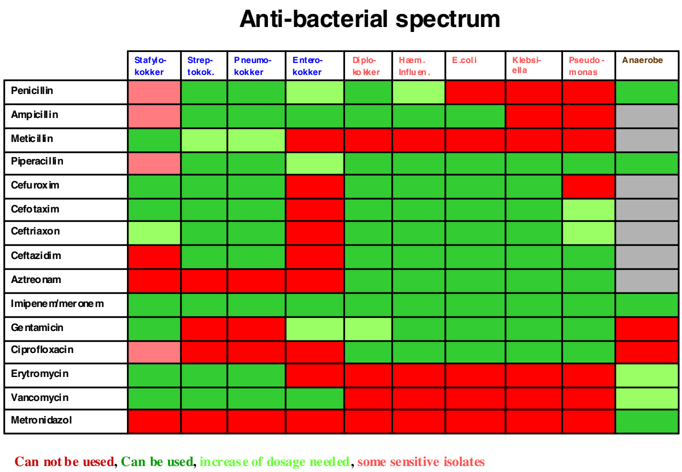


Fig. 4. A table from the ward instruction in full size

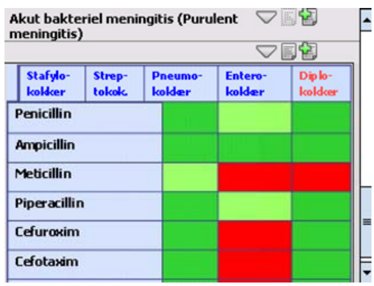


Fig. 5. Screen-dump of the table as shown on the small screen of the PDA

were quite large. There is an example of one of the tables in Fig. 4. If this table was transformed to be fully visible on the PDA screen, it would be impossible to read. Thus only part of it could be visible on the screen. Yet the problem with that is that the contents of a cell in the table only make sense if it is related to other cells or at least to the column and row headers.

The solution we implemented was a table with fixed column and row headers, where the user could scroll in both directions to get to a certain cell, see Fig. 5. The design of the digitized clinical ward instruction was implemented on a PDA. The implementation did not require changes to the design.

5 Evaluation

The prototype was distributed to the junior registrars at the ward and to the chief physician who had the educational responsibility for the registrars. They were not trained in the system, but they received a two-page laminated sheet that provided some information about getting started with the system and fitted in their pocket. They were instructed to work with the PDAs and explore their potentials. This exploration lasted for over one month and was followed by two evaluations.

17 physicians (16 registrars and one chief physician) participated in the evaluation of the digital ward instruction on the PDA. Eight registrars participated in the usability evaluation of the prototype. Seven registrars participated in the evaluation of the utility of the prototype. Both evaluations are described below.

5.1 Usability Evaluation

The first evaluation focused on usability. The digital ward instruction was evaluated in a classical usability evaluation [687]. This was conducted when the system had been used for 6 weeks.

The evaluation was conducted on site at the hospital, in the administrative section of the emergency ward in order to avoid disturbances. We were required to conduct the evaluation in this setting because we did not have the resources for taking the users to our usability lab. Given the exploratory nature of the study, we did not consider this as a major problem.

The users were seven junior registrars all on duty on the day of the evaluation. The evaluators were two usability experts who had a master thesis in HCI and significant experience with the classical approach. The two evaluators conducted a series of usability tests where each physician was sitting in a locker-room, solving a series of predefined tasks while thinking aloud. All tests were video recorded. The evaluators analyzed the video recordings separately. They also rated the severity on the scale: critical-serious-cosmetic [30]. Finally, they negotiated a joint list of usability problems with agreed severity ratings, see Table 1.

The evaluators identified 20 usability problems. This included 9 critical, 4 serious and 7 cosmetic problems, all related to the operability of the PDA interface. The list of usability problems emphasizes three main problems in the system:

- Navigation is not intuitive to the users
- The note function is poorly understood
- The table does not work satisfactorily, neither in terms of navigation nor in terms of content.

In addition, post-test interviews uncovered a fundamental problem. The idea of the system is to “transfer” the paper instructions onto the PDA. However, several users expressed that this did not work because they could not gain enough overview from the small screen and read the instructions from it. Therefore, they did not use that part of the system.

An important feature is that the system provides access to an on-line medicine catalogue. This catalogue provides detailed information about each medicament that is available. All users emphasized the importance of this feature and the high quality of the information provided.

Table 1. The problem list from the usability evaluation

#	Description	Severity
1	Navigation: does not know that "medicin.dk" is an item in the main menu	Cosmetic
2	Instructions: does not know the function of "quick checklist" in the instruction item on "acute serious asthma"	Critical
3	Navigation: is missing a "forward button" like the "back button"	Critical
4	Instructions: believes erroneously that a click on the name of an instruction will activate it	Cosmetic
5	Instructions: believes erroneously that the "+" sign by an instruction will activate it	Serious
6	Navigation: does not know that the logo activates the main menu	Cosmetic
7	Navigation: is unable to find the link to the table (bacteria and penicillin)	Critical
8	Physical: the physical start button does not activate the main menu	Serious
9	Notes: does not know that the PDA has Dictaphone, camera and video	Cosmetic
10	Notes: does not know that there is a note function	Cosmetic
11	Notes: does not know how a note is created	Cosmetic
12	Search: cannot get back from the search page as there is no logo	Critical
13	Search: cannot find an instruction as search only works with the correct combination and order of words	Critical
14	Table: cannot get to the relevant cell as there is no horizontal scroll-bar	Critical
15	Table: cannot control the pointer pen to get to the relevant cell	Critical
16	Table: cannot find the relevant cell as the leftmost column has partly disappeared	Critical
17	Table: cannot read the relevant cell as there is an automatic home function when the pen is lifted from the screen	Serious
18	Table: misses a legend for the colour codes in table cells	Critical
19	Table: misses information about dosage of medicaments	Serious
20	Table: misses information about package size for medicaments	Cosmetic

5.2 Utility Evaluation

The second evaluation focused on utility. The purpose was to identify how useful he digital ward instruction on the PDA was for the junior registrars in their work. It was conducted when the system had been used for six weeks.

This evaluation was based on a semi-structured qualitative interview. Two evaluators were present during an interview with one registrar at a time. It focused on the digital ward instruction, but there was a paper-made PDA artifact with a 'blank screen'/white paper for writing, and the registrars were encouraged to use it to illustrate their points and as a trigger for reflection and new ideas for alternative designs. Seven junior registrars and one chief physician who was responsible for the junior registrars' education participated. The two evaluators were HCI researchers with significant experience with the interpretive approach.

In the interview, each physician was asked to solve pre-defined tasks. The dialogue about user experiences during the 6 weeks of using the prototype was considered the

Fig. 6. A junior registrar writing and drawing design ideas on a paper-PDA

target information, capturing everyday obstacles and ideas. All interviews were video recorded. Based on notes and the video recordings, the evaluators produced a list of utility problems covering the following themes:

- Wireless network and IT-support
- Information access
- Mobile workers and technological islands
- High learning curve for beginners
- Conventional problems for mobile technologies

Wireless network and IT-support: Since one of the primary gains of the mobile technology was the access to on-line resources on the www, network problems were experienced as a serious obstacle. Only 3 of the junior registrars succeeded in getting online and only for a period, the longest being 3 weeks. The evaluation revealed a serious lack of infrastructure in the (rather old) hospital buildings. Only some parts of the buildings had wifi; the two medical wards had wifi while the casualty and the administrative building did not. Additionally, only 4 users were allowed on the network at a time, and 2-3 laptops on the wards were permanently used for accessing electronic health records, which left room for only 1-2 PDA users. The support of the PDA-users was also a serious problem. As explained by one of the junior registrars when asked why she did not ask the IT-department for support when she had network problems: "because then I was on night duty for a week and then I had 2 weeks on vacation and then we were so busy when I started up again so I just stopped using it".

Information access: A recurring reason for not using the PDA as the primary tool for information access was the lack of connection to on-line resources and the incomplete digitized ward instruction: only the 9 (out of 15) major and most used archetypes were transformed to the PDA due to time pressure on the side of both the interaction designers and the chief physicians at the ward. "I cannot use it when they are not all there" one of the junior registrars explained. "You like to think of a place where you have all your information" another registrar noticed.

Mobile workers and technological islands: The mobile doctors needed an application that was fully mobile in the sense of being used at whatever hospital or ward you are in contrast to the implemented ward specific 'island'. An IT strategy for introducing this type of technology must take the mobility of (junior) physicians into account. As explained by one of the junior registrars: "in two months I am leaving for general practice so I did not want to use energy on learning this technology and applications". During the first 18 month the registrars work at two different wards and in a general practice with different technologies, applications, policies etc. Thus supporting the mobility of junior registrars is not only about supporting the work within wards, but also about supporting mobility between wards.

High learning curve for beginners: "You really have to invest some time in it before it will pay off" one of the few experienced PDA-users explained. They did not experience a PDA as a 'walk-up-and-use' technology, but a demanding technology which users have to become familiar with, personalize or "tame" [6].

Conventional problems for mobile technologies: The evaluation also revealed a list of interaction problems: (a) too many clicks: "you have to start it up, open the program, log-on and make your search, while you would have been able to look it up in your book four times" as explained by one of the junior registrars; (b) too little screen space for reading long text; c) too difficult to find information with direct text search "we are used to looking at lists" one of the junior registrars explained referring to the contents in text books, reference books, indexes on the www, guidelines etc.; and (d) errors in two hyperlinks and one spelling mistake which reduced the value of the system.

6 Discussion

The results of the two evaluations came as a surprise. The usability problems uncovered a range of difficulties, although the transformation from paper to PDA and the design of the interaction was developed in close collaboration with the users.

The information implemented on the PDA is very complex. There is both a large amount of information, and the nature of the information very intricate and dense. This makes the transformation process interesting. The key decision was that the design was centered on archetypes. This provided a clear structuring mechanism when the paper instructions were cut into fragments that could be displayed on the mobile device. This worked well. The implementation of tables was less successful as the users experienced several usability problems when interacting with them.

The most striking result of the evaluation was the limited utility that the users had experienced. They had definitely not found the system particularly useful, and only a few of them had practiced use at a serious level. The study also confirmed the common experience that exploratory use of a prototype in the real context of use requires full access to information. In our case, this applies both to the information in the system and the access to the network.

In our prototype, neither the information in the system nor the network access were fully implemented. It has been emphasized that case studies are important for developing a better practice in usability evaluation [36], yet the argument also holds for design activities. In the case study, we have reported from, the most interesting aspect is the transformation of complex information on paper onto a PDA. It turned out to be virtually impossible to evaluate the usability and utility of the PDAs, because the users could not abstract from the limitation. This can be seen as a specific occurrence of the general design tension between attention to specific content versus attention to general content [32].

It could be discussed whether we should have conducted and discussed the utility evaluation before the usability evaluation. If the system does not offer the required functionality, the users might not want to use it. Yet if the system is not usable, the utility will be hard to discover. This reflects the more general challenge of founding design on organizational fervour for quality in use, fit to context and value in the world [713].

The main impact of this case study is a better understanding of interaction design for PDA-based system combined with experience about key obstacles for evaluating the utility and usability of such a system. Moreover, the case illustrates the difficulties faced in the actual evaluation, e.g. the range of problems with the hospital's infrastructure that lead to a high rate of abandonment.

7 Conclusion

We have presented experiences from a case study of the design and evaluation of a prototype on a PDA that implements a complex textual paper instruction. The instruction is a training document for junior registrars in an emergency medical ward. The interaction design has been described in detail.

The PDA was evaluated both in terms of usability and utility. The usability evaluation revealed a variety of basic problems with the transformation from paper to the PDA prototype, including problems with navigation, activating functions and interacting with information in tables. The utility evaluation uncovered fundamental problems in the application domain, including practical problems with the PDA technology as well as organizational problems that limited the usefulness of the prototype. The evaluations uncovered a more fundamental lesson full benefit from an evaluation in a context like ours, requires a complete computerized solution and not just a prototype that provides a good implementation of the novel aspects of the design.

The design and evaluation of the prototype has emphasized new areas for future research. There is clearly a need for more cases of designing systems for accessing complex information on a PDA in an application domain like the one we have worked with. It would also be interesting to develop more general and firm guidelines for transforming complex information into an interaction design for a PDA. This will facilitate development and deployment of mobile devices in new application domains. We have not dealt with the ethical aspects [13], but they need to be considered prior to full deployment of systems that support decision making and capturing of information about patients.

Acknowledgments. The research behind this paper was partly financed by the Danish Research Councils (grant number 2106-04-0022 and 2106-08-0011). The MINI-project was support by the Ministry of Science and Technology. We want to acknowledge the work of our colleagues Pernille Bertelsen, Ellen Christiansen, Thomas Christiansen, Lene Hofmeister, Christian Nøhr and Peter Risgaard who have contributed to ideas, design and evaluation. We are also grateful to Niels Boye and his colleagues at the medical ward for their participation in the project. Finally, we want to thank the anonymous reviewers.

References

1. Al-Ubaydli, M.: Handheld computers. In: BMJ 2004, vol. 328, pp. 1181–1184 (2004)
2. Al-Ubaydli, M.: Handheld Computers for Doctors. Wiley, Chichester (2003)
3. Borodin, Y., Mahmud, J., Ramakrishnan, I.V.: Context Browsing with Mobiles - When Less is More. In: MobiSys 2007, San Juan, Puerto Rico, USA, June 11-14, 2007, pp. 3–15 (2007)
4. Brewster, S.A., Lumsden, J., Bell, M., Hall, M., Tasker, S.: Multimodal 'Eyes-Free' Interaction Techniques for Mobile Devices. In: Proceedings of CHI 2003, pp. 473–480. ACM Press, New York (2003)
5. Burigat, S., Chittaro, L., Parlato, E.: Map, Diagram, and Web Page Navigation on Mobile Devices: the Effectiveness of Zoomable User Interfaces with Overviews. In: MobileHCI 2008, Amsterdam, The Netherlands, September 2-5, 2008, pp. 147–156 (2008)
6. Christiansen, E.: Tamed by a rose. In: Nardi, B. (ed.) Context and Concioussness. MIT Press, Cambridge (1996)
7. Cockton, G.: Focus, fit, and fervor: future factors beyond play with the interplay. International Journal of Human-Computer Interaction 21(2), 239–250 (2006)
8. Dahlerup, J.F., Roelsgaard, K.: Handheld computers in clinical training of doctors – can and will doctors use them (in Danish), November 1, 2004, pp. 4044–4047. Ugeskrift for læger (2004)
9. Dillon, A., Richardson, J., McKnight, C.: The effect of display size and text splitting on reading lengthy text from the screen. Behaviour and Information Technology 9(3), 215–227 (1990)
10. Duchnicky, R.L., Kolers, P.A.: Readability of text scrolled on visual display terminals as a function of window size. Human Factors 25, 683–692 (1983)
11. Fisher, S., Stewart, T.E., Mehta, S., Wax, R., Lapinsky, S.E.: Handheld Computing in Medicine. Journal of the American Medical Informatics Association 10(2), 139–149 (2003)
12. Geven, A., Sefelin, R., Tscheligi, M.: Depth and Breadth away from the Desktop – the Optimal Information Hierarchy for Mobile Use. In: MobileHCI 2006, Helsinki, Finland, September 12–15, 2006, pp. 157–164 (2006)
13. Hayes, G.R., Abowd, G.D.: Tensions in designing capture technologies for an evidence-based care community. In: Proceedings of CHI 2006, pp. 937–946. ACM, New York (2006)
14. Hejmadi, M., Winkel, B.G.: Tab 1750 gram. Stud. Med. 86, 42–50 (2003)
15. Jones, M., Marsden, G., Mohd-Nasir, N., Boone, K., Buchanan, G.: Improving Web interaction on small displays. Computer Networks 31, 1129–1137 (1999)
16. Kakihara, M., Sørensen, C.: Mobility: An Extended Perspective. In: Proceedings of the 35th Annual Hawaii International Conference on System Sciences HICSS 2002. IEEE, Los Alamitos (2002)
17. Kanstrup, A.M., Boye, N.: Theory meets practice in the design of e-support for junior registrar doctors. Artifact 1(3), 190–197 (2007)
18. Kanstrup, A.M., Boye, N., Bertelsen, P., Christiansen, E., Nøhr, C., Stage, J.: Mobile e-Learning for Young Physicians. In: Experiences from the MINI Project. Virtual Centre for Health Informatics, Aalborg University (2007)
19. Kanstrup, A.M., Boye, N., Nøhr, C.: User-driven design considerations of m-learning services for residents – activation of a theoretical model of clinical knowledge. In: MedInfo. 2007 (2007)

20. Kjeldskov, J., Stage, J.: New Techniques for Usability Evaluation of Mobile Systems. International Journal of Human-Computer Studies 60(4-5), 599–620 (2004)
21. Lam, H., Kirkpatrick, A.E., Dill, J., Atkins, M.S.: Effective Display of Medical Laboratory Report Results on Small Screens: Evaluation of Linear and Hierarchical Displays. International Journal of Human-Computer Interaction 21(1), 73–89 (2006)
22. Lu, Y., Xiao, Y., Sears, A., Jacko, J.A.: A review and a framework of handheld computer adoption in healthcare. International Journal of Medical Informatics 74, 409–422 (2005)
23. MacKay, B., Dearman, D., Inkpen, K., Watters, C.: Walk 'n scroll: a comparison of software-based navigation techniques for different levels of mobility. In: Proceedings of MobileHCI 2005, pp. 183–190. ACM Press, New York (2005)
24. Mathiassen, L., Stage, J.: The Principle of Limited Reduction in Software Design. Information Technology & People 6(2-3), 171–185 (1992)
25. Mizobuchi, S., Chignell, M., Newton, D.: Mobile text entry: relationship between walking speed and text input task difficulty. In: Proceedings of MobileHCI 2005, pp. 122–128. ACM Press, New York (2005)
26. Mustonen, T., Olkkonen, M., Häkkinen, J.: Examining Mobile Phone Text Legibility while Walking. In: Proceedings of CHI 2004, pp. 1243–1246. ACM Press, New York (2004)
27. Pirhonen, A., Brewster, S., Holguin, C.: Gestural and Audio Metaphors as Means of Control for Mobile Devices. In: Proceedings of CHI 2002, pp. 291–298. ACM Press, New York (2002)
28. Rothschild, J.M., Fang, E., Liu, V., Litvak, I., Yoon, C., Bates, D.W.: Use and Perceived Benefits of Handheld Computer-based Clinical References. Journal of the American Medical Informatics Association 13(6), 619–626 (2006)
29. Rubin, J.: Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. John Wiley & Sons, New York (1994)
30. Skov, M.B., Stage, J.: Supporting Problem Identification in Usability Evaluations. In: Proceedings of the Australian Computer-Human Interaction Conference 2005 (OzCHI 2005), ACM Press, New York (2005)
31. Shrestha, S.: Mobile Web Browsing: Usability Study. In: Proc. of the 4th Intl. Conf. on Mobile Technology, Applications and Systems (Mobility 2007), Singapore, September 10-12, 2007, pp. 187–194 (2007)
32. Tatar, D.: The Design Tensions Framework. Human-Computer Interaction 22(4), 413–451 (2007)
33. Vadas, K., Patel, N., Lyons, K., Starner, T., Jacko, J.: Reading On-the-Go: A Comparison of Audio and Hand-held Displays. In: MobileHCI 2006. ACM Press, New York
34. Watters, C., Duffy, J., Duffy, K.: Using large tables on small display devices. International Journal of Human Computer Studies 58(1), 21–37 (2003)
35. Watters, C., Duffy, J., Zhang, R.: Comparing Table Views for Small Devices. In: Proceedings of the ACM Symposium on Applied Computing SAC 2005 (2005)
36. Wixon, D.: Evaluating Usability Methods: Why the Current Literature Fails the Practitioner. Interactions 10(4), 29–34 (2003)

Designing User Interfaces for Smart-Applications for Operating Rooms and Intensive Care Units

Martin Christof Kindsmüller[1], Maral Haar[2], Hannes Schulz[3], and Michael Herczeg[1]

[1] University of Lübeck, Institute for Multimedia & Interactive Systems, Ratzeburger Allee 160, D-23538 Lübeck, Germany
[2] Drägerwerk AG, Moislinger Allee 53-55, 23542 Lübeck
[3] Dräger Medical AG & Co. KG, Moislinger Allee 53-55, 23542 Lübeck
{mck,herczeg}@imis.uni-luebeck.de,
{maral.haar,hannes.schulz}@draeger.com

Abstract. Today's physicians and nurses working in operating rooms and intensive care units have to deal with an ever increasing amount of data. More and more medical devices are delivering information, which has to be perceived and interpreted in regard to patient status and the necessity to adjust therapy. The combination of high information load and insufficient usability creates a severe challenge for the health personnel with respect to proper monitoring of these devices respective to acknowledging alarms and timely reaction to critical incidents. Smart Applications are a new kind of decision support systems that incorporate medical expertise in order to help health personnel in regard to diagnosis and therapy. By means of a User Centered Design process of two Smart Applications (anaesthesia monitor display, diagnosis display), we illustrate which approach should be followed and which processes and methods have been successfully applied in fostering the design of usable medical devices.

Keywords: Smart-Applications, Safety Critical Systems, Healthcare, User Interface, OR, ICU.

1 Working Conditions in ORs and ICUs

Backhaus and Friesdorf ([1], p.45, translation by the authors) describe the "megalopolis" intensive care unit (ICU) "as an unmanageable set of hoses, cables and wires interconnected with at least an equal amount of sensors, instruments, devices or any other technical equipment. Right in the middle, the »patient« agonisingly made stationary by the equipment that is required to keep him alive" (Fig. 1). These circumstances describe both the situation in ICUs and in operating rooms (ORs). The assembled multitude of technical support systems creates highly complex and highly dynamic working conditions for the attending physicians and nurses (in the following summarized as health personnel).

As a general rule, patient monitors are the central constituent of this environment. They measure and log vital signs (e.g. heartbeat, blood pressure, breathing and anaesthetic gas concentration levels), and thus form the basis of a reliable monitoring of the patient. Patient monitors are used for anesthetized patients in ORs and ICUs as well

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 684–695, 2009.

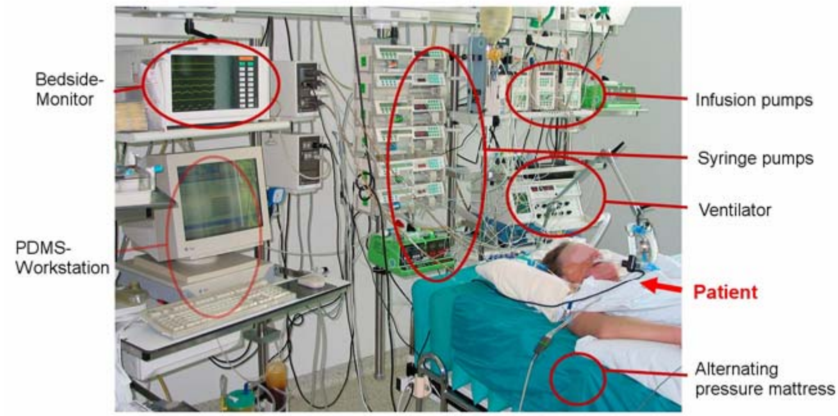


Fig. 1. Patient embedded between technical equipment on a modern ICU

as for conscious patients in general wards and allow health personnel to quickly diagnose and react on critical patient conditions. The interaction with medical devices is often difficult and cannot be effective without a huge amount of previous knowledge. Furthermore the medical devices from different manufacturers often vary in their operating philosophy. Patient monitors for example often have different controls/keys with a different layout. Thus nurses often have to adapt to a different handling while going from one patient to another or interacting with various devices (e.g. ventilator, patient monitor and syringe pump) at one patient. Obradovich and Woods [2] and Wears and Cook [3] show that these devices often suffer from bad usability. This deficiency is of particular importance because the workload of many users in the medical domain does not give them enough time to carry out in-depth examinations of these technical devices [4]. Health personnel often refuse to change visual settings of devices because they fear that they could end up with an even worse layout and not being able to change it back.

But even if every specific device would be ergonomically designed, numerous problems will only manifest themselves in the interaction with other devices in a certain working context. Some of the problems emerge from divergent design principles. In addition, there is insufficient interconnection and integration within the various devices compiling the patient monitoring system assembly (Fig. 1). This forces users to type in identical data into different devices, or to read a measurement at one device and type it into at another device. Taking these situations into account, the lack of acceptance for these devices is hardly surprising.

1.1 Stress and Habituation Effects

Numerous studies (cf. [5] for a review) document that the cognitive system of humans is generally not well adapted to monitoring activities. The errors occurring in monitoring tasks are mostly a consequence of the high attentional load, caused by the enormous amount of information and the large number of signals that have to be processed. In the case of health personnel using monitoring devices, Coiera et al. [6] found evidence for the following three possible types of errors: (1) Users see the information on a monitor, but they are not interpreting it. (2) Users concentrate on a very small amount of the data and ignore relevant information. (3) If users are tied up with performing a complex task, they avoid spontaneously addressing themselves to another task. Thus severe problems are sometimes detected too late.

It has been shown however that an adequately designed device can support the human information processing considerably in relation to monitoring tasks. Parasuraman et al. [5] emphasise the possibility of exploring the system behaviour, whereas Norman [7], recommends that the system should communicate its internal control decisions to the user.

1.2 Users in Healthcare as a Challenge

As in every design task that needs to incorporate the wide experience and expertise of users, the quality of the system design is directly dependent on how effective clinical expertise can be fed into the development process and applied to the product. The characteristics of this medical field are: the comprehensive education of the users, a language that is hard to understand for the medical laymen, as well as the far ranging legal requirements set by regulatory authorities. Due to the well known limitations in expressing expertise simply questioning the health personnel is not very promising [8] [9]. Many well established methods of user and task analysis like in situ observations and interviews or even usability tests are almost impossible to accomplish, because of the jeopardy of life and limb for at least some of the parties concerned. The everyday working life of health personnel is largely beyond the realm of experience of designers and developers. Therefore, it is a challenge to devise the design requirements that follow from this process and to turn these into an adequate design.

There are additional factors that make it difficult to optimise the design for a specific user group. For example, in Switzerland or France anaesthetic nurse specialists can (under the supervision of an anaesthesiologist) apply narcoses, whereas in Germany the application of a narcosis is mandatorily done by anaesthesiologists. This differentiation occurs from country to country, but also within each country, hospital, and even hospital ward when allocating tasks between medical specialist und clinical nurse specialists. This leads to the situation that persons with the same occupational title and the same job description take on different responsibilities, e.g. in adjusting the dose of the medication or adjusting the parameters of the respirator. Therefore, it is difficult to make generalisations from single observations or interviews in order to design one system for a certain user group.

1.3 Decision Support in Medical Devices Systems

When patients stay in ICUs and ORs they are often in critical conditions. Therefore, in some cases, the health personnel have to arrive at decisions very quickly. Nevertheless, depending on situation and case, they have to take up to 65 to 100 parameters spread over several devices into consideration [10]. Hecker and Hölscher [11] show that it takes a lot of effort for the health personnel to monitor this multitude of distributed information and parameters on the devices and – in case of an alarm – to efficiently identify the causes and react accordingly. To cope with the immense data volume on ICUs and ORs, decision support systems are currently being discussed

[12]. According to [13], [14], and [15] it is expected that these systems can increase the quality in healthcare by improving the performance of the health personnel, by enhancing patient safety, by saving time, and by cutting down treatment expenses at the same time.

However, there are great challenges for introducing decision support systems in the medical domain. Particularly the acceptance of a system by the users is often an issue [14]. The central question asks how a system could be integrated in the existing workflow. It not only has to be compatible with the existing hardware, but it has to be technically mature und updateable [14]. Mueller et al. [16] state that comprehensive analyses have to be carried out before one can integrate these systems into the workflow. Kawamoto et al. [17] specify this request and give three preconditions to advance the application of decision support systems in clinical practice: (1) decision support has to be provided automatically as part of clinician workflow, (2) decision support has to be delivered at the time and location of decision making, and (3) the system has to provide actionable recommendations. Applications that are designed in order to meet these preconditions are in the following referred as *Smart Applications*.

In summary, decision support systems seem to have the potential to improve the situation in the ICUs and in the ORs. The question remains: How can one design medical devices that achieve both a high degree of usability and a widely acceptance by the health personnel? In the following section we present two examples how the early involvement of users can help to deal with these challenges.

2 Designing Smart Applications

In the first example, the Anaesthesia Display „SmartPilot“, the challenge was to turn a scientific lab prototype into a product that supports anaesthesiologists in administering anaesthesia. The prototype uses mathematical models to calculate the concentration at the effect site and the resulting effect. This is based on statistical distributions. Based on drug dosage and patient traits the most probable course of drug concentration and resulting effect is displayed. The mathematical models also predict those values for the future. As the prototype displayed much information, which was not familiar to most anaesthesiologists, it was not very useful to them. Important requirements for a display were self-explainability and – especially in the case of real time data – allowing fast interpretation. The aim was to change the design of the display accordingly.

The second example is a Diagnosis Display, which in case of an alarm gives suggestions for the underlying cause, based on a multivariate analysis of the measured vital signs (blood pressure, heart rate, oxygenation, etc.). The focus of this research project was to study how diagnoses as well as their reliability can be displayed adequately and how such a system needs to represent itself to be considered as support by the users (and not e.g. as paternalism).

The two examples have been designed in a User Centered Design process. The process started with observations of the workflow, workplace and overall behaviour of 8 (SmartPilot) respectively 6 (Diagnosis Display) users in their everyday environment at work. Interviews with further health personnel completed the insights. Based on the results mock-ups for the devices have been created that were iteratively improved in dialogue with health personnel at different hospitals. As a next step prototypes have been implemented and tested in usability tests. The samples for the tests included 5 physicians and 4 nurses (Diagnosis Display) respectively 10 anaesthesiologists (SmartPilot). Finally an improved prototype of the SmartPilot has been tested in a simulator study with a high fidelity patient simulator with another 10 anaesthesiologists.

2.1 Anaesthesia Display "SmartPilot"

Both projects have in common, that new features are to be made available. To succeed it needs to be understood, how those features fit into the daily routine of the users, and to what extent they are already represented in their minds and working procedures. Self evidently, anaesthesiologists take into account how drugs distribute in the patient's body and which effect they cause. But to do so, they use rather heuristics because accurate calculations are almost impossible without technical support. Intra-operative awareness and pain result from low drug dosages and adverse effects like hemodynamic shock and breath depression are caused by high dosages. All of them can result in increasing healing time and/or permanent damage. The aim is to reduce the workload and to support the personnel to apply the adequate procedure.

First observations in ORs showed that anaesthesiologists do not want an additional device to be monitored at their workplace. Anaesthesiologists often work in crowded

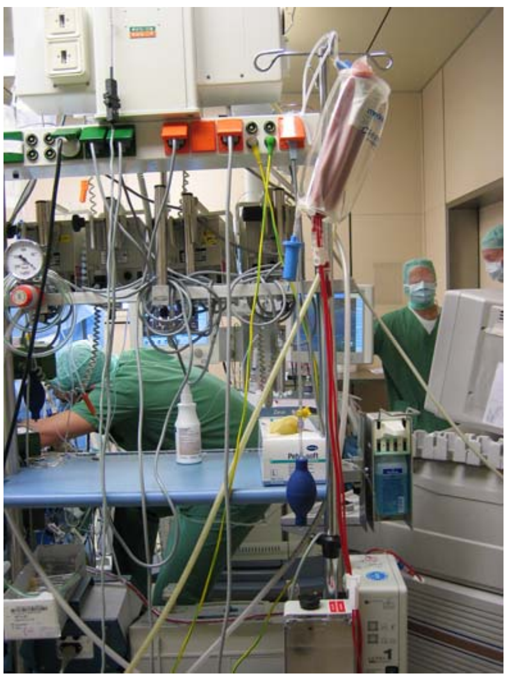


Fig. 2. Anaesthesia workplace in an OR

places, virtually in the middle of an equipment park consisting of anaesthesia machine, patient monitor, syringe pumps, warming therapy devices and more (Fig. 2). It seemed reasonable to add information the anaesthesiologist needs to track anyway to detect changes in the patient's state to the new display. This way the most important information is available in one place (the new "SmartPilot" display) and should only require looking at the other devices like anaesthesia machine, patient monitor, syringe pumps etc. to explore parameters in more detail or to adjust drug administration or change other settings.

As a simplified description of their major goal of work the interviewed anaesthesiologists could agree to the following statement: "anaesthesiologists want to keep the hemodynamic signs of the patient stable in adjusting the depth of anaesthesia to the changing pain level resulting from the surgery". In the mathematical models, the depth of anaesthesia is calculated as a probability to tolerate noxious stimuli. To allow the anaesthesiologist to stick with his mental models, this calculated anaesthetic depth was displayed with the most important hemodynamic parameters on a shared time scale (Fig. 4, lower right side). This allows putting the calculated effect into perspective based on hemodynamic reactions.

The drug dosage as well as the calculated concentration at effect site is displayed accordingly. That allows the anaesthesiologists to see the coherence of drug dosage, concentration at effect site, calculated effect, and observable effect on patient at a glance. That also helps to relate the individual patient to the standard, as described by the models. In a next step these findings were realized as sketches and discussed and revised iteratively with the help of several anaesthesiologists (Fig. 3).

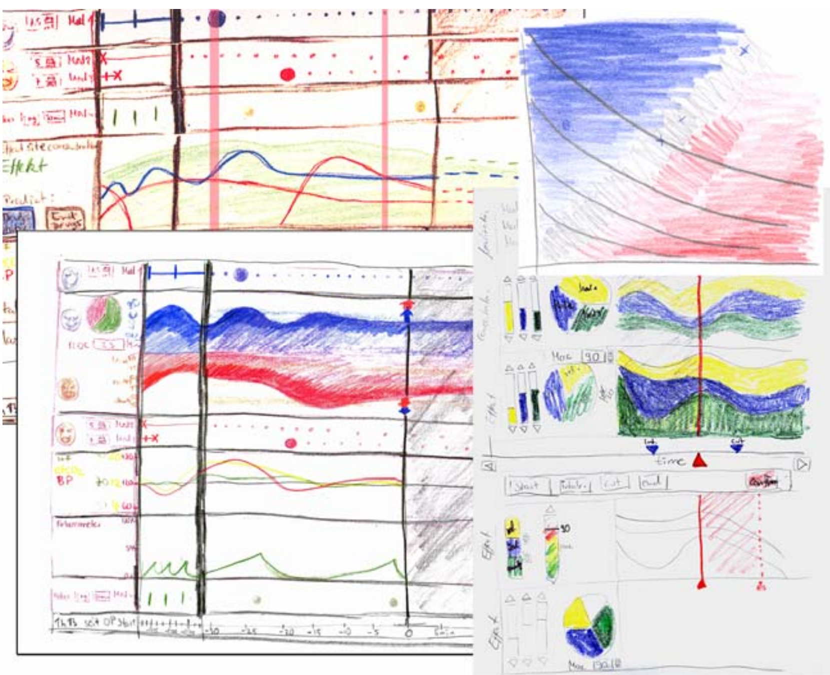


Fig. 3. Sketches which were iteratively discussed and improved with anaesthesiologists

Fig. 4. "SmartPilot" Prototype state of December 2007

In the next phase a functional prototype was developed and formal usability studies have been conducted. In the end, a display evolved which includes the most important vital signs as deviation from a norm, the combined effect of analgesic and hypnotic drugs as a graph as well as an index which was deliberately developed for this display to quantify the anaesthetic effect (Fig. 4). Furthermore, the concentration of each drug at effect site is displayed in relation to the according dosage. In relating the vital signs as an indicator for the current patient state with the calculated drug concentrations, the anaesthesiologist is able to relate the drug reaction of the individual to the norm. All information is displayed on a shared time scale to simplify the contextual interpretation. On this time scale additionally events like "skin incision" can be marked. As SmartPilot combines measured variables with model based calculations, it is crucial to visually separate the two. Therefore all vital signs are in a separate box, all calculated drug concentrations, and the indices BIS and NSRI also have a single box each.

2.2 Diagnosis Display

During the design process of the Diagnosis Display one central question was how the reliability of a diagnosis can be communicated to the user. This question occurs because it is rarely possible that an expert system can derive a diagnosis that is absolutely certain. Additionally some steps of the physicians' workflow (e.g. finding out what favours a diagnosis) are not supported when the plain diagnosis is displayed. Eventually theses actions may be even hampered. Thus the system should be designed in a way that "overtrust" or "undertrust" problems can be avoided [18]. Actually the

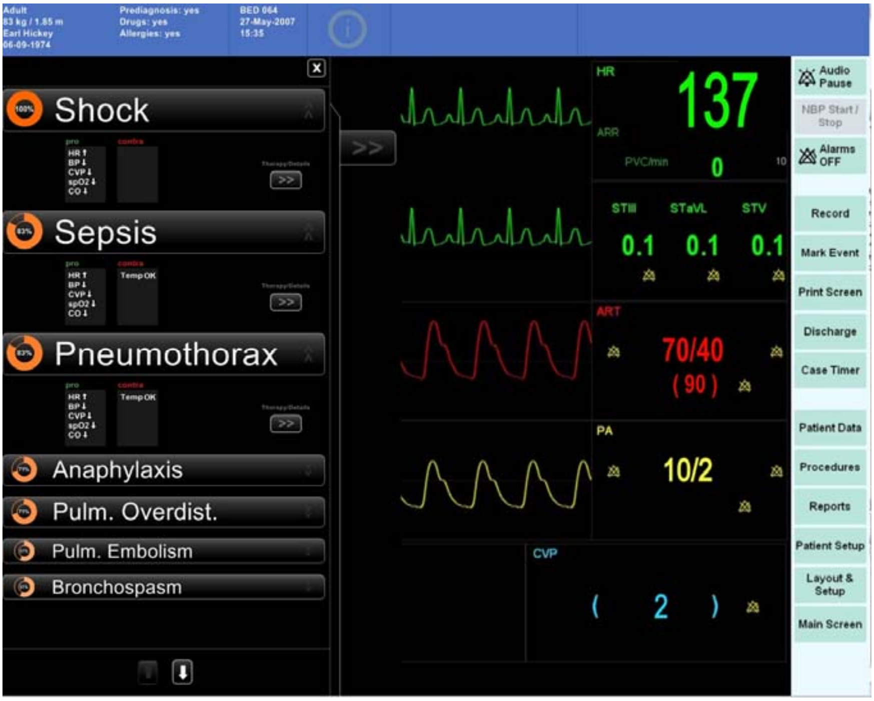


Fig. 5. Diagnosis Display with proposed diagnoses. The validity and reliability of a diagnosis is communicated by the orange segments of circles and by variations in font height and size.

results of usability-tests and interviews showed that the knowledge of a diagnosis' reliability is crucial for many participants. Furthermore, a visual presentation of the reliability is preferred over a numeric one. A numeric presentation usually stands for a high grade of accuracy that cannot be guaranteed in most cases. The validity and reliability of a diagnosis can be communicated by coloured geometrical shapes (orange segment of circle) and variations in font height and size. Numeric expressions like a percent value are only provided additionally (Fig.5).

Another important aspect was the acceptance of the Diagnosis Display. As Berner and Moss [13] point out, systems are particularly accepted, if the user is able to understand how the device works and how diagnoses are calculated. The extent of decision support can affect the acceptance as well. Several physicians (as participants in the usability-test) commented on certain drafts of the Diagnosis Display that the decision support would be too extensive. They said that the display shows trivial diagnosis and information that they already know for sure in this situation. Some nurses however considered this information as useful. A few physicians agreed and commented that the display would be useful for certain situations (e.g. admission of a patient, cognitive overload, education). This discrepancy shows that it is essential to focus on certain users and situations when a decision support system like this is introduced. The assumption is underlined by the following finding: The majority of the physicians consider the display as an inspiration and source for their own decision

Fig. 6. Diagnosis Display with the possibility to add clinical observations

making process. They would like to see lots of information and diagnosis on the display. The majority of the nurses however are primarily interested in symptoms, not diagnoses, because these symptoms trigger the nurses' treatment activities, without the prior involvement of a physician. E.g. one nurse stated: "diagnoses are not important for me".

The interviews and tests show that physicians and nurses are critical about additional devices at the workspace. In fact, a good integration into an existing device and the ability to communicate with other devices is required. Here the Diagnosis Display was integrated into a patient monitor to provide information at the point of care (Fig. 6). Asked about the integration the majority of the participants said that the Diagnosis Display should only be visible on demand. They wanted to keep the monitor they are used to as their primary device for supervising. However they expected some kind of notification when the Diagnosis Display has new information for them.

The participants also criticised that the display can only use parameters that are measured by devices. If the display needs to reproduce the process the physicians go through, clinical observations of the patient (e.g. tearing, sweating) are missing. This information is very important for the physicians. As a solution these observations can now be chosen from a list and affect the calculation of a diagnosis when using the display (Fig. 6). However these observations often must be entered into a Patient Data Management System too. This conflicts with the precondition that users do not want additional work with the Diagnosis Display. Therefore these clinical observations should be synchronized between the Diagnosis Display and the Patient Data Management System.

2.3 Conclusion from the Case Studies

One major challenge in the design of the Anaesthesia Display "SmartPilot" and the Diagnosis Display is to implement the support feature without suffering complacency or overtrust (as described above). The display needs to be designed in a way requiring the users to make use of their expertise despite all support. This necessity was mentioned by some of the participants of the usability tests as well. They frequently noted that especially inexperienced colleagues need to be challenged to learn diagnosing or administering anaesthesia appropriately and effectively.

Both projects show, that it is crucial to the users to have those new features and information available in parallel to their standard working procedures. They want a system that is – at first – only observed whether it comes to the same conclusions. The decisions are at that time still made by the users themselves. Therefore with "SmartPilot" a product is developed which visualizes the distribution and effects of the drugs based on their dosage according to the underlying models. This is displayed in context of selected vital signs. The dosage of the drugs is carried out as usual by the anaesthesiologists at the anaesthesia machine and the syringe pumps.

The diagnosis display only exists as a research prototype and is meant as an optional component of a patient monitor. The user can turn the display on and off as he or she wishes, allowing to compare the results of the display with their own conclusions without any pressure to make use of it.

3 Designing Working Conditions Instead of Optimizing Devices

As these examples already show, it is crucial in the development of new medical devices to incorporate the lessons learned from the adversities found in clinical workaday life. If, for example, a user already has to monitor more devices than he or she is able to, adding a new device further increases this cognitive overload. Therefore, any new device added should replace at least one other device. If users are already displeased by typing in the same information into several devices, this problem has to be avoided when reengineering devices or developing a new class of devices. Ideally, users enter data into one task-oriented device, which then propagates the data into all other devices. A common standard[1] for the interchange of patient data is required to achieve this.

Another important lesson learned from both our case studies, is that users within a safety critical clinical environment have to act conservatively. Therefore, they tend to be critical towards innovative devices. This is because they cannot be sure that their expertise is considered and supported by the devices. Furthermore, they have to make sure that theses devices will not become a security risk for their patients. Ideally, the health personnel should be able to actively validate, whether the new system comes to the same conclusions as themselves, or analyse, why the device comes to other

[1] DICOM [19] shows that it is possible to establish a common data structure and communication protocol within the field of medical imaging.

conclusions. Depending upon whether the system has proved its value to the user, they can gradually utilize more of the functions of the device. Thus, they can gradually shift their way of working with the device in way that feels save to them. Particularly in support systems, it is of utmost importance that the systems themselves know their limits and can communicate these limits to the users [20]. Hence, an important question in regard to the Diagnosis Display is: How to visualize the reliability of a diagnosis to enable the users to avoid relying on the system when they should not? This holds particularly in the case, when the system itself does not have sufficient data or if the data is of questionable quality.

It is important to always leave the full freedom of decision over the medical device to the user. Every respirator offers the possibility to run the respiration manually. Every control loop at a syringe pump that automatically adjusts a given control variable, can be interrupted at any time and overwritten with manual settings. Last but not least, physicians want to see and touch their patients; they never rely merely on the patients' records when making decisions. Physicians are more willing to consider devices for diagnosis support if they can enter their clinical observations into the device, so that this data will be considered in the diagnosis computation by the device.

In developing medical devices used by nurses, it should be considered that both the competence, as well as the area of accountability of nurses differs sometimes highly from country to country, from hospital to hospital, and even from ward to ward. Some nurses with certain skills and responsibilities need to be informed about the general state of the patient to make qualified decisions about the patient's care. Other nurses work conclusively when instructed to do so, and otherwise, they carry out only assigned tasks of basic patient care. Developing products for such heterogeneous organisational application contexts is therefore a great challenge.

To disclose theses problems, restrictions and characteristics, it is important to involve the users early, repeatedly and consistently during the development process of the device. This way, the User Centered Design process leads – through early feedback, short iteration cycles and usability tests – to the understanding of context, interrelationships, work conditions, and problems of the target domain. As a result, this acquired knowledge can be used to design a better device suited well for its users, their tasks and their work contexts.

Acknowledgements. We want to thank all the participants of our usability-tests, simulator-tests, interviews, surveys, and workshops. We appreciate their willingness to critically reflect and discuss the new systems and for their patience to communicate their concerns to non-medics. Especially the physicians and nurses of the Universitätsklinikum Schleswig-Holstein, Campus Lübeck were very helpful in allowing us to look over their shoulders and observe them in ORs and ICUs. Last not least we are very grateful to Dr. Michael Imhoff, who provided us with photos of ORs and ICUs.

References

1. Backhaus, C., Friesdorf, W.: Verloren im Ballungsraum der Intensivstation. In: Forschung Aktuell 2006 Gesundheitstechnologien. Technische Universität Berlin, Berlin (2006)
2. Obradovich, J.H., Woods, D.D.: Users as Designers: How People Cope with Poor HCI Design in Computer-Based Medical Devices. Human Factors 38(4), 574–592 (1996)

3. Wears, R.L., Cook, R.I.: Automation, Interaction, Complexity, and Failure: A Case Study. Reliability engineering & systems safety 91(12), 1494–1501 (2006)
4. Matern, U., Koneczny, S., Scherrer, M., Gerlings, T.: Arbeitsbedingungen am Arbeitsplatz OP. Deutsches Ärzteblatt 103(47), 3187–3192 (2006)
5. Parasuraman, R., Molloy, R., Mouloua, M., Hilburn, B.: Monitoring of Automated Systems. In: Parasuraman, R., Mouloua, M. (eds.) Automation and Human Performance: Theory and Applications. Lawrence Erlbaum, Mahwah (1996)
6. Coiera, E.W., Tombs, V.J., Clutton-Brock, T.H.: Attentional overload as a fundamental cause of human error in monitoring. Hewlett Packard Laboratories Technical Report (1996)
7. Norman, D.A.: The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation'. Philosophical Transactions of the Royal Society of London 327, 585–593 (1990)
8. Dreyfus, H.L., Dreyfus, S.E.: Mind over Machine. Free Press, New York (1986)
9. Ericsson, K.A., Charness, N.: Expert performance: Its structure and acquisition. American Psychologist 49, 725–747 (1994)
10. Friesdorf, W.: Überwachung am Intensivbett - Informationsrepräsentation. In: Friesdorf, W., Schwilk, B., Hähnel, J. (eds.) Ergonomie in der Intensivmedizin. Bibliomed Medizinische Verlagsgesellschaft, Melsungen (1990)
11. Hecker, E., Hölscher, U.: Informationsverarbeitung am Erwachsenen-Intensivbett – ein Lösungsansatz. In: Friesdorf, W., Schwilk, B., Hähnel, J. (eds.) Ergonomie in der Intensivmedizin. Bibliomed Medizinische Verlagsgesellschaft, Melsungen (1990)
12. Sailors, R.M., East, T.D.: Clinical informatics: 2000 and beyond. In: Proceedings AMIA Symposium (1999)
13. Berner, E.S., Moss, J.: Informatics Challenges for the Impending Patient Information Explosion. Journal of the American Medical Informatics Association 12, 614–617 (2005)
14. Garg, A.X., Adhikari, N.K.J., McDonald, H., Rosas-Arellano, M.P., Devereaux, P.J., Beyene, J., Sam, J., Haynes, R.B.: Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes. A Systematic Review. Journal of the American Medical Association 293, 1223–1238 (2005)
15. Payne, T.H.: Computer Decision Support Systems. Chest 118, 47–52 (2000)
16. Mueller, M.L., Ganslandt, T., Frankewitsch, T., Krieglstein, C.F., Senninger, N., Prokosch, H.U.: Workflow analysis and evidence-based medicine: towards integration of knowledge-based functions in hospital information systems. In: Proceedings of the AMIA Symposium, p. 330 (1999)
17. Kawamoto, K., Houlihan, C.A., Balas, E.A., Lobach, D.F.: Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. British Medical Journal 330, 765–772 (2005)
18. Parasuraman, R., Miller, C.A.: Trust and etiquette in high-criticality automated systems. Communications of the ACM 47(4), 51–55 (2004)
19. DICOM. Medical Imaging & Technology Alliance, `http://medical.nema.org`
20. Herczeg, M.: Intention-Based Supervisory Control - Kooperative Mensch-Maschine-Kommunikation in der Prozessführung. In: Grandt, M., Gärnter, K.-P. (eds.) Situation Awareness in der Fahrzeug- und Prozessführung. Deutsche Gesellschaft für Luft- und Raumfahrt, Bonn (2002)

Interactive Therapeutic Multi-sensory Environment for Cerebral Palsy People

Cesar Mauri[1], Agusti Solanas[2], Toni Granollers[1], Joan Bagés[3], and Mabel García[4]

[1] GRIHO, HCI research group (University of Lleida)
[2] CRISES Research Group (Rovira i Virgili University)
[3] Centre de Reserche Informatique et Creation Musicale (University of Paris)
[4] Associació Provincial de Paràlisi Cerebral, Tarragona
cesar.mauri@urv.cat, agusti.solanas@urv.cat, tonig@diei.udl.cat, joanbir@hotmail.com, mggs@tinet.org

Abstract. The Interactive Therapeutic Sensory Environment (ITSE) research project offers new opportunities on stimulation, interaction and interactive creation for people with moderate and severe mental and physical disabilities. Mainly based on computer vision techniques, the ITSE project allows the gathering of users' gestures and their transformation into images, sounds and vibrations. Currently, in the APPC[1], we are working in a prototype that is capable of generating sounds based on the users' motion and to process digitally the vocal sounds of the users. Tests with impaired users show that ITSE promotes participation, engagement and play. In this paper, we briefly describe the ITSE system, the experimental methodology, the preliminary results and some future goals.

Keywords: Cerebral Palsy, Elderly, Disabled People, Artificial Vision.

1 Introduction and Motivation

This article presents the Interactive Therapeutic Sensory Environment (ITSE) project. Its main aim is to investigate, develop and apply a system able to analyze the user's gestures and vocal sounds (by means of web cameras and microphones), and to transform them into images, sounds and vibration. By doing so, the system creates a sensitive space with a multimodal feedback highly coupled with the input.

The project was born in the APPC and the reduced number of users that were able to interact with computers was the main motivation. Although several assistive technology gadgets and special software applications were used, most users "simply" could not understand the interaction procedures. We believed that a high interactive activity (reinforced with sounds and images closely related with gestures) would be more easily accessible to most users despite their sensory, motor and cognitive impairments.

---

[1] APPC (in Catalan stands for Associació Provincial de Paràlisi Cerebral), Cerebral Palsy Centre of Tarragona. APPC is located in Catalonia (Spain).

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 696–699, 2009.

Our goal is to turn the system into a tool that has plenty of activities allowing users to explore, express, play and enjoy. Moreover, by means of these activities, users may develop capabilities –sometimes unconsciously– such as participation, communication or creativity and, consequently, they improve their quality of live and wellbeing. The project focuses on people with cerebral palsy but it could also benefit people with other cognitive disabilities.

2 State of the Art

In spite of the importance of studying the interaction capabilities of cerebral palsy people, the number of related references in the literature is quite limited. Next three main lines/projects constitute the most relevant advances up to now:

1. **P. Ellis and Soundbeam.** In the early 90s, P. Ellis started to use the so-called sound therapy with Soundbeam[2] on children with autism and profound and multiple learning disabilities. Sound therapy combines the power of a contactless interactive space with an aesthetic response to sound. This encourages users' interaction and communicative skills development 4. In 1998, Ellis introduced the VibroAcoustic Sound Therapy that included a microphone and a vibroacoustic device (VAD) to the system. Since 2004 Ellis, from the iMUSE laboratory, focuses on the elderly and studies the addition of image feedback to the system 5.
2. **T. Brooks and Soundscapes.** Soundscapes 1 refers to a body function capture library and to a collection of software able to generate a response from the gesture. One of the uses of Soundscapes is as an "expression amplifier" for people with disabilities. It allows users to generate images and sounds from body motion. The main idea is to create a virtual interactive space (VIS) where users like to be and where motivation and creativity are promoted through play and fun. Several European projects (e.g. 2) show the therapeutic use of Soundscapes.
3. The **Mediate** project 8 proposed a multi-sensory installation whose main goal is to let children with autism to have fun, play, explore and create in a controlled and secure environment. The project carried out psychological studies and provided parents with an environment where see their children playing.

In addition to these three main lines, other projects can be found in the literature, namely Intellivision 3 and Virtual Music Instrument (VMI) 6.

3 The ITSE Project

Promoting the participation of people with severe mental disabilities –that normally remain passive– would lead to a series of reactions that would foster their development. These people are usually withdrawn and isolated from their surroundings due to their limited communication abilities. Providing them with a channel that let them communicate, express feelings, and even create, will contribute to the improvement of their quality of life. To cope with this task it is necessary to have the right tools that

---

[2] Soundbeam is a device that uses an ultrasound sensor to capture user's motion and translates it into sounds (in form of MIDI notes).

provide multimodal interaction in an efficient, simple, and cheap way. Based on our previous experiences 7, we believe that computer vision techniques can play an important role due to their versatility. The ITSE system consists of two software modules: (a) The **motion capture**, an application that captures and processes a video stream in real time. It extracts several parameters such as the direction and quantity of movement, the position of a color marker and the index of global activity. And, (b) The **audiovisual generation**, an application developed using Max/MSP. It processes the information from the capture module and generates images and sounds. MIDI protocol is used for the communication between both modules. Its main features are: (i) a real-time digital sound effect processor (including effects like reverb, chorus, pitch change and echo) that allows users to play with their voice, (ii) a motion-based virtual instrument that allows users to play MIDI instruments by means of their gestures, and (iii) a sound visualizer (using an external application called R4[3]). In order to test the ITSE system we have equipped, in the APPC facilities, a quiet room with a computer, a sound system, a large screen, a web camera and a camcorder. Also we have selected several participants for the experiments. Currently, we distinguish among three groups of people according to their cognitive abilities: (i) people able to understand the procedure of the activities, (ii) people that initially do not understand the procedure of the activities but can understand it after some training and, (iii) severe cognitively impaired people that cannot understand the activities. Usually three people are present in the experiments: the researcher (managing the system and taking annotations), the facilitator (a professional close to the user, typically a teacher or therapist) and the user.

Experimentation sessions (between 15 and 30 min.) start welcoming the user and suggesting him an activity (essentially playing with his voice or playing a virtual instrument). This suggestion and the specific arrangements of the activity (level of complexity/difficulty and other parameters) depend on the abilities of the user and his success rate observed in previous sessions. The facilitator starts encouraging the user to interact, but without forcing him. In fact, we try to intervene as little as possible. Depending on the user behavior during the session, the researcher can smoothly change the parameters of the activity or switch to another one. The session ends when the assigned time span expires or when the user becomes tired. The whole session is recorded in video including the final comments made by the researcher and the facilitator.

The analyses of video recordings and annotations include studying several variables such as the type of attention and span, satisfaction, inhibition, etc. After a year of experimentation, we can conclude that the ITSE project promotes attention, perception, communicative intention, muscular control, expression, "connection" with the surrounding environment, imagination, participation, play and funny engagement. For instance, for some people we are specifically interested in seeing if action-reaction relationship is well understood. To prove that, we switch off audio and video for a while and look for user's behavior changes, like looking at the facilitator or changing the facial expression. It allowed us to demonstrate that some users are able to understand the action-reaction relationship. Other results include the presence of previously unseen behavior or physical movements. However, the studies are still in a preliminary stage. A demonstration video is available in 9.

[3] http://www.rabidhamster.org/R4/

4 Conclusions and Future Work

Although the idea of a contactless interactive multi-sensory environment is not new, our proposal tries to make such systems wide available reducing costs and simplifying the setups. This is a challenging project due to the severe impairments of the users and the interdisciplinarity of the task. However, after one year of experiments, results are very promising. Future improvements include to add a learning/adaptive module, to better characterize user's profiles by means of ontologies, to deploy a detailed analysis methodology, to improve usability to avoid the need for technical experts, and to develop a software framework in which new activities and sensors could be easily added.

References

1. Brooks, T.: Virtual interactive space (VIS) as a movement capture interface tool giving multimedia. In: Intl. Congr. World Confed Physical Therapy, Japan (1999)
2. Brooks, T.: CAREHERE. In: Intl. Conf. Disability, Virtual Reality & Assoc. Tech., Oxford, England, pp. 191–198 (2004)
3. Brumback, C., Borrisov, L., Galusha, J., DiIorio, A.: Intellivision. In: The 6th Intl. Conf. Ubiquitous Computing, UK (2004)
4. Ellis, P.: Incidental music; a case study in the development of sound therapy. The British Journal of Music Education 12, 59–70 (1995)
5. Ellis, P.: The development of interactive multisensory environments for expression. Keynote to Luxembourg Society for Music Therapy (2006)
6. Lamont, A., Knox, R., Chau, T., et al.: Converting movements to music new musical exploration opportunities for children in rehabilitation. In: Canadian Asso. for Music Therapy 29th Annual Conference, Regina, Canada, pp. 26–30 (2002)
7. Mauri, C., Granollers, T., Lores, J., García, M.: Computer vision interaction for people with severe movement restrictions. Human Tech. 2(1), 38–54 (2006)
8. Pares, N., Carreras, A., Durany, J., Ferrer, J., Freixa, P., Gomez, D., Kruglanski, O., Pares, R., Ribas, I.J., Soler, M., Sanjurjo, A.: Promotion of creative activity in children with severe autism through visuals in an interactive multisensory environment. In: Proc. of the 2005 conference on Interaction design and children, pp. 110–116. ACM, Boulder (2005)
9. `http://video.google.es/videoplay?docid=-6074387707245897409`

Designing Systems for Health Promotion and Autonomy in Older Adults

Helena Lindgren[1] and Ingeborg Nilsson[2]

[1] Department of Computing Science, Umeå University, SE-901 87 Umeå, Sweden
[2] Department of Community Medicine and Rehabilitation, Unit for Occupational Therapy, Umeå University, SE-901 87 Umeå, Sweden
helena@cs.umu.se, ingeborg.nilsson@umu.se

Abstract. The inclusion and autonomy of older people in the society where large parts of the life is organized with computer and Internet use as means is addressed in an ongoing project in the rehabilitation and health domains. Part from investigating the potentials of using ICT for rehabilitation of older people with limited or no computer skills, the aim for the project is to develop methods and tools for the purpose, and also for the interaction design domain where systems are developed for older people. The resulting methods are used for informing the design of the system in an iterative process.

Keywords: Interaction design, activity theory, health care, evaluation.

1 Introduction

Activities are, to an increasing extent, re-organized to involve the web as a means, as source of information and entertainment, channel for communication, base for social environments and for administration of individual's personal and social life. This trend excludes persons who have not access, knowledge or the skills to use computers and Internet facilities from a large part of the society and often also parts of the social life of the immediate family. To a large extent, older persons are characterized by this exclusion and would benefit from purposeful designed systems aimed at supporting the execution of activities of their choice. This category of aids to overcome hindrances to perform needed activities in daily life constitutes a significant asset for occupational therapy interventions. As part of the professional reasoning to evaluate activity barriers in individual clients, the adjustments of the environment in which the activity is to be performed is done on a regular basis in the rehabilitation process. Therefore, the professional assessment of the needs for certain computer and Internet based aids constitutes an activity that will increase in importance by the increased demands of ICT-mediated distributed activities in people's daily lives. However, this requires also the development of tools and methods for the rehabilitation domain for facilitating the integration of the task into daily practice.

The rehabilitation process is also a process distributed over time and location, sometimes involving additional actors such as family, where members often are located in other parts of the country. Computer and web based tools for investigation and follow-ups could easily overcome these factors, and may consequently

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 700–703, 2009.

enhance the client's role and influence over his or her rehabilitation process if used as a common tool for choices and adjustments of interventions. Therefore, a pilot study was initiated with the aims to investigate to what extent ICT can be used as a common tool in the rehabilitation process by both occupational therapist and an older patient. In this pilot study the possibilities are explored for such a tool in rehabilitation of activity impairments in older people and where the persons may not have any or very limited experience of computer use. Specifications and prototypes for such a tool are being developed. Of central concern in our work, is also the need for tools and guidelines for the process, both regarding the rehabilitation domain and for the domain of interaction design for older adults. Developing systems aimed to be used by older persons with limited experiences of computer use is a challenging task (e.g., [1, 2, 3, 4]). This process requires both a carefully organized design process methodologically, suitable for the inclusion of user representatives in the design process, and interaction designs that meet the special needs of older persons.

2 Methods, Material and Procedure

An activity-oriented participatory design process with heavy emphasis put on the activities in focus was applied, involving two occupational therapists, a computer scientist and a group of older volunteers. This was accomplished by using activity theory and the notion of ZPD (zone of proximal development) as the theoretical framework for data gathering, analysis and re-design [5, 6, 7]. The activities observed ranged from occupational therapists investigating the participants, the participants performing activities of their choice, and the dialogue between the two for the purpose of follow up on interventions. Interviews and questionnaires were used continuously. The occupational therapist's evaluation of activity status was used as starting point, in which activity hinders, personal resources, needs and interests were identified in eight older persons who volunteered to participate in the study. The persons comprised one man and seven women between the ages of 65-75, who did not have any or very limited experiences of using computers and none using the web.

A structured protocol emerged from the results of the initial assessment phase. The protocol were used for valuing and documenting the individual's development of skills and ability in Internet-mediated activity, identify reasons for breakdown situations, and to formulate questions for interviews and computer-based questionnaires. These were used after each session for assessing the subjective views and attitudes towards computer use. Furthermore, the protocol was used for informing the re-design of prototypes.

Based on the activity analyses and a literature review, initial prototypes of a virtual café environment were developed and integrated in an already existing physical Internet café environment for older people. The subjects participated in seven to eight organized occasions of 1,5 hours, during a period of two months. One woman did not proceed to participate in the café environment due to health reasons.

3 Results

The major purpose of occupational intervention is to increase autonomy in performing activities in fields of interest and of importance for the individual. To summarize the results, all of the participants discovered new values that could be added to their activities by using computers and the web as tools. This includes new and easily accessed sources of information and knowledge, channels to organizations of interest where they are engaged, finding and communicating with old friends and family, developing interests. In addition, four persons expressed that it is simply fun to use the computer and that the enjoyment increases with control and skills. This is said in contrast to the remaining skeptical person who saw little use for, or enjoyment in computer-based activities, part from using the computer for keeping up with news.

Three levels of complexity in web pages were described by Dickinson and colleagues [4] where the first level could be seen as a "walled garden", hiding large parts of the web from the user with content fully controlled by the development team. The second layer allowed some external content that adhered to high standards of accessibility and navigation structures, while the third allowed any content of the web. A similar approach was taken in this study, increasing the complexity judged based on what lied in the next zone of development and based on what the individuals' wished to do during the next session. All participants wanted to do things eventually, which involve external web pages (some tried on their own in their homes but got lost among frames, different routines for logging in, etc.), therefore, handling such events were included in later sessions as part of their main activities. Maintaining a consistent basic structure was necessary in order to reduce the amount of new things to internalize. A resistance to change of structures was expressed, the participants wanted to be able to reuse knowledge as much as possible and not having to deal with several different ways of doing what they perceived as being the same thing. The participants found it hard to grasp differences between similar phenomenon, and one way to distinguish between them was by the procedure to use them (e.g., web addresses vs. searching by keywords, logging in). Therefore, one strategy to handle breakdowns was to start over from a point where they were familiar, which was a role the basic structure served. Therefore, this basic structure, which essential served as a home page, was adjusted to become more distinct from the rest of the web pages and easy to recognize among many open windows.

The iterative process of evaluation and re-design during the period of use, enabled the participants to contribute to the design of the system by trying to do what they wanted to be able to do, reflecting on their own performance and experienced computer-related phenomenon, without having to express explicit opinions about design choices. The latter seemed to be very difficult, since they did not distinguish differences between pages, objects, their properties, etc., to the extent that was necessary for having a specific opinion, and even less having the vocabulary to express an opinion in interviews or in questionnaires. What they did express though, was general desires for alternative ways of interacting with the computer. In spite of the simplified initial designs with introduction and support to practice basic skills, they had significant difficulties in particularly controlling the mouse and using controls in the user interface. An alternative approach to the first acquaintance to the use of computers is being developed using touch functionality.

4 Conclusions

This paper reports on results from a pilot study aimed at developing prototypes for activity support involving older novice computer users in the process. A starting point for development is taken in the purposeful and desired activities to be supported, identified in eight individual cases. A protocol was developed based on activity theory and the notion of ZPD for informing the design of the system in an iterative process of re-design, use and evaluation. The seven older novice computer users who participated in the organized sessions increased their autonomy in computer-mediated activity to the extent that they are able to perform activities of their choice independently or with occasional support, and by partially using a structured web interface as starting point in case of breakdowns. By setting focus on the activity and using the protocol as base for assessing reasons for breakdown situations, the participants were able to contribute to the design process by reflecting on their performance of activities of their choice, without having to express explicit preferences among design choices or mastering computer-related phenomenon and terms. The method appears to be highly valuable in the context it has been used in this study, which is developing systems for the rehabilitation and health domain where the older novice actors need purposefully designed computer-based tools for increasing their autonomy and skills in daily activities. The protocol is being validated with additional users and activities in order to further investigate its usefulness for informing interaction design when involving potential users who may not be articulate about design choices. This is being done in the ongoing development of an extended system based on the results.

References

1. Dickinson, A., Arnott, J., Prior, S.: Methods for human – computer interaction research with older people. Behaviour & Information Technology 26, 343–352 (2007)
2. Turner, P., Turner, S., van de Walle, G.: How older account for their experiences with interactive technology. Behaviour & Information Technology 26, 287–296 (2007)
3. Hawthorn, D.: Interface design and engagement with older people. Behaviour & Information Technology 26, 333–341 (2007)
4. Dickinson, A., Smith, M., Arnott, J., Newell, A., Hill, R.: Approaches to Web Search and navigation for older computer novices. In: Proceedings of the SIGCHI conference on Human factors in computing systems CHI 2007, pp. 281–290 (2007)
5. Vygotsky, L.: Mind in Society: The Development of Higher Psychological Processes. Harvard University Press, Cambridge (1978)
6. Engeström, Y.: Learning by Expanding: An Activity - Theoretical Approach to Developmental Research. Helsinki, Orienta-Konsultit (1987)
7. Bødker, S.: A human activity approach to user interfaces. Human-Computer Interaction 3, 171–195 (1989)

CLINICAL SURFACES – Activity-Based Computing for Distributed Multi-Display Environments in Hospitals

Jakob E. Bardram, Jonathan Bunde-Pedersen, Afsaneh Doryab, and Steffen Sørensen

IT University of Copenhagen, Rued Langgaards Vej 7
DK-2300 Copenhagen, Denmark
{bardram,jonathan,adoryab,sts}@itu.dk

Abstract. A multi-display environment (MDE) is made up of co-located and networked personal and public devices that form an integrated workspace enabling co-located group work. Traditionally, MDEs have, however, mainly been designed to support a single "smart room", and have had little sense of the tasks and activities that the MDE is being used for. This paper presents a novel approach to support activity-based computing in distributed MDEs, where displays are physically distributed across a large building. CLINICAL SURFACES was designed for clinical work in hospitals, and enables context-sensitive retrieval and browsing of patient data on public displays. We present the design and implementation of CLINICAL SURFACES, and report from an evaluation of the system at a large hospital. The evaluation shows that using distributed public displays to support activity-based computing inside a hospital is very useful for clinical work, and that the apparent contradiction between maintaining privacy of medical data in a public display environment can be mitigated by the use of CLINICAL SURFACES.

1 Introduction

A multiple display environment (MDE) is comprised of co-located personal and public devices, ranging from small devices (e.g. PDAs, laptops, and tablet PC) to large wall-based displays, which are networked to form an integrated workspace [4]. The goal of such MDEs is to foster co-located group work utilizing available devices.

An on-going research challenge is to design the technologies and user interfaces for such MDEs. Many systems and interaction techniques have been proposed [18, 8, 10, 4, 16, 17], each addressing specific issues of creating an MDE. Existing research has, however, mainly focused on MDEs within a single physical space – often referred to as a *Smart Space*. Little work has addressed what we call a "Distributed Multi-Display Environment" (dMDE), i.e. an MDE that is distributed across several physical spaces inside e.g. a large building. Furthermore, most existing research has focused on information and device management for sharing basic data on the devices. For example, sharing and exchanging documents, spreadsheets, presentation, images, etc. Less work is available on supporting whole work tasks and activities. For example, how the different documents, data, and services involved in a specific work activity may belong together, and how an MDE may help collate and present these

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 704–717, 2009.

resources in a coherent and efficient manner for the people collaborating around this activity.

This paper proposes CLINICAL SURFACES, which is a set of user-interface technologies and infrastructure elements supporting a distributed multi-display environment with a special focus on supporting clinicians working inside hospitals. Hence, the type of dMDE we are designing for is a setup where displays are scattered around the entire hospital and are available in e.g. conference rooms, patient wards, hallways, and operating rooms.

The goal of CLINICAL SURFACES is to help clinicians use publicly available displays for managing large amount of medical data while moving between different physical spaces. Key features include support for easy access to an aggregation of all information and data relevant for a specific patient; access to this data from all displays in the dMDE; support for having several users actively participating in the work on a patient activity; support for moving patient activities and their related resources around in the dMDE; support for both fixed (wall-based) and mobile public displays; support for context-aware adaptation of the access to relevant activities; and support for a seamless transition between using the display in a public, personal, or private mode.

2 Related Work

Research on MDEs or Smart Spaces dates back to the original work on Ubiquitous Computing at PARC, which supported an integrated user experience ranging from small-size “pads” to medium-size “tabs” to large-size wall-based displays [19]. Similarly, the iLand [18], the Interactive Space [10] and the Gaia smart space projects [8] have been researching the infrastructure and interaction technologies for seamless integration of mobile and fixed displays in MDEs. Technologies like PointRight [11], ARIS [3], and IMPROMPTU [4] enable users to bind devices together, redirect control and user interfaces, and to share and move information across multiple displays. These projects all hold two things in common. One is that the technology is deployed inside one room only; the other is that besides from simple knowledge about who is in the room, the technology has no knowledge of the activity of the users. Our work deliberately focus on dMDEs as well as creating support for handling devices, information, services, and users in relation to the activities that is taking place in the dMDE.

Explicit computational support for human activity is the main focus for activity-based computing approaches. For example, activity-based computing support for the personal computer [2] illustrates how the computer can model the human activity and use this for information and window management in the Windows XP operating system. Similarly, the Unified Activity Management project [14] uses activities for information management and collaboration inside large organizations. More automatic systems based on sensor inference have demonstrated how activities can be inferred from low-level user interaction with the computer. For example, the context-aware activity display (CAAD) [15] is able to detect, cluster, and highlight activities that the user might be engaged in. This is used for semi-automatic information management of documents, files, web pages, and emails associated with an activity.

None of these activity-based computing approaches address large interactive displays which can be used for *public* purposes in a dMDE, as CLINICAL SURFACES does. Furthermore, CLINICAL SURFACES is designed to highlight relevant activities in a specific contextual situation using context sensing of the physical world. This is different from the CAAD system, which is targeted towards activity recognition and highlighting based on low-level user input to a personal computer.

A particular set of semi-public display has been categorized as "awareness displays", i.e. displays that provide the user with peripheral information on the flow of work. Kimura [12] is an augmented office environment using interactive personal peripheral displays to manage "working contexts". The Notification Collage [7] is a system for users to communicate through a network of desk-mounted peripheral displays and public large displays. Because CLINICAL SURFACES constantly lists and highlights relevant activities for its current context, it could work as an awareness display. The aim of CLINICAL SURFACES is, however, to support a more active engagement with the displays. Dynamo [9] enables collaboration around a large interactive display, but relies on the users to provide data and organize applications and content at each new session. Thus, Dynamo does not relieve the user of the overhead that also exists in regular PC usage. CLINICAL SURFACES similarly supports collaboration in a public setting, but in addition it explicitly supports the activity of a user and thereby decrease the work involved in starting applications, finding relevant data, etc.

3 Background

This work is based on extensive observations of clinical work in large hospitals, and on the design of pervasive computing technology for hospitals [1]. In many ways, the work of hospital clinicians is very different from that of information workers – the latter having been the primary target for the design of interactive computer technology. First of all, work in hospitals is nomadic, i.e. clinicians constantly move around inside the hospital visiting different patients, departments, wards, conference rooms, and colleagues. In contrast to mobile work, nomadic work is characterized by the fact that the clinicians do not have any place to sit and work, i.e. most clinicians do not have a desk or an office of any kind. Second, clinical work is intensively collaborative; patient care happens in close collaboration between a team of highly specialized clinicians who constantly need to align, coordinate, and articulate their respective part of the overall activity of treating the patient. This collaboration typically pivots around shared medical information available in medical records, medical charts, radiology images, etc. Moreover, collaboration is often co-located, i.e. clinicians meet for planned or ad-hoc conferences and discuss the patient case. Examples of such conference situations range from the formal medical conference every morning, to improvised and ad-hoc conferences between clinicians in the hallway of the ward.

Analyzing the work artifacts used in hospitals, one finds that publicly available "displays" play a core role in the execution and coordination of work. Examples include large whiteboards listing operations, patients, beds, and other critical resources; medical records; notebooks; and small notes. Some of these artifacts are fixed to the wall while others are mobile. However, common to them all is their public nature, that they are used as a shared collaborative resource.

4 CLINICAL SURFACES

The design of CLINICAL SURFACES is based on the principles of (i) *activity-based computing*, use of (ii) *fixed and mobile displays*, (iii) *context-based adaptation*, and a separation between (iv) *public, personal, and private* use of these public displays.

Activity-based computing [2] includes support for activity centered computing, where resources and participants associated with a real-world human activity are modeled and aggregated in a corresponding computational activity. For example, the treatment of a patient may be modeled as a computational activity by aggregating all relevant medical information (e.g. medical records, radiology images, care plans, and medicine charts) as well as the participants involved in this activity. Each user participates in many activities, and the same activity may have several participants; the latter approach designed to support collaborative sharing of the resources associated with the activity. Moreover, activity-based computing includes support for activity roaming, i.e. support for moving a computational activity from one device to another; we say that an activity can be suspended on one machine and resumed on another. Activity-based computing is specifically targeted at supporting collaborative management of the myriad of heterogeneous data associated with a patient case, and to support the nomadic work inside e.g. hospitals.

The second principle is to support public displays. This is similar to the support for public wall-based displays in other MDEs, like iRoom or iLand. However, in addition to this previous work on supporting fixed wall-based displays in smart space technology, CLINICAL SURFACES is deliberately designed to also support *mobile* public displays. This resembles e.g. the use of the medical record in a hospital, which is a shared clinical artifact that is carried around by the clinicians.

Public displays constantly change usage context. Fixed public displays are frequently used by many different users for different activities, and mobile public displays are furthermore continuously changing their physical context. Hence, the design for public displays necessitates the support for these displays to adapt to their usage context. In CLINICAL SURFACES this is achieved by monitoring the context of the display, and continuously highlighting and providing access to the most relevant activities in the current context.

Finally, even though CLINICAL SURFACES is primarily designed to support public displays, it is also designed to support a seamless transition to more personal and private use of it. Hence, a user may choose to appropriate any public display and transform it into a personal device.

4.1 Interaction Design

Figure 1 shows the user interface of CLINICAL SURFACES consisting of four main parts: (A) the context bar showing this public display's current context; (B) the activity list showing an overview of relevant activities within this context; (C) a resumed activity providing an overview of its associated resources; and (D) two applications launched as part of the resumed activity. This interface is used on all public displays and is able to scale from portable tablet PC to large wall-based displays.

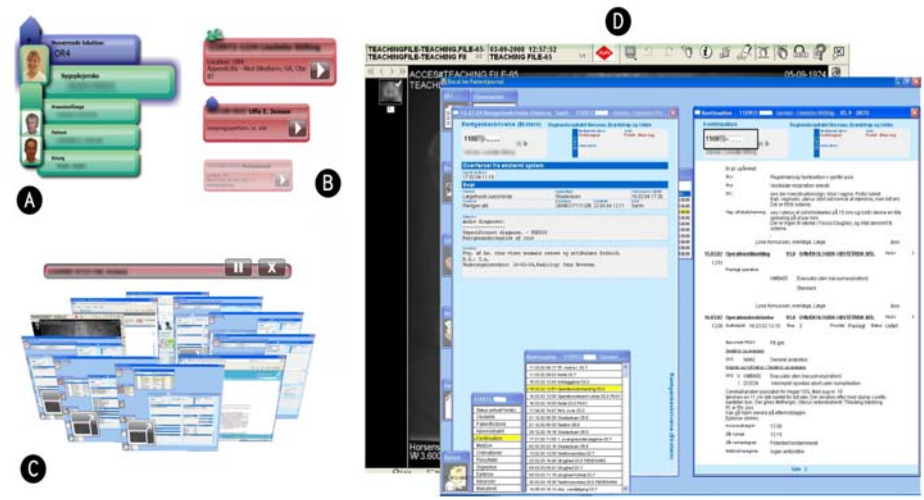


Fig. 1. The CLINICAL SURFACES user interface – (A) Context Bar (B) List of activities, (C) a resumed activity with associated resources displayed in a carrousel, and (D) two running applications

The context bar (A) shows the current context of this public display. This includes displaying the current location using a blue "house" icon, and nearby persons (including patients) using a green label and a picture. If the public display is fixed to e.g. a wall, the current location seldom changes. However, if this is a mobile public display, the current location is constantly tracked. Similarly, the display constantly tracks and displays a list of persons co-located with this public display.

The activity list (B) is constantly adjusted to show and highlight which activities are most relevant in this context. The current implementation operates with three levels of relevance shown as a large, small, and transparent red box. The icons attached to each activity reveal the type of relevancy: the icon with the two users illustrate that several persons co-located with the public display participate in this activity; the house icon illustrates that this activity is relevant for the current location of this public display.

An activity can be resumed by clicking the "Play" icon. Then CLINICAL SURFACES shows previews of its associated resources, as shown in Figure 1 (C). The user can now choose to either restore each resource individually by clicking the "Enlarge" icon on each preview, or s/he can restore the whole activity, i.e. all of the associated resources by clicking on the "Play" icon on the activity's title bar. Restoring a resource means that this resource is fetched and displayed using the local device's relevant application. Hence, a URL link opens in a web browser, a Word document in MS Word, etc. In Figure 1 (D), an electronic patient record and an X-ray image relevant to this patient are shown. The activity is suspended again by clicking the "Pause" icon, and the state of the activity is saved. This includes information about the placement and state of any resumed applications. Clicking on the "X" dismisses the "carousel" view. Each display can handle multiple resumed activities simultaneously, thereby enabling co-located use.

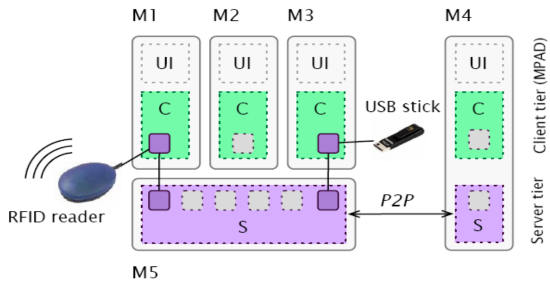


Fig. 2. The CLINICAL SURFACES architecture showing the (S)erver, (C)lient, and user interface (UI) components, as well as the extension points with e.g. RFID reader interface. This setup has five machines.

CLINICAL SURFACES distinguishes between three levels of privacy; public, personal, and private. *Public* is the default mode where the relevancy algorithm puts equal weight to the information presence of all close-by personnel. The *personal* mode is entered by double-clicking on a person in the context bar. This is, for example, used when a clinician picks up a mobile public display. In the personal mode, the display shows all the activities that this person participates in, including the ones not relevant for the current context. *Private* mode is activated when a USB stick containing private activities is inserted into a display. Then context information is ignored, and the user can work only on the private activities stored or referenced on the USB stick. Once the USB stick is removed, the display reverts back to operate in a public mode.

The focus of CLINICAL SURFACES is to support easy access to activities and associated resources and participants. Hence, focus has not been on activity management, i.e. the creation and maintenance of activities. This is done in a semi-automatic fashion using context-aware technologies [2], which is beyond the scope of this paper.

4.2 Technical Implementation

CLINICAL SURFACES is implemented on top of a distributed storage and event system called AEXO. As illustrated in Figure 2, the system contains two tiers; a server tier and a client tier; the latter containing the user-interface (UI) is shown in Figure 1.

The server tier maintains a hierarchical map data structure holding all persistent data in the system, including users, activities, (links to) resources, and location and context information. AEXO embeds an event system that enables clients to listen for changes in the map. The server tier can be distributed in a peer-to-peer fashion across several host machines; like between M4 and M5 in Figure 2. AEXO dynamically listens for and mounts data, which may reside on external data sources such as laptops, USB drives, and RFID readers. This data is incorporated into an existing map. This feature is used to implement the feature of having private activities on a USB stick.

Communication between the AEXO client and server tiers can use several protocols, but the AEXO C# client library layer wraps all aspects of the lower level communication details. The main responsibility of the AEXO client tier is to ensure that the CLINICAL SURFACES object model is synchronized with the hierarchical map in AEXO. This is done by adding C# annotations of relevant properties in the data model. The user interface of CLINICAL SURFACES is implemented in Windows Presentation Foundation (WPF) using data binding to the underlying data model store in AEXO. This three-layer model ensures that events from e.g. changes in the AEXO data map are propagated to the AEXO clients, which updates the client object model, which again propagates changes to the CLINICAL SURFACES WPF user interface.

5 Evaluation

CLINICAL SURFACES has been evaluated in a large hospital. It was used during a full day (8 hours) by clinical personnel from the orthopedic department, including surgeons, operating nurses, anesthesiologists, and ward nurses. In total 8 clinicians participated. Since patients were not intended users, no real patients were involved; two staff members of the hospital helped playing the patient role. The acting patients, were, however treated as any real patient, including being physically examined by the doctors and prepared for surgery.

5.1 Evaluation Setup

The deployment of CLINICAL SURFACES in the hospital is illustrated in Figure 3. The locations include the conference room (A) used by the surgeons for the morning conference; the patient ward (B) where admitted patients are hospitalized; the hallway

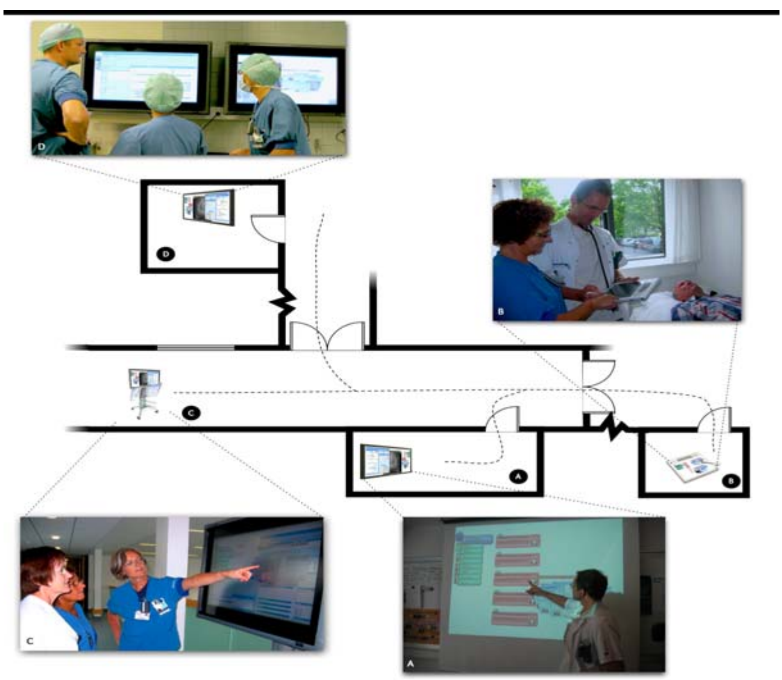


Fig. 3. The area of the hospital in which we evaluated the clinical prototype. (A) is the conference room, (B) the patient-ward, (C) the hallway, and (D) the operating room.

leading to the patient ward (C); and inside an operating room (D). Wall-based public displays using 42” touch screens were deployed in the conference room, the hallway, and inside the OR. A Lenovo tablet PC and a Motion Computing C5 Clinical tablet PC were used as mobile public displays. The mobile displays were used mainly at the patient ward. All displays (both mobile and fixed) and all clinical personnel were tracked using active RFID tags (WaveTrend LRX400), which had a reading range of ca. 10 meters.

In total 15 real patient cases were modeled in CLINICAL SURFACES as activities. All of these activities had associated medical data from the electronic patient record, the radiology system, and various web-based medical applications. Figure 1 shows an example of the user interface of CLINICAL SURFACES with this data. The creation and maintenance of the activities was performed in a Wizard-of-Oz manner by a human operator [13]. As the focus was only browsing the relevant patient data and not the usage of the applications, screen shots of the real applications were used.

5.2 Evaluation Methods

The evaluation was not intended to be a strict usability evaluation but was adopted to test the design of CLINICAL SURFACES in a real clinical environment. The aim was thus to assess the feasibility of a new innovative design, in which case applying traditional usability evaluation in a lab setup would be rather inappropriate [6].

The evaluation was centered around a real patient case in an anonymized form. The entire patient case was authentic, including all medical data and events. The main flow of the case was:

1. The patient is admitted to the hospital for acute appendicitis during the night
2. In the morning, surgery is planned during the medical conference
3. The patient is visited at the bed ward
4. An ad-hoc discussion of the case takes place in the hallway
5. Surgery is prepared in the operation room
6. Surgery is performed
7. The patient is taken back to the bed ward

The goal was to provide objective measurements on the usefulness and usability of our design while, at the same time, investigating the underlying detailed user reaction to the system and the user interface in a more qualitative fashion. For this purpose, we used a *multi-method evaluation* setup where we (i) asked the users to enact the patient case as realistically as possible while thinking aloud, (ii) investigated perceived usefulness and usability based on a questionnaire [5], and (iii) made a semi-structured follow-up interview.

6 Results

The results from the perceived usefulness/ease of use questionnaire are shown in Table 1. Overall, the system is perceived very useful (4.16, std. dev. of 0.65) and easy to use and learn (4.13, std. dev. 0.57). Looking at the specific aspects of the system, Table 1 also shows the perceived usefulness of the four main features of the system.

Table 1. Perceived usefulness summary (N=8, scale = 1–5)

Perceived usefulness	*Avg.*	*Std. dev.*
System usefulness	4.16	0.65
Ease of using and learning	4.13	0.57
Activity Management	4.25	0.52
Public Displays	4.34	0.69
Context-based adaptation	3.56	0.84
Public-Personal-Private	4.22	0.79

6.1 Activity Management

Support for activity management scores high – 4.25 (std. dev 0.52). Hence, the clinicians either agreed or strongly agreed that it would be useful to gather all relevant patient information in an activity, that an activity would help recover relevant patient data quickly, and that it would be easy to get access to patient data while moving around inside the hospital. As argued by a surgical nurse in the subsequent interview:

Such a system will definitely be a huge advantage. Everything would be collated in one place. Today, we spend a lot of time in getting access to the different systems, and to find the right documentation in them.

Similarly, the simple overview that CLINICAL SURFACES provides was highly appreciated by the clinicians as an important feature in a busy work environment:

What I like about the system is that we all use different aspect of the clinical data, and I can just select the parts that I need. I think the user interface is extremely simple.

During the discussion, the clinicians also suggested things for improvement. For example, having mechanisms for displaying activities and services in fixed locations on the screen might increase the overview and help the clinicians to quickly locate important information.

6.2 Public Displays

Using public displays in a clinical environment was considered very useful (4.34, std. dev 0.69). Hence, the clinicians agreed that fixed and mobile displays would be a good fit to the working conditions of a hospital; that it would be important to have the same interface on both the wall-based and the mobile displays; and that being able to just walk up to a public display and start using it would be a crucial feature. As expressed by a nurse:

Another good thing about this system is, that in a busy hospital [...] you do not need to spend time accessing data. You just go to the screen and get the data. [...] One can imagine that there are 10 of these mobile displays for each [patient] ward and you just take one from a re-charger station when you enter the ward.

The participants also discussed where to place wall-based and mobile displays. They concluded that the smaller mobile displays would suit the work in a patient ward, whereas it would be of little use inside the operating room. In the operating room, the large wall-based displays would be useful since they display more data and can be viewed from a distance. Furthermore, carrying mobile display in and out of operating room also implies a contamination hazard.

6.3 Context-Based Adaptation

The feature of context-based adaptation scores consistently low (3.56, std. dev 0.84). In particular, the clinicians did not agree that CLINICAL SURFACES was displaying the most relevant activities for a specific situation, and that using location tracking of personnel would help find the most relevant activities. Rather than co-location of people in front of the display, it seemed that the relevance of activities – or more specifically the patient cases – depended more on clinical issues like urgency of the case, or the order of the operations to be done in the operating room. As argued by a surgeon:

On the operation hallway – if we want to have information about a patient then we can go to the display and activate the name. But in the operation room, it is perhaps better to have the patients' [activities] fixed on the screen so that they will not change or move.

Another issue with the adaptation is that the content and appearance of the activity list would change based on people entering and leaving the room – as expressed by an operation nurse:

When, for example, [the surgeon] enters the operation room he can see his own activities and can see what he is going to do after this process. And it is really good for [him]. But the team in the operating room may be confused if his presence in the room starts to affect the tasks, which are linked to the people in this room.

6.4 Public-Personal-Private

Smooth transition between a public, personal, and private use of the displays also scores high on perceived usefulness (4.22, std. dev 0.65). The clinicians clearly agreed that it would be beneficial to see both shared as well as personal activities, and they also argued that in most cases, privacy is less of a problem, since ownership of patient related clinical data needs to be shared among all parties involved in the treating of a patient. As argued by a ward nurse:

...it is good that only you have access to the private data on the USB [stick] and not others, but as long as it is patient information and something which everybody has an interest in, it is nice to be able to transfer the screen to another [member of the staff].

The clinicians, however, pointed out that it was important to consider exactly where to put the public wall-based displays; they should be located in places where people, irrelevant for the patient case (e.g. other patients and relatives), would not accidentally view the data or overhear a conversation. In the evaluation setup, the display in the hallway was placed in a space too crowded and too public.

7 Discussion

In this section, we want to discuss lessons from the evaluation of CLINICAL SURFACES in relation to existing and future research.

7.1 Activity-Based dMDEs

CLINICAL SURFACES supports bundling of related services and data which is associated with some real-world human activity, like the patient cases used in the evaluation. The clinicians agreed that this support for activity-centered bundling of resources is very useful. Furthermore, the support for roaming the activities between public displays, as enabled by the AEXO platform, was deemed essential for work inside the hospital.

The distributed nature of our interactive space is very unlike most related research on MDEs. The physical disconnectedness of the devices and displays in a dMDE is a crucial point of difference. The structure of hospital work and the architecture of hospitals themselves do not easily lend themselves to the application of MDE as a technology confined in isolated rooms. Instead hospitals have vast corridors, offices, operating rooms and patient wards. The approach we have taken is to consider the entire hospital as an interactive space and deploy multiple displays in such an environment.

This has raised a number of new research challenges, which traditional MDEs do not address. Among these, we found two issues to be of central importance for a dMDE; how to enable the flow of activities and data between the public displays, and how to enable users to access and start using the displays in the dMDE. In CLINICAL SURFACES, the first challenge is addressed by using location information to automatically direct relevant activities onto a wall-based public display. Hence, activities – and their associated resources – were simply just following the users around inside the dMDE. Based on the feedback from the clinicians, this approach seemed to work well.

Furthermore, by using mobile devices, activities could be "carried" wherever they were needed, and this fact was used in the case of the ward-rounds. In contrast to traditional use of tablet PCs, CLINICAL SURFACES also turned this device into a public display participating in the dMDE. In several cases, the tablet PCs were handed over from one clinician to another (as shown in Figure 3B). In these 'hand-over' situations, the display simply switched context and was carried for use elsewhere in the dMDE.

Finally, to support accessing and carrying private activities – with associated data – the USB stick served as a "pocket-book"; i.e. a private store for activities which could be inserted in all fixed and mobile screens thereby converting them to personal devices. This kind of local adaptation of specific displays in the dMDE is another important difference from a traditional MDE.

7.2 Public Displays and Privacy

There seems to be an inherent conflict between using public displays and maintaining privacy. The latter is especially important in a clinical domain where sensitive patient data is displayed and discussed. As outlined in previous section, this issue was also discussed by the clinicians while using CLINICAL SURFACES in the hallway. It was a deliberate choice to put the display in this rather public location, hoping to spark a discussion of the feasibility of having such public displays around the hospital. The discussion, however, was concerned with several aspects of privacy. The first aspect was whether it is appropriate that one participant can see and access the activities of another participant present in front of the display. For example, in Figure 1 the surgeon has access to the nurse's activities. After some discussion, the clinicians concluded that this was actually not a privacy issue. Partly because all patients hospitalized at the hospital in principle should be accessible by all clinicians, and partly because the activities were only accessible if the user was in physical proximity of the display. Once the user moves away, his or her activities are no longer present in the display. Even though the reading range of the active RFID tags seemed to be too long in the deployment setup, the clinicians agreed that this design was suitable for upholding of privacy in hospital-specific activities. By using the USB stick – either as a key or as a holder of data – the users were able to more precisely control when private activities were displayed.

Another aspect of privacy was concerned with the physical deployment of such large wall-based displays. The deployment location of the hallway display was not appropriate in the opinion of the users; it was simply put in a location where too many people were passing by, including relatives and other patients. Hence, it would not be feasible to discuss a patient case in this location. The clinicians concluded that such displays would be very useful if put in the hallway of the operation ward or the patient ward. Inside the operation room, there were no privacy issues.

In summary, the inherent conflict between using public displays and the concern for privacy can be mitigated through the proper modeling of public, personal, and private activities in CLINICAL SURFACES and its infrastructure, combined with an appropriate deployment of the displays and tracking technology.

7.3 Context-Based Relevance

In the current implementation, context-based adaptation is linked to the physical co-presence of participants of an activity; the more participants present in the same location as the public display, the more relevant the activity becomes. For the clinicians, this model was too simple – physical co-location could be a relevance-indicator, but a series of other more clinically oriented issues seems to be more important to incorporate. For example, the operation schedule of the day, the urgency of the patient case, and whether the case needed to be discussed on the morning conference. The current relevance score based on co-location also had an annoying side-effect, because when an RFID tag became invisible due to other reasons than the person leaving the room (e.g. being covered by something), then the relevance of the activities would immediately change, which would resize and re-organize the list of activities. This caused a lot of confusion during the evaluation. Since the technical implementation of CLINICAL SURFACES allows for replacement of the relevance algorithm, a natural next step would be to investigate how to improve this based on e.g. more clinical context information.

8 Conclusion

In this paper, we have presented a novel approach and user interface technology for supporting activity-based computing in distributed multi-display environments named CLINICAL SURFACES. Taking an outset in the nomadic and collaborative work inside hospitals, CLINICAL SURFACES was designed to enable easy access to medical data across physically distributed public displays. CLINICAL SURFACES was evaluated in a large hospital and the evaluation revealed that the clinicians found the system very useful and easy to use. The support for activity management, distribution of clinical data, and the ability to quickly access all relevant data on any public display were considered as strong features. The evaluation, however, also pointed to areas for further research and improvements, including improving the current implementation of context-based adaptation.

References

1. Bardram, J.E., Christensen, H.B.: Pervasive Computing Support for Hospitals: An Overview of the Activity-Based Computing Project. In: IEEE Pervasive Computing, pp. 44–51 (2007)
2. Bardram, J.E., Bunde-Pedersen, J., Soegaard, M.: Support for activity-based computing in a personal computing operating system. In: Proceedings of CHI 2006, pp. 211–220 (2006)
3. Biehl, J.T., Bailey, B.P.: Aris: An interface for application relocation in an interactive space. In: Proceedings of Graphics Interface 2004, pp. 107–116 (2004)
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: Proceedings of CHI 2008, pp. 939–948 (2008)
5. Davis: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13(3), 318–340 (1989)
6. Greenberg, S., Buxton, B.: Usability evaluation considered harmful (some of the time). In: Proceeding of CHI 2008, New York, NY, USA, pp. 111–120 (2008)
7. Greenberg, S., Rounding, M.: The notification collage: posting information to public and personal displays. In: Proceedings of CHI 2001, USA, pp. 514–521 (2001)
8. Hess, C., Romn, M., Campbell, R.: Building Applications for Ubiquitous Computing Environments. In: Proceedings of Pervasive 2002, pp. 16–29. Springer, Heidelberg (2002)
9. Izadi, S., Brignull, H., Rodden, T., Rogers, Y., Underwood, M.: Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media. In: Proceedings of UIST 2003, pp. 159–168 (2003)
10. Johanson, B., Fox, A., Winograd, T.: The interactive workspaces project: Experiences with ubiquitous computing rooms. IEEE Pervasive Computing 1(2), 67–74 (2002)
11. Johanson, B., Hutchins, G., Winograd, T., Stone, M.: PointRight: experience with flexible input redirection in interactive workspaces. In: Proceedings of UIST 2002, pp. 227–234 (2002)
12. MacIntyre, B., Mynatt, E.D., Vodia, S., Hansen, K.M., Tullio, J., Corso, G.M.: Support for Multitasking and Background Awareness Using Interactive Peripheral Displays. In: Proceeding of UIST 2001, pp. 11–14 (2001)
13. Molin, L.: Wizard-of-oz prototyping for co-operative interaction design of graphical user interfaces. In: Proceedings of NordiCHI 2004, pp. 425–428 (2004)

14. Moody, P., et al.: Business activity patterns: A new model for collaborative business applications. IBM Systems Journal 45(4), 683–694 (2006)
15. Rattenbury, T., Canny, J.: CAAD: An Automatic Task Support System. In: Proceedings of CHI 2007, pp. 687–696 (2007)
16. Rekimoto, J., Saitoh, M.: Augmented surfaces: a spatially continuous work space for hybrid computing environments. In: Proceedings of CHI 1999, pp. 378–385 (1999)
17. Stefik, M., et al.: Beyond the chalkboard: computer support for collaboration and problem solving in meetings. Commun. ACM 30(1), 32–47 (1987)
18. Streitz, N.A., et al.: i-land: an interactive landscape for creativity and innovation. In: Proceedings of CHI 1999, pp. 120–127 (1999)
19. Weiser, M.: The Computer for the 21st Century. Scientific American 265(3), 66–75 (1991)

Designing a Safer Interactive Healthcare System - The Impact of Authentic User Participation

Kathryn L. Went, Peter Gregor, and Ian W. Ricketts

School of Computing, University of Dundee, Dundee,
DD1 4HN, UK
{kwent,pgregor,ricketts}@computing.dundee.ac.uk

Abstract. Information technology has been widely promoted in the healthcare sector to improve current practice and patient safety. However, end users are seldom involved extensively in the design and development of healthcare systems, with lip service often paid to the idea of true user involvement. In this case study the impact of sustained authentic user participation was explored using an interdisciplinary team, consisting of experts both in interaction and healthcare design and consultant anaesthetists, nurses, and pharmacists, to create an electronic prescribing and administration system. This paper details the interface that was created and provides examples of the way in which the design evolved in response to the sustained authentic user participation methods. The working prototype both reduced the opportunity for user error and was preferred by its users to the existing manual system.

Keywords: Healthcare, user engagement, participatory design.

1 Introduction

The appropriate use of information technology has been promoted in the healthcare sector, both in the UK and abroad, to improve current practice and patient safety. The UK National Programme for IT in the NHS has been widely criticised for running behind schedule and producing solutions that have not always been deemed as satisfactory [1]. Failure to engage with clinicians during the initial stages of the design has been cited as a reason.

This study was based in a hospital setting and created a prototype to allow optimum engagement of users from multiple disciplines in the design and development of an interactive healthcare system. The aim was to design an interface to actively reduce errors, be highly usable and not negatively affect patient care.

2 Background

2.1 The Incidence of Hospital Medication Errors

2.1.1 Medication Errors

The Department of Health [2] estimates that one in 10 patients admitted to NHS hospitals will be unintentionally harmed, with medication errors being the second highest reported cause of harm.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 718–730, 2009.

Medication errors are a global problem [3,4]. It is estimated that between 44,000 - 98,000 deaths per year are attributable to medication errors in the United States of America [5] with an unknown number in the UK, due to low reporting rates [2].

Medication errors adversely affect the patient, their family and friends, and the professionals concerned. As well as the human cost, they also impose a financial burden on the NHS. The Department of Health estimated that the direct cost of medication errors in 2004 was £200-£400 million per year [6].

2.1.2 Prescribing Errors

The growth in the number of available medicines over recent years makes it increasingly difficult for prescribers to be alert to all potential medication interactions. Consequently prescribing has become very complex. When looking at the types of medication error it is believed that prescribing errors are not only the most likely to cause an adverse drug event but are also on the whole avoidable [7].

Common factors associated with prescribing errors have been reported [4] as:

- Decline in renal or hepatic function requiring alteration of drug therapy (13.9%)
- Patient history of allergy to the same medication class (12.1%)
- Using the wrong drug name, dosage form or abbreviation (11.4%)
- Incorrect dosage calculation (11.1%)
- Unusual and critical dosage frequency considerations (10.8%).

Approximately 200 million prescriptions are generated in NHS hospitals each year [6]. Almost all are handwritten. A study assessing the quality of handwritten prescriptions found that out of 4,536 inpatient prescriptions, 4-10% were illegible or ambiguous [8], which can result in wrong drugs or dosages being given. The positioning of zeros and decimal points also give rise to errors. Trailing zeros are sometimes used, for example 1.0mg, intended to mean 1mg, may be misread as 10mg if the decimal point is not seen, resulting in a ten fold overdose. Handwritten prescription sheets are also subject to transcription errors [7]. These occur when a prescription sheet becomes full and current prescriptions are rewritten onto another sheet, increasing the opportunity for human error.

Prescribing and recording the administration of medicines in intensive care is a complex process, involving a number of disciplines working in a stressful clinical environment. Patients in an intensive care unit are frequently prescribed multiple medicines, often in complicated combinations, therefore increasing the possibility of interactions. This complexity increases the risk of medication errors, which may have catastrophic consequences [6]. This would suggest the Intensive Care Unit (ICU) is a good setting [9] for studying the effectiveness of a suitably designed electronic prescribing system.

2.2 The Manual Prescribing System

The ICU in Ninewells Hospital, Dundee uses a manual prescribing system comprising

i) A paper medication chart (Kardex®) to record the patients' prescriptions (figure 1)

ii) A separate card (the Medicine Recording sheet) to record drug administration times.

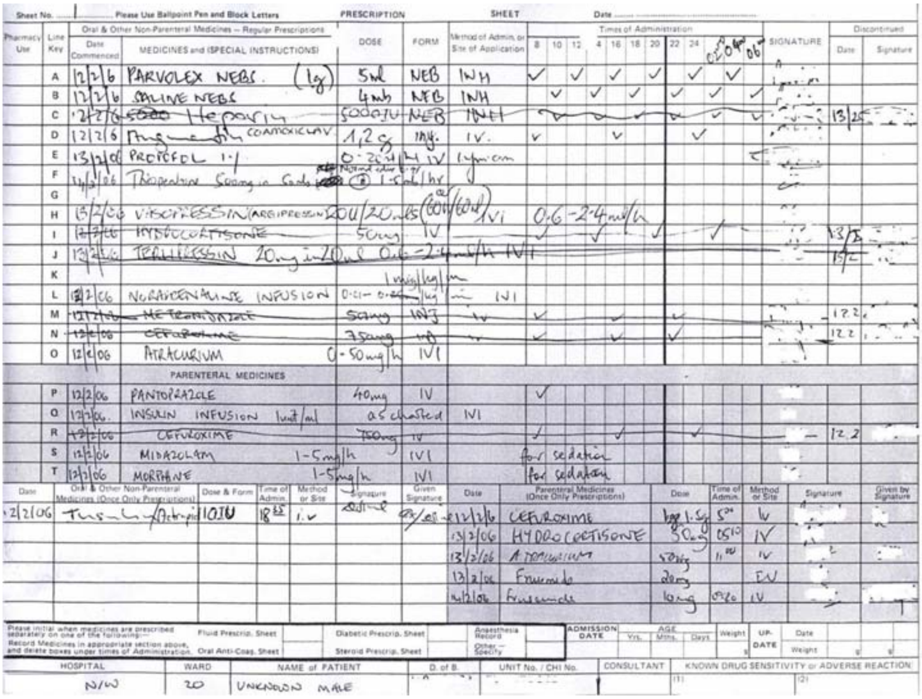


Fig. 1. Patient's Medication Chart from the ICU

2.2.1 Paper Medication Chart

The paper medication chart records the start date, name of the drug, dose, route, formulation, time/s of administration and the signature of the prescriber. The details of discontinued prescriptions are also recorded on the chart.

The prescriptions on the chart are separated into two sections. The top section details the regular prescriptions that are required at specified times until discontinued. The lower section details prescriptions that are given only once at the time and date specified.

The chart also contains the patient's personal details recording their name, date of birth, a unique patient identifier used in Scotland (known as the Community Health Index or CHI), consultant, ward, hospital, weight and any allergies.

2.2.2 Medicine Recording Sheet

A Medicine Recording sheet is used by the nursing staff to document the times that medications are administered and reasons for non or late administration e.g. "once only medication given" or "dose withheld on prescriber's instruction".

2.3 Problems Associated with the Manual Prescribing System

In addition to the problems associated with illegible handwriting, the following system failures are also associated with the paper system:

- The medication chart and medicine recording sheet fail to provide any backup record.
- If there is damage or spoilage to the chart, it can be difficult to ascertain what medicines a patient has been prescribed or administered, particularly if documentation needs to be reviewed at a later date, for example during the investigation of a complaint or incident.
- There is no mechanism to ensure all appropriate details are provided; for example, a prescription may be discontinued but not clearly crossed out and mistakenly interpreted as a prescription still to be administered.
- Problems can arise when the dosage unit “micrograms” is not written in full, i.e. written as “µg”, which can be misinterpreted as “mg” (milligrams), increasing the likelihood of an overdose being administered.
- If the generic name is not recorded, this can result in the same prescription being entered twice i.e. once with the generic name and a second time using a brand name.
- When the medication chart becomes full, the current prescriptions (i.e. those which have not been discontinued) are transcribed to a new card. There is further potential for errors at this point, in that the information may be transcribed incorrectly as a result of illegible handwriting, human error or incomplete prescriptions.

An electronic system could potentially reduce or eliminate the problems that are prevalent with the paper system and improve patient safety. However, studies in the UK and abroad have identified that such systems can introduce new types of errors [10-15]. Koppel et al [13] discovered 22 types of medication error risks that were facilitated by an electronic prescribing system. Causes of these errors were found to be “human-machine interface flaws” and “information errors”. Not only were errors prevalent but they also occurred frequently: this study looked specifically at reducing the opportunity for error by following interaction design practice combined with early and continual authentic user participation.

3 Designing the System with Authentic User Participation

The paper medication chart is normally kept in a folder at the end of the patient’s bed. The chart is accessed by nurses, doctors and pharmacists and can be viewed by multiple users simultaneously. The electronic system needed to replicate this and also be portable to prevent restricting the normal interaction process that was observed with the paper system e.g. nurses taking the paper chart to the doctor if they had a query about a prescription or passing the chart between the team in the morning ward round. This led to the decision to use a tablet pc.

The system was designed as a client server application, allowing several client machines to be connected to the server, providing each patient with a tablet machine, with a tablet pen for input and an incorporated barcode scanner to allow users to log on to the system. The tablet pc could be mounted at the end of the patients’ bed to provide easy viewing and be easily removed from the mount. It would be roughly the same size as the paper chart. A rugged tablet pc was required to endure the Intensive Care environment, sustaining extensive cleaning and being dropped. A tablet pen was used for input instead of a keyboard for portability and to reduce the risk of infection.

3.1 Authentic User Participation

Involving users as a central part of the development process can lead to more effective, efficient and safer systems [16]. A user-involved approach was adopted for the design, development and implementation of the electronic prescribing and administration system.

At the beginning of the project an interdisciplinary team was formed comprising two consultants in anaesthesia and intensive care, the principal clinical pharmacist for critical care, the intensive care specialist liaison nurse, a nursing education specialist, a professor of interactive systems design, a professor of assistive systems and healthcare computing and a PhD computing student. The established team was authentically involved throughout the design, development and implementation of the system and met 48 times over a three-year period. The frequency of the meetings ranged between fortnightly and monthly.

A participatory design approach was central to the process. Evolutionary prototyping, a participatory design method where the "system is delivered to the end users as a series of iterative prototypes, each of which gradually adds functionality" [17], was employed. The design began as a set of paper-based prototypes that quickly enabled a visualisation of the system providing a basis for agreeing the requirements. These paper prototypes led to the development of an evolutionary software prototype.

To sustain engagement and enthusiasm, regular meetings were held at which the evolving prototype was evaluated by the interdisciplinary team. The typical structure of these meetings was a walk-through of the prototype, with particular emphasis on the amendments made since the last meeting, followed by a managed discussion with the team that identified improvements and agreed recommendations for changes to the interface, to support best practice. Adopting this process authentically ensured that the evolving system focused on the clinicians' requirements.

3.2 Interface Design

The interface of the system was designed by taking cognisance of best practice in interaction design in combination with authentic user participation.

The final system (figure 2) was designed to be familiar and emulate the original paper medication chart with the additional functionality of confirming that medication had been administered. The main functionality of the system was that it allowed

- Doctors to create and discontinue prescriptions
- Nurses to record the administration of medication
- Pharmacists to record that they had checked a prescription and add additional notes to the prescription
- New patients to be added to the system
- Authorised users to amend patient details

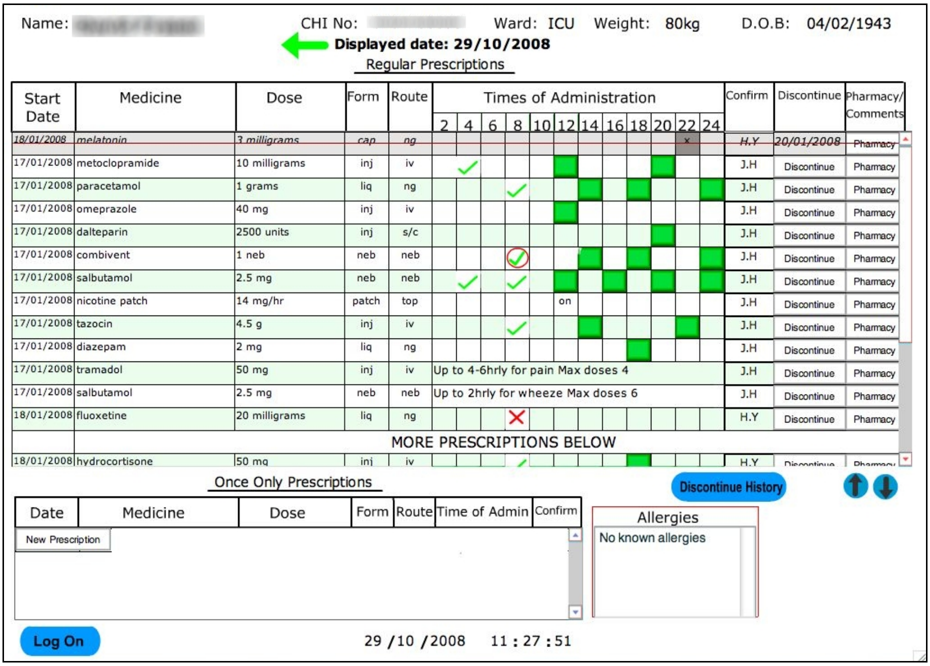


Fig. 2. The Final System

The interface was organized into two sections, similar to the paper system, with the top section displaying regular prescriptions and the bottom section displaying the "once only" prescriptions. Within the regular prescription section three subsections displayed discontinued, current and future prescriptions.

Within the "regular prescriptions" continuous infusions were grouped together as a consequence of nursing staff feedback, to make the administration process easier and to reflect the sequence for administering medication. The interface was designed to minimise free text input, to avoid frustration and save time entering text via an on-screen keyboard.

3.3 Decision Support

Some commonly prescribed medications such as insulin and gentamicin have associated algorithms to determine the amount and frequency of doses given to the patient, based for example on patient weight. Patients in intensive care often behave as if they are diabetic. It has been found that a lower target range than normally aimed for of blood glucose in these patients can improve their outcome [18,19].

It was initially decided that because insulin was frequently prescribed in ICU, it would be beneficial to incorporate an insulin-prescribing algorithm [20] into the system.

The clinicians sourced and evaluated an algorithm, which the team further refined during the course of eleven meetings, so that it could be successfully incorporated into the evolutionary prototype.

There is also a high prevalence of renal and hepatic impairments in ICU patients, requiring the usual medication dose to be altered. When creating a new prescription, the team decided to display relevant dose recommendations, as specified by the British National Formulary (BNF), after a medication had been selected.

Both the insulin prescribing algorithm and the medication dose recommendations are examples of decision support. During the course of numerous meetings, the advantages and disadvantages of incorporating decision support were debated. It was highlighted that the dose recommendations included in the system conflicted with other guidelines in use in the ICU. After careful consideration by the team, it was decided that implementing decision support, in a piecemeal fashion, could introduce errors and cause annoyance to users by providing inconsistent information. Furthermore, it could introduce an additional variable, which could impact on the comparison of the paper and electronic systems. Therefore, it was proposed to concentrate on creating an interface that minimised prescribing errors and to evaluate its effectiveness before implementing decision support. The team agreed to remove the insulin prescribing algorithm and any medication impairment recommendations.

3.4 Wish List

Throughout the process, requirements were identified that the team believed would be useful (e.g. incorporating the 24-hour chart functionality) but would not directly affect the level of prescribing errors or the usability. As the purpose was to design a usable system to reduce prescribing errors, it was decided by the team that these requirements could be implemented at a later date and would be part of an "ideal" system. A wish list was suggested by one of the clinicians to prevent these ideas from being forgotten and to also lessen the risk of demotivation, which could lead to a reduction in contributions from group members.

3.5 New Prescription Design

Following one of the evaluation sessions, it became evident that multiple selections were needed to create a new prescription and that this could be avoided by using a software "wizard" for the task. In addition to reducing the number of screen selections, this would ensure that the users would adopt the correct order for creating a new prescription, by breaking down the process into the logical order for construction of a prescription.

Creating a prescription with the original prototype involved selecting the *start date*, *medicine*, *times of administration*, and *confirm* and selecting options from the route and form drop down menus, all of which were located underneath the last completed prescription line as shown in figure 3.

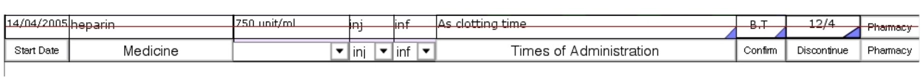


Fig. 3. Initial "New Prescription" Design

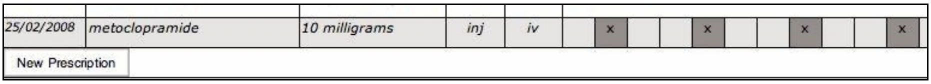


Fig. 4. Revised "New Prescription" Design

The revised "new prescription" design allowed the user to initially select the *new prescription* (figure 4). The location of the new prescription was retained as the eye was naturally drawn to the point under the last completed prescription.

After selecting the *new prescription*, the prescription "wizard" was displayed in the middle of the screen to attract the users attention.

- Firstly a calendar was displayed prompting the user to enter a start date.
- Next, the screen refreshed to prompt the user to select a medicine. There are over 100 commonly prescribed medicines in ICU, which led to a discussion on how best to display the different medicines. This resulted in the decision to use a filing card layout to allow the user to easily select the required medicine. They would select which letter range the medicine was in and then select the appropriate medicine from a list of options. This screen also displayed the start date selected and a back arrow providing the opportunity to return to the previous screen.
- To avoid duplicate prescriptions, if the medicine was currently prescribed, a warning was displayed to the prescriber presenting two options to either continue creating the prescription or cancel creating the new prescription.
- If they selected "continue" or if there were no duplicate prescriptions, a screen prompting the user to select the appropriate route of the medicine was presented, followed by a screen to prompt selection of the appropriate formulation of the medicine.

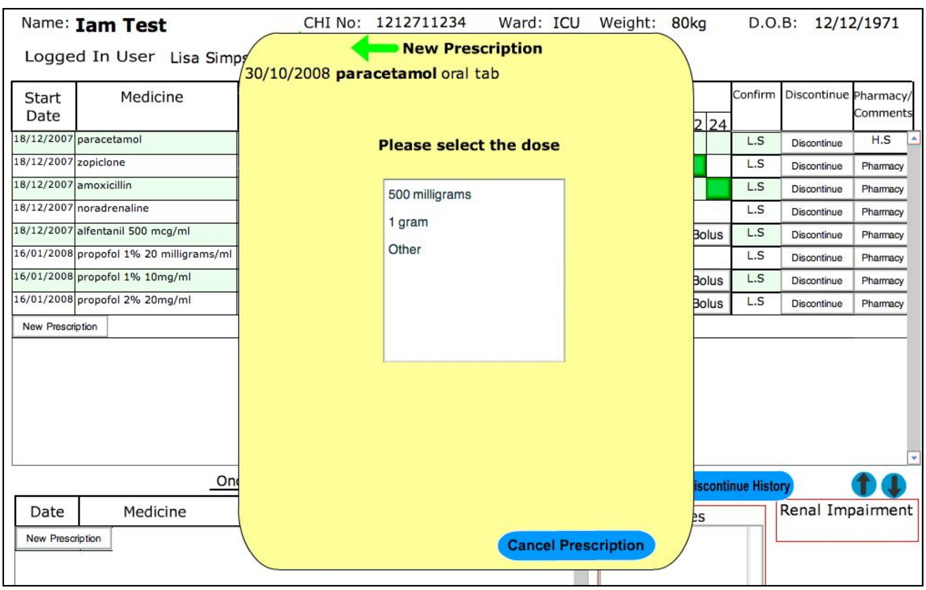


Fig. 5. New Prescription "Wizard"

– Based on these selections, the next screen displayed the appropriate dose in addition to the selected start date, medicine, route and form (figure 5).
– This was then followed with a screen prompting the user to select the times of administration and then a screen to confirm the prescription.

3.6 Minimising Error

The prescription wizard further reduced opportunities for errors. From a survey of prescription errors carried out in the ICU [21], the omission of important information occurred in over a quarter of the non-compliant prescriptions e.g. missing dose. The prescription wizard was developed to not only lessen the number of screen selections but to minimise the risk of important omissions.

In addition, the wizard was also designed to reduce the likelihood of entering inappropriate or incomplete information by providing suitable formulations, route and dose defaults, appropriate to the selected medicine. Once the medication was selected, the screen displaying the appropriate route options would be populated based on the medication. Subsequent to the route selection, the screen would then refresh to display the available formulations resulting from previous selections of medication and route. Underneath any default lists an "other" option was always provided to maintain flexibility and allow the user to prescribe a non-routine selection. This error prevention mechanism of displaying dynamic default options not only reduced the number of physical actions necessary to input the required prescription information but also assisted the prescriber by recognition.

However, although displaying a list of default options can reduce the likelihood of users entering inappropriate and incomplete information it can introduce human error via "slips" [22] if an unintended option is selected. For example Shulman et al [14] reported that one of the major errors caused by an electronic prescribing system occurred when the wrong dose was selected from a pull down menu and could have resulted in a 70 times overdose To minimise this, when an option was selected via the "wizard" it was again displayed in the next window with an option to return to the previous screen if an amendment is required. Once all the required prescription details had been selected, the created prescription was then displayed in a different format for the user to check and confirm, before the prescription was finalised. This double check mechanism was incorporated into the wizard to reduce the likelihood of erroneous selections, which could cause serious patient harm.

4 Evaluation Results

Expert analysis and user participation methods [23] were used to evaluate the evolutionary prototype electronic prescribing and administration system. The expert analysis method comprised three heuristic evaluations conducted at regular intervals throughout the development of the prototype to evaluate whether the system adhered to accepted usability principles [24]. In addition to the evaluation sessions held with the clinicians in the interdisciplinary team, end users outwith the team evaluated the system in both a laboratory and clinical setting where observational methods (cooperative evaluation) and query techniques (user interviews) were used respectively.

Subsequent to the team decision that the prototype of the electronic system appeared to meet the necessary requirements of the different user groups, two evaluation sessions with 30 clinicians were held; one with prescribers and one with administrators.

Following 59 evaluation sessions (including the 48 held with the interdisciplinary team) the prototype electronic prescribing and administration system was evaluated using live prescription information in the ICU. The prototype was used with 16 patients over a 5-month period (14th January – 5th June 2008). Paper and electronic prescription charts were run in parallel within ICU.

4.1 Levels of Prescribing Error

On completion of the 5-month trial both the electronic and paper medication charts were surveyed for levels of prescribing errors. The results of the study demonstrated that the overall level of prescription compliance was significantly higher with the electronic system (91.67%) compared to the paper system (46.73%). Electronically generated prescriptions were unambiguous, legible and resulted in a significant reduction in identified prescribing deviations.

However, there was a decrease in the level of compliance for infusions created on the electronic system, where both the rate and concentration of the medication was not always stated. This was due to a flaw in the new infusion prescription design and the integrity of data used to populate the drug database. Evaluating the system in parallel with the paper chart allowed this flaw to be identified without any risk to the level of patient care.

4.2 Interviews

From a total population of 62 clinicians who used the system, a sample was interviewed, comprising 9 doctors, 10 nurses and a clinical pharmacist. Nurses were the most frequent users of the system. Interviews took place between 28th April – 23rd July 2008. Participants had different levels of use of the system and only one participant had previous experience of electronic prescribing.

Overall the electronic system was preferred (60% of participants, 6 doctors, 6 nurses). Examples of the comments from the staff are shown in box 1.

The users' experience was on the whole positive. Generally they found the system intuitive, easy to learn and they were comfortable using it (box 2).

None of the participants interviewed expressed discomfort with the system. Five of the participants expressed initial reservations, with three of the five expressing a change of impression to a positive reaction after using the system.
"*Anything that's new you think it's going to go wrong and because it is prescribing as well, prescribing is important to be right… I suppose there was a concern that it might cause a problem with prescribing*" When asked if this impression changed over time "*Yes, I like your system*!" (Specialist registrar)

The other 2 participants changed to expressing an acceptance for the system but did not express a liking.

Eleven of the twenty participants said they believed the electronic prescribing and administration system had a positive impact on patient safety as it "*reduce(d) errors*" and was *"a safer system to use"*. Although the remaining participants did not explicitly say that the electronic system had a positive impact they implied that it would improve patient safety and would not have a negative effect.

Box 1. Comments from the staff about the electronic system

Examples of preference for using the system

"*I found it easier to go through the electronic one and see what had been given throughout the day instead of flicking through pages and pages of paper. It's a lot clearer and easier to decipher the information*" (Staff nurse)

"*I liked it, it's good, it's concise, it's a joy to be able to read a prescription clearly, understand what it says, the units it's supposed to be in and who has prescribed it. Yeah I really liked it*" (Staff nurse)

"*I probably prefer for ease of use the electronic one*" (Specialist registrar)

"One thing is it's definitely safe, … Patient safety - I think it's immensely improved and the other thing is it's very legible... and everybody can understand what's happening very clearly" (Specialist registrar)

Example of hesitance towards the electronic system

"I think probably I prefer the paper because of practice and what I'm used to, but I think if we didn't have to do the paper as well I wouldn't have a problem with doing the electronic" (Staff nurse)

"Writing does seem quicker but I can see that there is probably less room for errors if you do it electronically. They were quite similar really you put the same details in (for both systems). (I preferred using the) Paper, I think it's just getting used to something … I only had a short period on it" (Specialist registrar)

Box 2. Examples of Users' Experience of the Interaction

"*Ideas were really good, the format of the (electronic) Kardex® was good and after a small amount of training you could use the basics. I thought it was quite easy really, to use*" *(*Charge nurse)

"*I would say very comfortable it didn't take too long to get the hang of it… It was good, it was simple to use, it was a very intuitive set of interfaces, it was very clear*"(Specialist registrar)

"It was quite straightforward actually. Once I got over the shock of an electronic system …. it's actually very straightforward" (Staff nurse)

"Very user friendly…(The layout) mirrored a normal prescription so that's good. The thing that I like about it were that it keeps your allergies, renal impairments, pregnant, hepatic impairment to one bit… I think it worked well" (Specialist registrar)

5 Discussion

The study clearly demonstrates that the participatory approach produced a prototype which reduced prescribing errors by almost doubling levels of prescription compliance and importantly, was accepted by the end users: the doctors, pharmacists and nurses.

This paper illustrates the benefit of sustained user engagement in the design process. The inclusion of the users throughout the design and development process generated an evolving understanding of the process and encouraged ownership of the project.

Although the design decisions made were not necessarily novel, the process taken to reach these decisions highlights the benefits of ongoing participatory user involvement. Sustained engagement allowed the members of an interdisciplinary team to share their ideas, work through a number of options and refinements and delivered a solution that was designed both with and for the users, to meet their various needs. The participants understanding of the process was evidenced by the agreement of the group to create a wish list and remove decision support.

Authentic user participation shaped a system preferred overall by the clinicians to complete their everyday tasks and demonstrated a significant reduction in prescribing errors.

Acknowledgments. We wish to express our gratitude to the clinicians on the development team from the ICU, Ninewells Hospital, Dundee, who made this study possible by giving so generously of both their expertise and their time: P Antoniewicz, DA Corner, S Dailly, J Joss, F McIntyre, S McLeod and AJ Shearer. We also want to extend our thanks to all the clinicians who evaluated the prototypes.

References

1. Cross, M.: Benefits of £12bn IT Programme in NHS are "unclear," MPs Say. BMJ 334, 815 (2007)
2. Department of Health.: A Safer Place for Patients: Learning to Improve Patient Safety. The Stationery Office, London (2005)
3. Bates, D.W., Cullen, D.J., Laird, N., et al.: Incidence of Adverse Drug Events and Potential Adverse Drug Events: Implications for Prevention. JAMA 275, 29–34 (1995)
4. Lesar, T.S., Briceland, L., Stein, D.S.: Factors Related to Errors in Medication Prescribing. JAMA 277, 312–317 (1997)
5. Institute of Medicine: To Err is Human: Building a Safer Health System. National Academic Press, Washington DC (2000)
6. Department of Health: Building a Safer NHS for Patients: Improving Medication Safety. The Stationery Office, London (2004)
7. Wheeler, S.J., Wheeler, D.W.: Medication Errors in Anaesthesia. Anaesthesia 60, 257–273 (2005)
8. Jenkins, D., Cairns, C., Barber, N.: The Quality of Written Inpatient Prescriptions. Int. J. Pharm. Pract. 2, 176–179 (1993)
9. Barber, N.: Designing Information Technology to Support Prescribing Decision Making. Qual. Saf. Health Care 13, 450–454 (2004)

10. Ash, J.S., Berg, M., Coiera, E.: Some Unintended Consequences of Information Technology in Health Care: The Nature of Patient Care Information System-related Errors. J. Am. Inform. Assoc. 11, 104–112 (2004)
11. Campbell, E.M., Sittig, D.F., Ash, J.A., et al.: Types of Unintended Consequences Related to Computerized Provider Order Entry. J. Am. Inform. Assoc. 13, 547–556 (2006)
12. Donyai, P., O'Grady, K., Jacklin, A., et al.: The Effects of Electronic Prescribing on the Quality of Prescribing. Br. J. Clin. Pharmacol. 65, 230–237 (2008)
13. Koppel, R., Metlay, J.P., Cohen, A., et al.: Role of Computerized Physician Order Entry Systems in Facilitating Medication Errors. JAMA 293, 1197–1203 (2005)
14. Shulman, R., Singer, M., Goldstone, J., et al.: Medication Errors: A Prospective Cohort Study of Hand-written and Computerised Physician Order Entry in the Intensive Care Unit. Crit Care 9, 516–521 (2005)
15. Walsh, K.E., Adams, W.G., Bauchner, H., et al.: Medication Errors Related to Computerized Order Entry for Children. Pediatrics 118, 1872–1879 (2006)
16. Preece, J., Rogers, Y., Sharp, H.: Interaction Design: Beyond Human-Computer Interaction, 2nd edn. John Wiley & Sons, New York (2007)
17. Muller, M.J.: Participatory Design: The Third Space in HCI. In: Jacko, J.A., Sears, A. (eds.) The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications. Human Factors and Ergonomics, pp. 1051–1068. Lawrence Erlbaum Associates, Hillsdale (2003)
18. Finney, S.J., Zekveld, C., Elia, A., et al.: Glucose Control and Mortality in Critically Ill Patients. JAMA 290, 2041–2047 (2003)
19. Van den Berghe, G., Wouters, P., Weekers, F., et al.: Intensive Insulin Therapy in Critically Ill Patients. N. Engl. J. Med. 345, 1359–1367 (2001)
20. Thomas, A.N., Marchant, A.E., Ogden, M.C., et al.: Implementation of a Tight Glycaemic Control Protocol using a Web-based Insulin Dose Calculator. Anaesthesia 60, 1093–1100 (2005)
21. Went, K., Antoniewicz, P., Corner, D.A., et al.: A Survey of Prescribing Errors in Intensive Care. Scot. Med. J. 52, 50 (2007)
22. Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)
23. Dix, A., Finlay, J., Abowd, G.D., et al.: Human-Computer Interaction, 3rd edn. Pearson Education Ltd., London (2004)
24. Nielsen, J.: Heuristic Evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods, pp. 25–62. John Wiley, New York (1994)

A Novel Approach for Creating Activity-Aware Applications in a Hospital Environment

Jakob E. Bardram

IT University of Copenhagen
Rued Langgaards Vej 7, DK-2300 Copenhagen, Denmark
bardram@itu.dk

Abstract. Context-aware and activity-aware computing has been proposed as a way to adapt the computer to the user's ongoing activity. However, deductively moving from physical context – like location – to establishing human activity has proved difficult. This paper proposes a novel approach to activity-aware computing. Instead of inferring activities, this approach enables the user to explicitly model their activity, and then use sensor-based events to create, manage, and use these computational activities adjusted to a specific context. This approach was crafted through a user-centered design process in collaboration with a hospital department. We propose three strategies for activity-awareness: context-based activity matching, context-based activity creation, and context-based activity adaptation. We present the implementation of these strategies and present an experimental evaluation of them. The experiments demonstrate that rather than considering context as information, context can be a relational property that links 'real-world activities' with their 'computational activities'.

1 Introduction

Context-aware computing has been proposed as a way to design mobile computer applications that adapt according to the physical context of the user, like his or her location. For example, if a mobile phone is inside a cinema, then it is automatically silenced [17]. In the nomadic work environment of a hospital, context-aware technologies seem a promising approach. For example, a physician using a PDA may be presented with relevant medical data on the patient, he or she is approaching [1, 19, 9].

However, deductively moving from physical context – like location – to establishing human activity has proved difficult [7]. The mobile phone in the cinema may belong to a service technician performing a repair activity, in which case silencing the phone is annoying. Similarly, a physician may be in front of a patient in order to prepare some medical equipment, like a catheter insertion, and would be more interested in a clinical guideline for this procedure rather than the patient's medical record [9].

In order to mitigate these challenges, several researchers have proposed to build activity-aware systems, i.e. systems which are able to recognize not only the context of the user, but the overall activity that the user is engaged in. For example, Favela et. al [9] argues that the challenge in the medical scenario is to make better sensor systems and machine learning algorithms for deducing the 'right' activity. Activity-aware

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 731–744, 2009.

computing typically entails sensing physical and digital context like location, time, movement, and usage of e.g. files, and then trying to deduce the overall activity of a person from a synthesis of current and historical sensor input. This synthesis is then used to adapt the computer to the activity of a user. In this paper, we propose another approach to activity-aware computing; instead of trying to deduce activities from lower-level sensor input, users help to explicitly define their activities as part of a computer system. This approach entails a range of benefits, both for system designers as well as for users, because it removes a huge part of the uncertainty associated with activity-aware computing based on activity inferring. It is relevant to note that this approach is not in conflict with context-aware computing and activity inference based on sensor input – as we shall see, our approach also make use of such sensor techniques. However, we view activity-aware computing as a way to apply these techniques within a framework for creating and managing real-world activities, which are defined by the users and not the computer.

2 Related Work

There is a wide range of systems and applications that support context-aware adaptation of an application to the specific context of the user. In the medical domain, researchers have been investigating how to build context-aware systems for mobile technologies in hospitals [13, 19, 1]. However, because these approaches to context awareness treat 'context' as a matter of representing context and then deduce human activity from it, they often run into the problem that the scope of contextual features is defined dynamically [7], in which case it is difficult to establish what a context event 'means', i.e. what kind of human activity is going on.

In order to support human activity more directly, a range of approaches seek to support activity management and representation. These are systems like Unified Activity Manager [16], Activity Explorer [18], and Activity Based Computing [3]. Because these systems require users to manually create, maintain, and clean up the computational representations of activities, they provide little or no support for connecting the computer representation of activities to the real-world human activities.

To address these issues, a set of systems has been proposed to automatically generate its task or activity representations from logs of low-level, interaction events. These systems include TaskTracer [8], UMEA [11], Kimura [15], and CAAD [20]. However, these kinds of low-level interaction event with a computer are of little use in e.g. a hospital environment where the important context events are related to the broader clinical context. Another set of systems tries to use extensive sensoring of an environment and run machine learning algorithms to deduce human activities of e.g. daily living (e.g.[14]). In a hospital domain, Favela et al. [9] has presented a neural network system that infers activities of clinicians. Our observations are similar to Favela et al., but we come to different conclusions. Whereas Favela et al. emphasize the need for better sensing and machine learning technology, we rather move in the direction of having the user to help explicitly represent relevant activities and intentions in the computer system.

Fig. 1. Inside the office at the ward. Coordination and information management done by the physician and the nurse during the ward round.

In summary, there are three types of systems. First, context-aware systems that tries to sense, model, and reason about context, and use this for adapting an application. Second, systems which provide support for activity management. And third, systems for activity deduction and inference. However, there are no systems that uses context events to help users bridge between 'real-world activities' and 'computational activities', i.e. representations of activities which are already modeled and created by the user in the computer system. This latter approach of combining a real-world activity with its computational counterpart is the focus of our approach to activity-aware systems.

3 Research Background

The present approach to activity-awareness is rooted in a long-term engagement in the design of pervasive computing technologies for hospital work. Clinicians in hospitals attend many concurrent activities in a very mobile, collaborative, and hectic environment, while accessing a wide range of information sources. For example, Figure 1 shows a nurse and a physician preparing for the ward round. In this case, they need to juggle several patient cases while using a large set of data sources and tools: the medical record, the medicine chart, the computerized lab system, the medicine handbook, referral letters, etc. A core challenge to the current work environment is that there is a substantial overhead associated with finding and aligning all the electronic and physical material that is needed in order to discuss a patient case and get ready for e.g. the ward round. Furthermore, once the physician and nurse leave the ward office, they need to get access to this information for each of the patients they may visit at their bedside.

The research site was an internal medicine department, specialized in hematology at a large metropolitan teaching hospital. The work at a hematology department is highly collaborative and involves a wide range of specialized doctors, like pathologists, hematologists, radiologists, and hematopathologists, as well as lab technicians, nurses, and care assistants. In order to investigate how activity-aware technologies could be designed for this time-critical, mobile, and collaborative work setting, we engaged in a user-centered participatory design process, which consisted of three main strands of research: (i) ethnographic field studies of work in hospitals, including participant observation and video analysis [21]; (ii) a series of design workshops applying the future workshop methodology [12]; and (iii) scenario-based evaluation [5] of prototypes in a simulated hospital environment. The project took place over a two-year period with 3two-month periods of ethnographic fieldwork, 12 open-ended interviews, and 11 design and evaluation workshops involving ca. 20 different clinicians from the hospital.

4 Designing for Activity Awareness

The field studies at the hematology department showed that clinicians need support for;(i) timely retrieval of large amount of medical information, which is relevant for specific patient activities; (ii) frequent and fast switching between activities; (iii) mobile support for activities; and (iv) continuous multi-tasking. This raised a number of questions on how to create computer support for activities, and especially how to design systems that are somehow aware of these activities and the context in which they take place. This was subsequently the focus of the design process.

The design for activity-aware computing builds upon the Activity-Based Computing (ABC) framework [4, 2, 3]. The ABC framework has been designed to support users– and especially clinicians working in hospitals – to manage a set of parallel and collaborative activities. In the framework, a real-world activity is supported by grouping resources like files, documents, records, and services into a bundle which is called a computational activity – or just 'activity'. For example, at the hematology department, at least one activity would be associated with each patient, and such an activity would bundle relevant medical data like the medical record, the medicine chart, relevant clinical guidelines, x-ray images, etc. Clinicians access these activities through activity resume and suspend, which enables the user to bring up a computational activity and its associated resources on publicly available displays within the hospital, and dismiss of it once finished using it. This has the effect that patient information relevant for the ongoing treatment is fetched and displayed. An activity can be shared amongst a set of participants which all can resume and use the activity during patient treatment. For example, the physician and the nurse doing the ward round together would share activities on relevant patients, and hence both have access to the medical resources for the patients.

The goal of activity-based computing is to help users manage complex real-world activities that involve heavy information management on computers. In contrast to systems that try to deduce real-world human activity though sensing technology, the approach in the ABC framework is to have users explicitly create, define, and maintain the activities, which they find relevant. The framework has mechanisms for users to define their activities, arrange the computational resources associated with these activities, and manage the set of people who participate in an activity. As such, the ABC framework has mechanisms for supporting the handling of large bundles of medical information relevant for different patient cases, and provides support for mobility and collaboration.

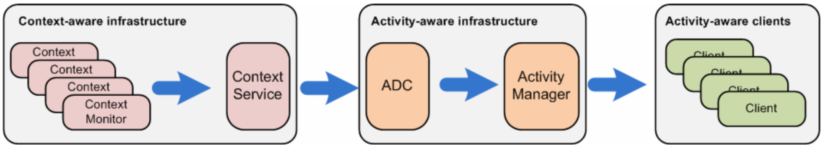


Fig. 2. A simplified picture of the ABC activity-aware architecture

During the design and evaluation phases, it was, however, realized that a core challenge to the ABC framework was to help users manage the dynamic complexity of these clinical activities; i.e. to help users create, maintain, navigate, and handle such activities, their resources, and participants in order to constantly ensure that the computational activities are aligned with the real-world activities taking place in the work environment. To help users with this task of maintaining an association between real-world activities and their computational counterparts, we made the ABC framework activity-aware by having the framework maintain an awareness of the real-world activities going on, and thereby help the users to define, maintain, and resume/suspend appropriate computational activities. Based on the series of design workshops with the clinicians, we identified three relevant strategies for activity awareness: (i) context-based activity matching;(ii) context-based activity creation; and (iii) context-based activity adaptation.

4.1 Context-Based Activity

The goal of context-based activity matching is to help users retrieve relevant information in a specific context. This is one of the earliest goals of context-aware computing, and has also been proposed in a medical setting – e.g. displaying the patient's medical record when the physician is near the patient's bed [1, 19]. In general, the goal is to somehow infer the activity of the user and use this for 'activity-aware information retrieval' [9].

From the field studies and the design workshops, it was evident that this kind of activity-aware information retrieval would be highly relevant in order to assist clinicians in handling the vast amount of medical information associated with patient treatment and care. However, taking a closer look, it is actually difficult to establish exactly what we mean by "display[ing] the medical record when a physician is in front of a patient's bed" [9;p.166]. The patient's medical record is in most cases extremely voluminous containing a wide variety of information ranging from hundreds of test results, tens of x-ray images, a very long medical continuation, medicine charts for dozens of medical prescriptions, extensive nursing notes, etc. Clearly we do not want to automatically bring all that up in front of the physician just because he happens to be approaching a patient's bed.

The approach that we propose by the concept of context-based activity matching is not to try to figure out what information is relevant for a specific activity, because this has already been defined by the user in the creation and population of his or her activities. The approach is therefore simply to make a link between specific context events – like the physician's location in front of the patient's bed – to a relevant activity. For example, the presence of the patient may trigger a suggestion to resume the computational activity associated with this patient's treatment, which will bring up the information that the physician has deemed to be relevant for the treatment of this particular patient. Another example is to associate the activity of handling medicine for a patient with the patient's pill tray, i.e. the tray nurses are using to hand out medicine according to the physician's prescriptions. When this pill tray is used in a specific context– e.g. in the medicine room or at the patient's bed side – the ABC framework will suggest the user to resume this 'handling medicine' activity.

The implementation of context-based activity matching makes use of the architectural components illustrated in Figure 2. Information about context is maintained by the context-aware infrastructure, which uses context monitors to sense context entities. An RFID scanner scanning an RFID-enabled pill tray is an example of a context monitor and the pill tray itself is an example of a context entity. Context information is stored and managed by a context service. The activity detection component (ADC) listens to context events in the context service, and based on a mapping between context entities (like the pill tray), the ADC can match context events with one or more activities. If a matching activity is discovered, the activity manager is notified, which again notifies relevant activity-aware clients.

4.2 Context-Based Activity Creation

The goal of context-based activity creation is to assist users in creating relevant activities in different context. This is relevant because activity-based computing relies on users to create their own computational activities and populate them with relevant resources. Context-based activity creation simply comes up with suggestions for relevant activities including resources and participants. For example, creating 'Hematology Treatment' activities to be used for patients admitted to the hematology department.

The approach proposed by the concept of context-based activity creation is to use generic activity templates. These templates can be used by the user when manually creating a new activity. For example, a 'Hematology Treatment' template would be used by the physician to create an activity for the treatment of a specific patient. These activity templates can, however, also be used for context-based activity creation. The activity templates can take as input current context information about an entity and based on this context, the template populates and returns an instance of an activity. For example, once a patient is being admitted to the hematology department, this event can trigger the creation of a 'Hematology Treatment' for this patient, and the template would be able to fill in all the details about this patient, seek out relevant information in the medical record, set up a medicine chart, and prepare the ordering of relevant lab tests. Once this patient's physician and nurse is known, they will become participants of the activity, and hence be able to access the activity and its associated medical resources. Another example is to use an 'Operation Template' to create activities based on scheduled operations at a surgical ward. This template would populate an operation activity with relevant patient information and information about the operation, and add the scheduled clinicians as participant. This activity can then be used for easy information access during the actual surgical procedure, and for later follow up in the outpatient clinic.

In the ABC framework, context-based activity creation is implemented as part of the ADC. The ADC listens to relevant context events generated by different monitors, and when a certain context event occurs, the ADC uses an appropriate activity template to create a new activity based on the context information. This new activity is then forwarded to the activity manager, which again notifies the activity-aware clients. For example, an 'Operation Scheduler Monitor' may listen to the operation scheduling system. For every new patient being scheduled, the Context Service notifies the ADC which again uses an 'Operation Activity' template to create a new activity, which links to medical data relevant for the surgery for each patient.

4.3 Context-Based Activity Adaptation

The goal of context-based activity adaptation is to help users to maintain their activities. Activity-based computing fundamentally relies on users to maintain and nurture their computational activities according to their need for supporting different real-world activities. Context-based activity adaptation is aimed at helping users in this maintenance task by using contextual information to continuously adapt the activities used in different context. As also pointed out by Favela et. al [9], the physician may be in front of a patient in order to prepare a catheter insertion, in which case the clinical guidelines and pharmacological databases, rather than patient's health record, may be appropriate in this context.

The approach taken in context-based activity adaptation is to investigate the context in which an activity is being resumed or suspended. If there are resources, services, or information that seems to be relevant in the current situation, these are added to the activity. For example, if we have some knowledge about the presence of a catheter insertion needle, then it would be relevant to add the guideline as a resource to the activity. Another example is when resuming an activity inside the operating room. In the operating room there are different services, which are relevant for the operation activity, but only available in here. For example, different audio and video recording services. These services are added to the activity if the activity is resumed inside the OR, but are not part of the activity otherwise. Similarly, inside the OR different resources like a clinical guideline for this particular type of operation and a patient safety checklist are also relevant in this context, and may thus be added to the activity.

Context-based activity adaptation is implemented using the activity template introduced above. Each activity maintains a link to its originating template, and this template contains the logic, which can adapt the activity according to the context in which it is being resumed or suspended. For example, when an activity based on the surgical template is being resumed inside an operating room, links to resources like video cameras and operating manuals for the equipment inside this specific operating room are added to the activity. These resources are removed again once the activity is suspended and no longer used inside the operating room.

5 Experiments and Evaluations

The concepts of activity-based computing, activity-awareness, and the three strategies presented above are of an abstract, conceptual nature. In order to establish how these principles could be implemented in use, a range of experiments and evaluations have

Fig. 3. Experimental setup. Left: the 'Medicine Room' showing the nurse, a table with an RFID reader, a pill tray on the table, and a display. Right: the 'Patient Bed' showing the nurse, the patient, and a build-in display.

been conducted. These experiments have been done throughout the two-year period. In total 11 whole-day workshops were conducted in an experimental facility, which was equipped like a medical ward, including bed rooms, a hallway, the ward office, a medical conference room, and a medicine room. In all these places, different public displays were available. For example, a large wall-display in the wards office, a build-in display in the hospital bed (Figure 3, right), and a touch display in the medicine room (Figure 3, left).

The workshops applied scenario-based design and evaluation methods [5], allowing the clinicians to try out different design alternatives. The scenarios were created based on the field studies and refined together with the clinicians before they were enacted. We expected the clinicians to relate primarily to the usability aspects – i.e. issues related to the user interface of our prototypes. We were, however, quite surprised how well we were able to present and discuss the conceptual aspects of our design with them, and we were often engaged in detailed and very sophisticated discussions of appropriate solutions to activity-based computing and activity-awareness in a clinical environment. From a methodological point of view, we hence did not make any formal 'evaluation' as such. From a exploratory research perspective, however, we had substantial feedback on our proposed designs, which – we would argue – to a much larger degree helped to establish an appropriate design for activity-aware computing (c.f. [10]).

5.1 Context-Based Activity Matching

Context-based activity matching tries to match context events to relevant activities, which in the ABC framework is done through linkage between a context event and a

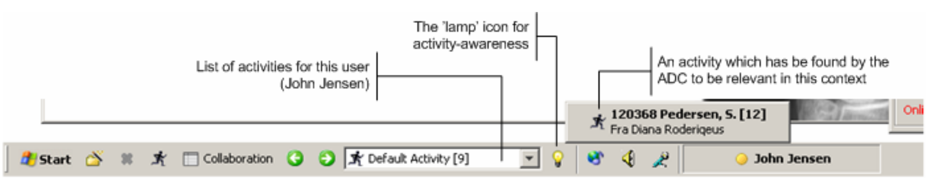


Fig. 4. The Activity Bar showing the 'lamp' icon, which is used to list the activities discovered by the ADC as relevant in this context

computational activity. In the experimental setup, we created support for this kind of linkage based on the location of different physical objects. This was implemented using passive RFID tagging of various clinical objects, such as the patient, the patient's paper-based medical record, x-ray photographs, and various medical utensils like pill trays and surgical instruments. In the event that such an object was seen, the ABC framework would notify the user that it had discovered an activity relevant in this context.

Figure 3 (left) shows a picture from one of our workshops where the nurse is preparing medicine inside the medicine room. She is using a pill tray for a specific patient and by placing this pill tray on the table (which has a build-in RFID scanner), the public display in the medicine room provides access to this patient's medication activity.

One of the open questions was how to present to the user that an activity had been discovered. We experimented with three prototypes, each taking into account the lessons learned from the previous one. In the first prototype, just placing the pill tray on the table would trigger the ongoing activity to be suspended and the activity related to the pill tray (typically an activity on a patient) would be resumed. This had the effect that all previous medical data on the display were removed and medical data related to the patient would be displayed.

This approach of automatically displaying information based on the presence of a tagged physical object has been suggested by a number of researchers [22] and also in the medical domain [19]. Our experiments, however, showed that this was not an appropriate design solution; physical objects may be in proximity of a public display for many reasons and may hence not justify that the user's current activity is being resumed and disappears from the screen. For example, the nurse working in a medicine room will be handling dozens of pill tray and some of these may be placed on the table just for convenience sake.

Based on this input, the second prototype presented a list of discovered activities as buttons in an 'Activity Bar' and the user could select the activity by pressing the corresponding button. Now the user would decide on the activity switching while easy access to the activity was maintained. The clinicians agreed that this solution was to prefer over the first one due to its non-intrusive nature. This activity bar had, however, some limitations regarding the number of activities it could hold. The third design of the prototype therefore listed the set of discovered activities in a pop-up menu accessible from a button on the activity bar. This third prototype implementation is shown in Figure 4 where the 'lamp' button blinks if there are any discovered activities. When clicking on the icon, the user has access to a list of activities, which the activity-awareness system deems to be relevant in this context.

A couple of lessons came out of these experiments. First, the users found the notion of linking activities and their associated resources to real-world objects and events very useful. It provided the important connection between the material world of patient treatment and the increasing amount of digital information and services being used as part of patient treatment. Second, the experiments showed that automatic activity switching should be done with outmost caution, and that activity matching with real-world context events should be done in a non-intrusive way.

5.2 Context-Based Activity Creation

Context-based activity creation aims at creating activities based on relevant context events, which in the ABC framework is done using activity templates. In our experimental setup we created a set of templates, which could generate relevant activities indifferent context. For example, a template for handling medicine for a patient. This 'medicine handling' template was used to generate an activity in the medicine room; when the nurse placed a patient's medicine tray on the table, this template was used to generate a new activity which was populated with resources relevant for medication of this particular patient. Resources were things like the medicine chart, relevant medical guidelines, and the medicine handbook (which holds a list of all available medicine). This activity was then added to the user's list of discovered activities, and depending on the version of the prototype, the user had access to this activity in different ways as described above.

Another experiment was to use a logic inference engine as part of the activity discovery component (ADC) [6]. This was done in order to extend the one-to-one relationship between a context event and an activity by allowing a set of context events to be parsed into creating an activity. For example, in order to bring up relevant information when the nurse was taking the prepared medicine to the patient's bedside, the ADC would react on the combination of the right pill tray for the patient, a certified nurse, and the presence of the patient. In this case, the ADC would create an activity, which could display the medicine chart for this patient, and in this chart highlight the medicine in the pill tray. This activity was then shown on the bed's build-in display as shown in Figure 3 (right). This approach has the benefit of presenting and highlighting relevant information in this specific context and furthermore has the benefit of being able to detect, if the nurse is trying to give the patient the wrong pill tray.

The clinicians clearly acknowledged the benefit of increased patient safety during medication administration. However, they did not find the solution to scale to the work in a hospital, because several contingencies could arise even with this rather simple scenario; the patient could be away from the bed, the person giving the medicine may be a nursing assistant, and the pill tray may be missing. Furthermore, the clinicians argued that there was actually no need for generating new activities every time a context event occurred – the activity for handling medicine in the medicine room is the same activity as giving the medicine to the patient at the bedside. This approach would therefore automatically generate a large number of minor computational activities belonging to the same overall human activity. Consequently, the clinicians preferred the simple object-to-activity linking where both the pill tray as well as the patient would link to the overall activity of treating the patient.

These lessons lead to the current design of activity templates that help users create computational activities, which are aligned with real-world activities. One example is a template suited for the medical treatment done at the hematology department, which can be used when a patient is admitted to the department. Another template could retrieve and present a set of x-ray images in a way, which is suited for the radiology conference. The clinicians deemed these kinds of activity templates to be more suitable than the detailed ones regarding e.g. medicine handling.

5.3 Context-Based Activity Adaptation

An important lesson coming out of our experiments with context-based activity matching and creation is, that context events not necessarily justify activity switching; in many cases there is no need for shifting from one overall activity to another, but rather to adjust the current activity. For example, when the nurse is placing a pill tray on the table she needs to view the medicine chart. However, in the case where the medicine chart for some reason is not part of the patient's activity, the appropriate course of action would not be to switch activity, but merely to add the medicine chart to the patient's activity. A similar argument can be made in the case where a physician is using the catheter insertion equipment in front of the patient. In this case, relevant medical guidelines can be added to the patient's current activity. This strategy ensures that the overall context of the treatment is maintained, but supplemented with resources relevant in this specific situation. Observations like these lead to the design of context-based activity adaptation.

During the workshops we discussed this adaptation approach, and this kind of adaptation of the activity to its use in a specific context was in general appreciated by the clinicians as a solution to some of the issues raised above. Basically, the idea of having fever overall activities that evolve over time was preferred to the idea of creating several smaller activities suited to a specific context. The adaptation involving adding and removing resources and services like the medicine chart was also deemed as less intrusive compared to automatic activity switching; the overall activity was maintained active and only supplemented with additional resources. If these resources – like the medicine chart – were considered inappropriate in this situation, it was quite easy to just close the medicine chart window. In other words; the consequence of showing additional resources to an activity is more balanced with the accuracy of scanning passive RFID tags.

6 Discussion and Conclusion

Traditionally in context-awareness research, context is viewed as a representational problem [7]. This means that context-awareness is inherently restricted to the context information that can be sensed and modeled, and progress in context-aware computing therefore relies on the basic expectation that we can make better sensors, models, and reasoning engines. In this line of reasoning, several researchers have proposed to create context-aware systems for use in hospitals, which reveal patient-related medical data when a clinician is close to the patient [13, 19, 1]. However, despite any technological development in sensing, modeling, and reasoning technologies, context-aware computing will always be limited to what has been sensed – any reasoning is no better that its raw data. Thus – if the patient is not nearby the physician, it is hard to use sensing of the patient as a means of locating the patient's medical record. Or – if the proximity of the patient in the bed wards brings up the medical record, is it then feasible to bring up the medicine chart as well?

Our approach to activity-aware computing is simpler; it basically asks the user to help define the computational activities that match whatever real-world activity he or she is doing, or wants to do. These computational activities embrace and provide easy access to resources, material, and services, which are relevant for performing the corresponding human activity. By having a model of the activity in the computer, the whole context can be restored – also resources which in no way can be deduced from the current context. Hence, physical context events, like the location of the patient, become links into a much richer computational context.

The drawback of this approach is, however, that users need to model, define, create, and nurture these computational activities. To mitigate this drawback, our approach to activity-aware computing helps users to continuously maintain and adapt their activities according to relevant information and events arising from the real-world context in which the corresponding human activities unfold. We have proposed three strategies for supporting activity-aware computing: context-based activity matching, context-based activity creation, and context-based activity adaptation. Through a user-centered design process, these strategies have been implemented in a set of prototypes, which were evaluated experimentally with a group of hospital clinicians.

Several lessons came out of these experiments. First, the support for context-aware activity matching based on linking computational activities with their associated resources and data, to context events from real-world objects were considered a useful and conceptually understandable strategy. It provided the linkage between the highly material world of patient treatment and the increasing amount of digital information and services being used as part of patient treatment. Second, the experiments showed that automatic activity switching should be done with outmost caution, and that activity matching with real-world context events should be done in a non-intrusive way. Third, the experiments showed that more advanced rules regarding the linkage between a set of context-events and a relevant activity may be too complicated to maintain and to use. The clinicians all agreed that having a one-to-one linkage between a physical object and an activity was more appropriate both conceptually as well as in use. Fourth, creating computational activities based on events in the real world was considered very useful for activity management in a hectic and ever-changing environment like a hospital. Based on our experiments, however, we also found that support for e.g. context-aware information display should be tied into the overall activity in which it makes sense; for example, presenting the medicine chart in the medicine room or the medical record at the patient's bedside does not in itself constitute an 'activity', but the medical chart and record are important resources that need to be made available in these settings. This leads to the fifth lesson learned, which revealed that the computational infrastructure should help users to constantly maintain and adapt their computational activities according to the flow of events happening as work unfold in the real world. Activity adaptation involving adding and removing resources and services, like the medicine chart in the medicine room, was deemed as less intrusive compared to automatic activity switching, because the overall activity was maintained active and supplemented with additional material.

The proposed approach to activity-aware computing as presented in this paper has only been implementing simple scenarios with easy detectable objects (like the pill tray) and has been subject to tests in an artificial lab test setup. A concern that can be raised about the application of this approach to the everyday working practices of complex environments, like hospital units, is whether the user intervention required to model the activity in question may become too demanding in terms of costs/benefits ratio, once the activity-based computing system becomes more sophisticated and it comprises several computational activities to be maintained. To address this scalability and complexity question, we are currently undertaking a new implementation of the proposed techniques that aims at handling far more complex context-based activity matching, context-based activity creation, and context-based activity adaptation. This would involve a large amount of real-world activities and be tested inside a hospital.

Acknowledgments. Claus Bossen and Henrik B. Christensen were participating in the study and design for hospital work. We are indebted for the enthusiastic participation from the clinicians at the Hematology department at the University Hospital in Aarhus – especially Jørgen S. Christensen. This work was funded by the Danish Research Council.

References

1. Bardram, J.E.: Applications of Context-Aware Computing in Hospital Work - Examples and Design Principles. In: Proceedings of the 2004 ACM Symposium on Applied Computing, pp. 1574–1579. ACM Press, New York (2004)
2. Bardram, J.E.: Activity-Based Computing: Support for Mobility and Collaboration in Ubiquitous Computing. Personal and Ubiquitous Computing 9(5), 312–322 (2005)
3. Bardram, J.E., Bunde-Pedersen, J., Soegaard, M.: Support for activity-based computing in a personal computing operating system. In: Proceedings of CHI 2006, pp. 211–220. ACM Press, New York (2006)
4. Bardram, J.E., Christensen, H.B.: Pervasive computing support for hospitals: An overview of the activity-based computing project. IEEE Pervasive Computing 6(1), 44–51 (2007)
5. Bødker, S., Christiansen, E.: Scenarios as springboards in design. In: Bowker, G., Gasser, L., Star, L., Turner, W. (eds.) Social science research, technical systems and cooperative work, pp. 217–234. Erlbaum, NJ (1997)
6. Christensen, H.B.: Using Logic Programming to Detect Activities in Pervasive Healthcare. In: Stuckey, P.J. (ed.) ICLP 2002. LNCS, vol. 2401, p. 421. Springer, Heidelberg (2002)
7. Dourish, P.: What we talk about when we talk about context. Personal and Ubiquitous Computing 8, 9–30 (2005)
8. Dragunov, A.N., Dietterich, T.G., Johnsrude, K., McLaughlin, M., Li, L., Herlocker, J.L.: TaskTracer: a desktop environment to support multi-tasking knowledge workers. In: Proceedings of the 10th international conference on Intelligent user interfaces, pp. 75–82. ACM Press, New York (2005)
9. Favela, J., Tentori, M., Castro, L., Gonzalez, V., Moran, E., Martinez-Garcia, A.: Activity recognition for context-aware hospital applications: Issues and opportunities for the deployment of pervasive networks. Mobile Networks and Applications 12(2), 155–171 (2007)
10. Greenberg, S., Buxton, B.: Usability evaluation considered harmful (some of the time). In: Proceeding of CHI 2008, pp. 111–120. ACM, New York (2008)
11. Kaptelinin, V.: UMEA: Translating interaction histories into project contexts. In: Proceedings of the CHI 2003, pp. 353–360. ACM Press, New York (2003)
12. Kensing, F., Halskov Madsen, K.: Generating Visions: Future Workshops and Metaphorical Design. In: Greenbaum, J., Kyng, M. (eds.) Design at Work: Cooperative Design of Computer Systems, pp. 155–168. Lawrence Erlbaum Associates, Hillsdale (1991)

13. Kjeldskov, J., Skov, M.: Supporting work activities in healthcare by mobile electronic patient records. In: Proceedings of the 6th Asia-Pacific Conference on Human-Computer Interaction (2004)
14. Logan, B., Healy, J., Philipose, M., Tapia, E.M., Intille, S.: A long-term evaluation of sensing modilitis for activity recognition. In: Proceedings UbiComp 2007, pp. 483–500 (2007)
15. MacIntyre, B., Mynatt, E.D., Vodia, S., Hansen, K.M., Tullio, J., Corso, G.M.: Support for Multitasking and Background Awareness Using Interactive Peripheral Displays. In: Proceeding of ACM User Interface Software and Technology 2001 (UIST 2001), Orlando, Florida, USA, November 2001, pp. 11–14 (2001)
16. Moody, P., Gruen, D., Muller, M.J., Tang, J., Moran, T.P.: Business activity patterns: A new model for collaborative business applications. IBM Systems Journal 45(4), 683–694 (2006)
17. Moran, T., Dourish, P.: Introduction to this special issue on context-aware computing. Human-Computer Interaction 16, 87–95 (2001)
18. Muller, M.J., Geyer, W., Brownholtz, B., Wilcox, E., Millen, D.R.: One-hundred days in an activity-centric collaboration environment based on shared objects. In: Proceedings of CHI 2004, pp. 375–382. ACM Press, New York (2004)
19. Munoz, M., Rodriguez, M., Favela, J., Martinez-Garcia, A., Gonzalez, V.: Context-aware mobile communication in hospitals. IEEE Computer 36(9), 38–46 (2003)
20. Rattenbury, T., Canny, J.: CAAD: An Automatic Task Support System. In: Proceedings of CHI 2007, pp. 687–696. ACM Press, New York (2007)
21. Suchman, L.A., Trigg, R.H.: Understanding Practise: Video as a medium for Reflection and Design. In: Greenbaum, J., Kyng, M. (eds.) Design at Work: Cooperative Design of Computer Systems, pp. 65–90. Lawrence Erlbaum Associates, Hillsdale (1991)
22. Want, R., Fishkin, K.P., Gujar, A., Harrison, B.L.: Bridging physical and virtual worlds with electronic tags. In: Proceedings of CHI 1999, pp. 370–377. ACM Press, New York (1999)

Investigating CAPTCHAs Based on Visual Phenomena

Anja B. Naumann[1], Thomas Franke[2], and Christian Bauckhage[3]

[1] Deutsche Telekom Laboratories, Ernst-Reuter-Platz 7, 10587 Berlin, Germany
[2] Cognitive and Engineering Psychology, Chemnitz University of Technology
[3] Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS
anja.naumann@telekom.de

Abstract. We propose and evaluate several novel types of CAPTCHAs (test to tell computers and humans apart) that exploit characteristics of the human visual system. Perceptions caused by the effect of lightness constancy or grouping phenomena due to transparent motion are hard to emulate on computers and may thus provide novel authentication mechanisms.

Keywords: CAPTCHA, visual phenomena, optical illusions.

1 Introduction: Visual Phenomena for CAPTCHA Design

Upon registration to a service, users often are faced with a little cognitive puzzle which is supposed to be difficult to solve algorithmically and thus is meant to prevent bots from misusing the service. However, the designer of such CAPTCHAs (*completely automated public turing test to tell computers and humans apart*) faces a trade-off decision. On the one hand, the principle behind the puzzle must be difficult enough in order to resist attacks through pattern recognition algorithms. On the other hand, it must be simple enough to allow for the automatic generation of many unambiguous instances which even naive users can solve quickly. Unfortunately, attempts of automatically breaking common types of simple cognitive challenges have become more successful recently. Straightforward countermeasures to these attacks, however, have led to CAPTCHAs of decreased usability for humans. This is especially evident for challenges that require the user to read a distorted text. In reaction to a growing number of reports on successfully using optical character recognition to break this challenge (see e.g. [1]), more recent types of CAPTCHAs therefore require the user to distinguish pictures [2] or to annotate a picture using a given list of terms [3]. However, mathematical combinatorics suggests that text-based CAPTCHAs can pose a more secure challenge [2]. Furthermore, people are used to solving these and will probably better understand and accept them [4].

In this paper, we investigate a novel approach to the design of reading-based cognitive challenges. Our idea is to exploit specific capabilities of the human visual system which are hard to emulate algorithmically and therefore will increase the difficulty for machines while they are supposed to be still easy to solve for humans. The visual phenomena we used are the following:

Lightness Constancy. In the physical world, the visual appearance of objects changes under varying illumination. Nevertheless, we are not constantly puzzled, our perceptual

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 745–748, 2009.

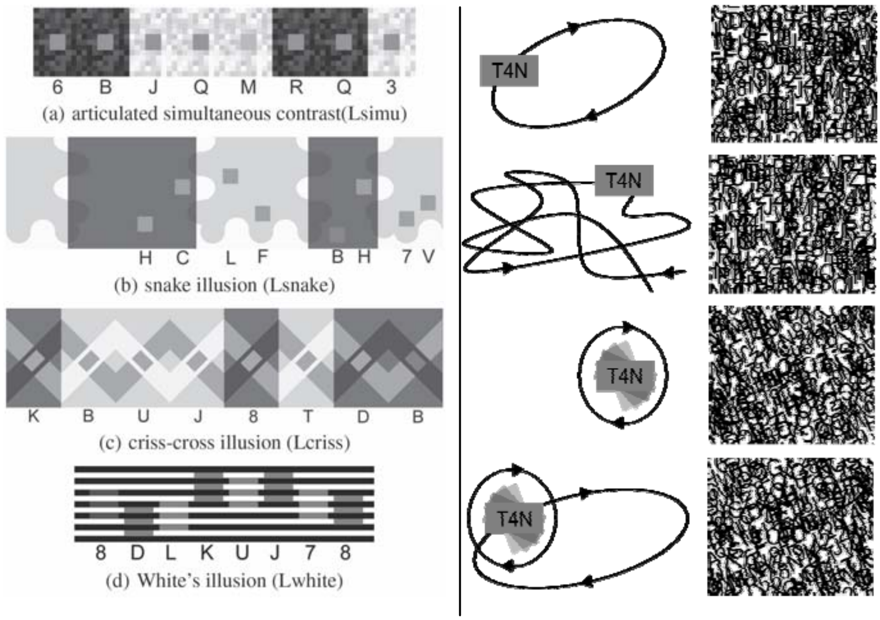


Fig. 1. Left: *Lightness illusions*. The small rectangles in 1(a) to 1(c) and the grey stripes in 1(d) are in fact all of the same brightness, even if it appears not to be the case. When asked to select the characters below the brighter rectangles, people thus select only a subset of characters. Right: The four types of *transparent motion CAPTCHAs*. A text shown in front of a noisy background becomes legible if text and background are moving differently. From top to bottom the different motion patterns are: circular motion (*Mcircle*), random motion (*Mrand*), rotation (*Mrotate*), and circular motion plus rotation (*Mcirro*).

system subconsciously compensates for these variations. Part of this effect is known as the phenomenon of lightness constancy. Under certain conditions, however, the mechanisms providing constancy may also cause illusions. In this work, we draw on four illusions that are known to be persistent and thus may apply in the context of CAPTCHA design. Details on the cognitive mechanisms behind these effects can be found in [5]. Examples are shown in Fig. 1. In practice, any number of similar pictures can be produced (e.g. using circles or ellipses instead of squares, or varying sizes). Hence, for breaking the CAPTCHAs, looking for regularities in pictures is not sufficient.

Motion-Defined Form. The second type of CAPTCHAs we introduce is based on the phenomenon of motion-defined form [6] which is based on perceptual grouping: since our visual system tends to group different entities that move together, sketches or letters superimposed over a noisy background of the same color become visible once they are moving. We experimented with a moving group of three characters superimposed over a moving picture of many cluttered characters (see Fig. 1.)

2 Experimental Evaluation

Evaluation Criteria. In order to assess the utility and usability of the CAPTCHAs illustrated in Fig. 1, we defined minimum requirements to be met: According to [1], humans should solve a CAPTCHA in more than 80% of the cases. Also, many interaction proofs have been reported to be solvable within 15 seconds [e.g. 2]. With respect to subjective difficulty, we demand that on a scale ranging from 1 (very easy) to 5 (impossible), a novel CAPTCHA should be at least rated to be manageable (score 3). For subjective usability, given a scale ranging from 1 (not at all problematic) to 5 (problematic), most users should assign scores lower than 3.

Setup and Procedure. A total of 28 subjects with normal vision (19-31 years; 17 female, 11 male) participated in our study. On average, they rated their computer- and Internet skills as moderate to moderately high (3.62 and 3.79 on a scale from 1 to 5). Most participants indicated that they had solved CAPTCHAs 10–50 times before, only 3 participants had solved less.

Our study was conducted in a laboratory setting. The CAPTCHAs were displayed using a web browser on a 19" CRT monitor (1024 x 768 pixels). All of them were positioned and accompanied by textual elements and entry boxes just as on a real website. They were normalized such that the entities that had to be identified (characters or rectangles) were of the same size across all experimental conditions.

When one of our lightness CAPTCHAs (see Fig. 1) was presented, the subjects were asked to only type the characters below a light or dark rectangle. The attribute light or dark was randomly changed from trial to trial. The challenges required to correctly enter 3 characters. In each image, there were 8 rectangles, 4 of which appeared in a brighter context and the remaining 4 appeared in a darker context. Also, in order to prevent simple automatic solutions, the brightness of one of the rectangles was slightly increased or decreased. For our motion CAPTCHAs (MPEG movies, 320 x 240 pixel), subjects were instructed to enter the characters that stood out through their motion. The challenges required to correctly enter 3 characters moving as a group over a differently moving cluttered background of randomly arranged characters. Speed of motion was 125 pixels per second for the *Mcircle* and *Mrand* variants. For *Mrotate*, the background was rotated at a speed of 60° per second while the foreground rotated with 160° per second. For the combined circular and rotating motion (*Mcirro*), the motion speeds were chosen as in the individual cases.

In each trial, the subjects entered what they believed was the correct answer. Their responses, i.e. response accuracies and response times were recorded. Then they were asked to subjectively rate the difficulty of the CAPTCHA. For each type of CAPTCHAs used for comparison and for each subtype of our proposed CAPTCHAs, a 10-item questionnaire was handed out to assess subjective usability. The questionnaire was referring to CATPCHA acceptance, i.e. it regarded work load, joy of use, and satisfaction, and was motivated by established questionnaires (SUS, NASA-TLX). All item scores were aggregated to one scale ranging from 1 (not at all problematic) to 5 (very problematic). Accordingly, the higher the score for a CAPTCHA the more problems in terms of acceptance by users are to be expected.

After a trial with a CAPTCHA that was not used further within the experiment (Gmail CAPTCHA), the experimenter made sure that the participants had understood the experimental procedure. Finally, before beginning the experiments, our subjects went through a practice block for each of the considered CAPTCHAs. The following experimental trials consisted of two main blocks corresponding to the two types of considered CAPTCHAs and were presented in random order. The motion and lightness trials consisted of four sub-blocks that were presented in random order. A complete within-subjects design was used in the study.

Results and conclusion. In general, the results we obtained revealed a positive impression of the newly developed CAPTCHAs. For all of them, the response accuracy was at least 80%, except for the *Lcriss* variant (69%). The motion CAPTCHAs were generally solved reliably more accurately than 80%, however, for the lightness CAPTCHAs it was only the *Lwhite* variant (88%). Response times were significantly shorter than 15 seconds for all newly developed CAPTCHAs (5.2 to 12.4s; $t(28)$=-46.74 to -5.34, p<.001). In particular, the motion CAPTCHAs stand out for their remarkably short response times. What is more, all but one (*Mcirro*: 3.6 out of 5; $t(28)$=4.56, p<.001) of the CAPTCHAs satisfied the criterion of not being rated worse than manageable in terms of subjective difficulty. Finally, all but one (*Mcirro*: 3.8 out of 5; $t(28)$=5.50, p<.001) of our CAPTCHAs did satisfy the criterion of not being rated worse than partly problematic and therefore sufficient in terms of subjective usability. Hence, except for two variants, the proposed CAPTCHAs satisfy the criteria discussed above. The low response accuracy for *Lcriss* appears to be problematic and suggests that this variant should be further developed to meet the criteria. The difficulty of *Mcirro* to satisfy the subjective variables points to potential problems in terms of acceptability. In this paper, we could demonstrate that focusing on the strengths of human visual perception by utilizing perceptual phenomena and visual illusions provides a feasible and viable avenue to suitable and usable CAPTCHAs. Next step will be to attempt to attack our approach using sophisticated machine vision techniques. From our experience, our CAPTCHAs pose severe for standard vision approaches. Currently we are experimenting with rather involved machine learning methods in order to further evaluate the robustness of our approach.

References

1. Chellapilla, K., Simard, P.Y.: Using Machine Learning to Break Visual Human Interaction Proofs (HIPs). In: Proc. Advances in Neural Information Processing Systems, pp. 265–272 (2004)
2. Elson, J., Douceur, J., Saul, A.J.: A CAPTCHA that Exploits Interest-aligned Manual Image Categorization. In: Proc. ACM Conf. on Computer and Communications Security, pp. 366–374 (2007)
3. Datta, R., Li, J., Wang, J.: IMAGINATION: A Robust Image-based CAPTCHA Generation System. In: Proc. Int Conf. on Multimedia, pp. 331–334 (2005)
4. Kolupaev, A., Ogijenko, J.: CAPTCHAs: Humans vs. Bots. IEEE Security & Privacy 6(1), 68–70 (2008)
5. Hoffman, D.: Visual Intelligence. W.W. Norton, NY (1998)
6. Regan, D.: Detection and Discrimination of Motion-defined and Luminance-defined Two-dimensional Form. In: Proc. York Conf. on Spacial Vision in Humans and Robots (1991)

Reflection of a Year Long Model-Driven Business and UI Modeling Development Project

Noi Sukaviriya[1], Senthil Mani[2], and Vibha Sinha[2]

[1] IBM TJ Watson Research Center
[2] IBM India Research Lab
noi@us.ibm.com, {sentmani,vibha.sinha}@in.ibm.com

Abstract. Model-driven software development enables users to specify an application at a high level – a level that better matches problem domain. It also promises the users with better analysis and automation. Our work embarks on two collaborating domains – business process and human interactions – to build an application. Business modeling expresses business operations and flows then creates business flow implementation. Human interaction modeling expresses a UI design, its relationship with business data, logic, and flow, and can generate working UI. This double modeling approach automates the production of a working system with UI and business logic connected. This paper discusses the human aspects of this modeling approach after a year long of building a procurement outsourcing contract application using the approach – the result of which was deployed in December 2008. The paper discusses in multiple areas the happy endings and some heartache. We end with insights on how a model-driven approach could do better for humans in the process.

Keywords: Business process model, model-driven user interface, UI design, solution design.

1 Introduction

Model-driven techniques have been applied in many areas such as software engineering, software architecture, service-oriented architecture, user interface development environments, and recently business process modeling and design. While model-driven techniques enable information organization and downstream automation, their benefits to human are enablement of automated analysis and the expressive power at the level that is closer to the users' conceptual models. In our work, we bring about model-driven benefits in two domains – business process modeling and user interface design.

Model-Driven Business Transformation (MDBT) is the work that our collaborator has been carrying out for a few years prior [1,2] with several successful completed and on-going customer engagements [3,4,5]. MDBT supports business process modeling using an artifact-centric approach [6] – the approach which will be briefly described in the next section. Once a process model is defined, it is transformed into an executable IT solution. While the MDBT framework allows business

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 749–762, 2009.

owners to dictate and essentially specify their IT solutions, it also reduces the time to develop the basic application logic such as business flow and the retrieval and storing of business data. With this approach, what remains time consuming is building solution UIs to go with the business logic and flow that are automatically created.

In 2006, we created a vision that a model-driven UI development environment would reduce the UI production time in the context of business design. Our first goal was to develop a methodology that guides UI solution design to meet business design goals, and continuing to meet changing goals through iterations. Our UI specific goal was to automate the tedious part of the design process while still enabling the UI designer to freely express their design. We want to pass the design integrity from UI designers directly to the solution developers. We want to support design simulation that can be demonstrated live to business stakeholders without implementing the backend business logic. We also want to build a platform which lends itself to reusable assets. While these goals are not fully realized currently, the work and the experience reported in this paper reflects a step towards these goals.

By September 2007, our model-driven UI environment [7] was ready to be tested. Our team at IBM research partnered with an IBM IT consulting team to apply the technology to build a real world application for a business customer. Our task was to build a global procurement application that enabled the process of drafting outsourcing requests, circulating for procurement inspections and approvals, facilitating supplier bidding, supporting multi-tiered reviews, and winner selection, and submitting purchase requisitions. The IBM IT consulting division was to be the recipient of the research technology transfer and would own the solution for the customer after the research team left. We released the pilot solution in December 2008 and the major release is on ramp for May 2009. IBM Research will remain involved in 2009 to assist the consulting team in delivering the second release that includes Asia Pacific countries. This second release would allow us to observe how the current model and tools adapt to additional complex business logic, which will definitely require challenging modifications to the current solution.

2 The Dual Model Approach and Methodology

In our approach, business requirements were gathered at the beginning through business process modeling. Once requirements became sufficiently clear, a storyboard and UI mock-ups were created to confirm and further drill down on the requirements. The implementation phase differs in our approach in that the user interface is explicitly modeled (described in the following sections.) By following the methodology, executable business solution is created as business services. The UI solution is generated as a set of pages with appropriate calls to the generated business services. This work was previously presented in detail in [7.] Figure 1 illustrates key participants in the creation of the business and UI models. The total process involved a larger set of people including business stakeholders, subject matter experts, IT architects, project manager, and developers.

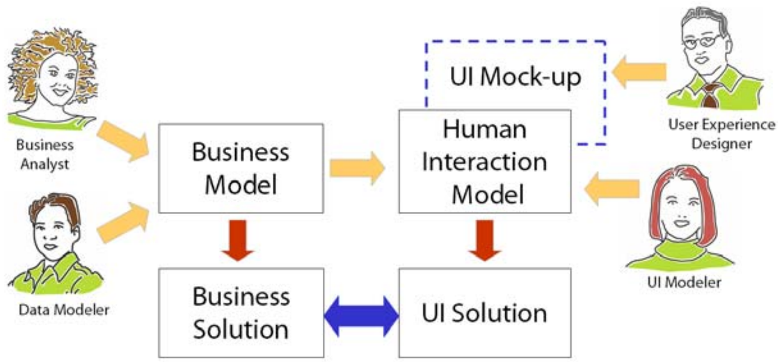


Fig. 1. Participants in the creation of a business model and a human interaction model

2.1 Business Modeling

In our approach, a business engagement starts with business process modeling workshops in which business analysts, subject matter experts, and user experience designers conduct face-to-face meetings to identify key business artifacts and their life cycles. These workshops are essentially requirement gathering workshops. To give the readers a sense of the complexity of the procurement application in this paper, it took about 5 workshops, 3-4 days each, to get a good grasp of the overall business process. The workshops were led by business analysts while the research team helped guide them through the modeling exercise.

The business process model was defined using a methodology called Artifact-centric modeling [3], of which focus is on modeling the life cycles of key business artifacts – artifacts that are the heart of a business. The methodology leads to a clean way of thinking about business operations without IT interference – by enforcing the thinking only around what happens to these key business artifacts. The use of artifacts as nouns and business tasks as the actions to the nouns naturally shakes the model free of IT-specific implementation such as button clicks or how data are navigated. The focus on key artifacts also helps avoid small tasks on non-key artifacts that may be specific to how a particular organization operates. The full life cycle of the "Fixed Price Request" for our application was drawn in several diagrams. The application had two main artifacts - Fixed Price Request and Supplier Response. Figure 2 shows a snippet of the "Fixed Price Request" artifact life cycle. "Create" and "Update" are tasks and the users who perform the tasks are shown on top of the tasks. Though the artifact life cycle diagrams may appear visually similar to other process diagrams, the differentiations as mentioned is in the task semantics and granularity.

As part of completing the business model, the data aspect of each artifact, as shown in Figure 3, is also filled in by business analysts and/or data modelers. This information is very essential for bridging the business model to the user interface model later on. We currently use IBM Websphere Business Modeler (WBM) as a front end tool to model business flow and the data model.

Our application has lots of requirements on when and whether certain user roles can create, view or edit the artifact data. The business team did not see how such requirements could be captured in the business diagramming tool. Complimentarily,

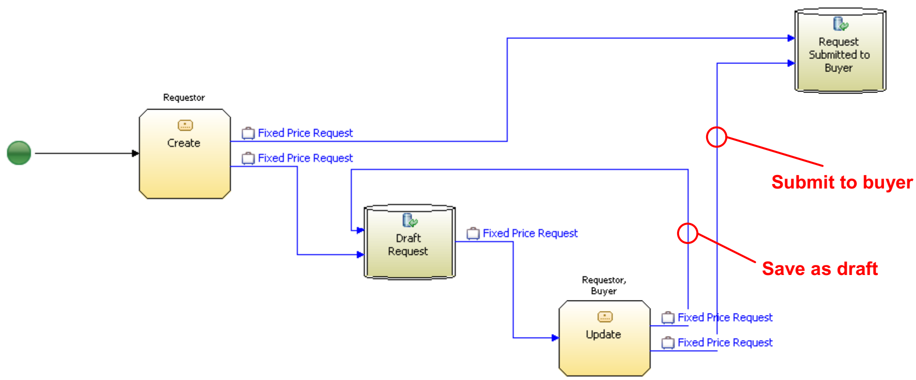


Fig. 2. An example of an artifact-centric business process model. The output ports of each task are labeled but not visible on the diagram.

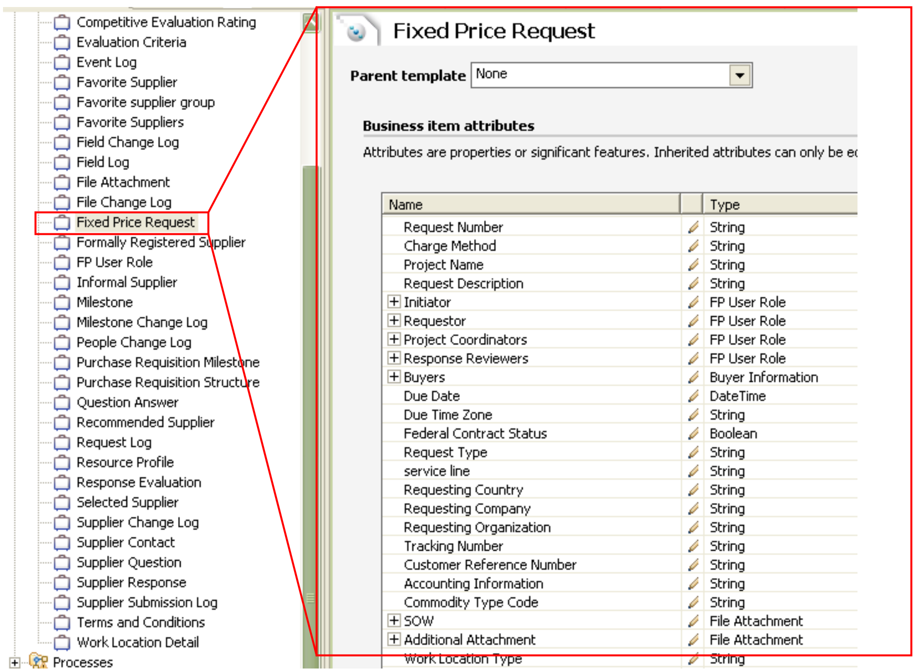


Fig. 3. An example of the information model defined in the business model

a spreadsheet is used to capture these requirements. These readability and edit ability details of each business task from the business model were captured on an individual worksheet in the spreadsheet. Figure 4 shows an example of the read/write requirements for a task. Additional detailed descriptions of business tasks were kept in a Microsoft Word document, again, as the tool does not facilitate easy navigation back and forth among task descriptions and appeared to poorly support the conversation flow.

1	TASK	Data Attribute	Created in Task ?	Optional / Mandatory					Read / Write					
2				"Reject"	"Return to RPIC for Additional Info"	"Submit to Suppliers"	"Save"	"Submit Fast Path"	I	R	PC	Rev	Buyer	Supp
3	Review & Update	Running Comments	Yes	Optional	Optional	Optional	Optional	Optional	W	W	W	NA	W	NA
4	Review & Update	Request Number	No	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	R	R	R	R	R	NA
5	Review & Update	Requester	No	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	W	W	W	R	W	NA
6	Create	Buyer(s)	Yes	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	TBD	TBD	TBD	TBD	TBD	TBD
7	Create	Buyer(s) type	Yes	Optional	Mandatory	Referenced	R	R	R	R	NA	NA	R	
8	Review & Update	Initiator	No	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	R	R	R	R	R	R
9	Review & Update	Project Name	Yes	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	R	R	R	R	W	NA
10	Review & Update	Request Description	Yes	Optional	Optional	Optional	Optional	Optional	R	R	R	R	W	NA
11	Review & Update	Due Date	Yes	Manadatory	Manadatory	Manadatory	Manadatory	Optional	R	R	R	R	W	NA
12	Review & Update	Government or Federal Contract ?	No	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	R	R	R	R	W	NA
13	Review & Update	RFI or RFQ ?	No	Mandatory	Mandatory	Mandatory	Mandatory	Optional	R	R	R	R	W	NA

Fig. 4. A spreadsheet business users specified for the task "review and update"

2.2 User Interface Design

Once the requirements became sufficiently clear, a storyboard and the UI mock ups were created. There is nothing very new in this process, except the user experience designer in our team needed to be totally aware of the business process flow and the artifact data model.

A couple of points are worth mentioning here. First, our intention was to use the UI model to mock-up the UI and present it to the business team. However, the lack of a visual tool in our toolset prevented this exercise from happening in a timely fashion, hence this idea was dropped. Instead, we used a static PowerPoint storyboard to present the UI mock-ups. The storyboard provided synopses of what would be months of activities by multiple people (request drafting, the review process, supplier bid revisions, cancellations, exceptions, contract renewal, etc.) into hundreds of charts. Several multi-hour sessions were spent validating the UI with the subject matter experts. We found it still very hard to follow the storyboard for an application of this size with many variations. An interactive storyboard which allows use cases to be selected and followed would be more appropriate.

2.3 User Interface Modeling

After the subject matter experts validated the initial UI design, the UI modeling process began. We used our custom UI modeling tool to automatically extract user roles, tasks for each role, artifact details, and artifacts each role is entitled to access from the business model. The tool presented all of this information in the Human Interaction perspective. The technical detail of this tool was presented in detail in [7.] Let us briefly summarize here for the context of this paper.

The UI modeler had 3 tasks. The first task, which is optional, is to model the read/write access for each user role and for each specific task or group of tasks. This task is primarily encoding the spreadsheet inputs from the business team into a more optimized form in the tool. The tool allows added efficiency such as allowing specifications of similar access requirements across user roles and tasks. The specifications could have been done by business analysts but since the tool requires an installation of Rational Software Architect with a large footprint, the business analysts were more than happy to have this work done by the UI modeler.

The second task for the UI modeler was to model screens as pages associated with each user role. Each page was modeled as a hierarchy of page fragments, each of which contained either further fragments, or a layout and UI elements. While modeling each page, the UI modeler associated the UI elements to data attributes from the artifact data model. Alternatively, she can use the information model to automate filling in the content. For example, a set of data attributes can be associated to a UI container, and the system will automatically fill the container with appropriate UI elements. Had the modeler done the first task, the automatic generation can be more accurate in selecting appropriate UI elements for viewing or editing. These elements can be altered later if the modeler wishes. The connections from UI elements to data attributes are kept in the model hence allowed the system to generate appropriate service calls to fetch or store data at run-time.

Lastly, the UI modeler defined page flow as links between pages. Some links maybe coupled with a side effect of saving or removing data. Some may couple with outputs of business tasks (i.e. “Approve” output of a “Review” task) hence signifies business logic connections. The page content and flow, with connections to business data and logic, enables code generation to construct appropriate web service calls that fill the screen with business data and enact appropriate business logic.

2.4 Simulating the UI Design

With the UI model, UI code can be generated to render pages without the backend business logic. In our current implementation, the output from the UI model is an XML model feed to IBM Websphere Portlet Factory (an IBM product), which in turns generated JavaServer Page (JSP) code. When running in a simulation mode, our system pulled sample data from XML-based data sets. The data are selected based on the semantic types enhanced to the data model by the UI modeler [7.] This feature was designed to enable UI viewing without business logic and can be used to show low-fidelity UI to business stakeholders. It turned out that this feature was not sufficiently designed for its job. The sample data feed became confusing quickly after viewing a few pages since there was no continuity from one page to the next. Secondly, most of the use cases in our application were rather lengthy with many possible variations. The simulation could not present these variations in a way that made sense as it did not have application-specific data that would trigger the right behavior. We later implemented support for adding application-specific values to the sample data set. However, this was still insufficient as dynamic UI behaviors such as turning on/off portions of the screen or enabling/disabling buttons often depended on specific values. When these values were randomly picked, the dynamic behaviors became confusing. We need to enhance this simulation further to make it truly useful for business users. Positively though, the feature was heavily used by the UI modeler before we started backend integration as a way to see whether the basic layout and flow worked as planned. The UI modeler made a good use of this feature.

2.5 Finishing Up the Solution

For the finale, we pulled together the generated outcomes from the business model and the UI model. Two products were generated from the business model – DB2 database definitions generated from the artifact data model, and business services based on the outputs of business tasks in the model. The database definitions were manually enhanced by the database administrator before they were used to generate the actual database tables. The business flow transformed into a business state machine and the business services were merely API for invoking state transitions. We will not dive into further technical details since it is not the interest of this community.

Additional development work was needed in this step to complete the solution. Business logic that was not based on flow need to be implemented, for example, routing a request to appropriate procurement teams, sending e-mail notifications to various parties when an event happened, calculations of some data values based on page inputs, etc. Also, since our UI modeling and generation is currently limited to static screens, dynamic behaviors needed to be programmed in JavaScript thus required additional UI developers. This is one area of our future work.

3 Related Work

Many business modeling languages exist such as UML, BPEL, CogNIAM, IDEF0, XPDL, [8,9,10,11,12] but one that is becoming a standard is BPMN [13]. While these languages provide variations of expressions and notations, with them come specific methodologies the languages enforce. The model-driven business modeling approach used in this work is rather unique in that the methodology coincides well with business user thinking, while the system can generate executable code. We do not intend to dive into related work in business modeling research here since we want to focus on the human interaction and human aspects of this work.

Many aspects of our UI modeling work have related work in the area of model-driven user interface environments. The use of task-driven modeling as a way to tie to UI model was reported in [14, 15, 16, 17, 18.] However, these research systems depend on the task and task hierarchy defined using the traditional HCI method [19.] The UI Pilot system [20] in particular was very close to our UI modeling environment in that it provided a 360-view of task, UI pages, users, and data components. Our work differed in that we extract such information from the business model and it remains so throughout the development cycle. In our methodology, UI modelers do not dictate the basis of the business tasks against which UI is to be designed. Some other related model-driven UI model and environments can be found in [21, 22, 23, 24.]

Our UI tool has an assistance that uses selected data attributes to fill in UI elements automatically. This work is similar to the work in the past in [25] though we use the semantic data types that are more familiar to business users (such as address, first name, last name, social security number, etc.) to drive UI element selection. We do not concern with a comprehensive automatic layout compared to [26.] Lastly, the specification of read/write access in the data model uses a similar concept as [27] with a conceptually similar hierarchy of read/write access. Again, our difference is the heavy tie to the business vocabulary and such a perspective it brings and enforces.

Lastly, new research work is emerging that ties business process modeling to UI modeling [28, 29.] We see our work as being rather similar though our work is guided by a business modeling methodology that prevents IT concerns or particular choices of interface design, hence the business model remains agile as UI design may change over time.

4 Analysis of the Methodology and the Process

In this section we discuss features that worked well and those needing improvements.

4.1 Business Model and UI Modeling Methodology Contains Solution Deviations

Business stakeholders found our approach extremely valuable. They resonated with the idea that the business model, which they helped create and refine throughout its evolution, provided a control over the solution. Business users can easily grasp the overview of the end-to-end process. The process remained at a level high enough for business users to identify where in the process that required modifications.

On one hand, the process model remained agile and all participants were comfortable with understanding the model implications. On the other hand, more detailed descriptions of what entailed the high-level tasks were captured deeper in the model. This was not seen by the business users. With business model as the original point of development and it remained so throughout the development process, and with the UI model integrated tightly with the business elements, many good behaviors followed. We observed that all team members talked in the same language – that was the language of the business elements. The artifact and attribute names given in the business model were used in the UI design language as well as by developers who programmed in cognizant of the same vocabulary. Business tasks were referenced all aboard. When a data attribute or a function was missing, it was easily traced back to its absence from the business model. Business analysts and UI design were then engaged to resolve the problems by changing the business and/or the UI models.

Interestingly, database administrators on the contrary requested the ability to map selected business data attribute names to other labels – specifically the table and column names in their master reference databases implemented prior to this application. The prompted reason was to reduce database personnel training since they were already trained to maintain these databases across applications. The database team members were aware of the mappings from the business vocabulary to their familiar database vocabulary, but this mapping had no effects on developers.

It was still possible to have communication breakdowns and wrong information entered in lower-level model. With a tight integration from the UI model components to the business model components, we had a control over the solution scope. All user roles were derived from the business model. Pages were assigned to appropriate user roles. The tool constrained each page flow to only link to tasks entitled to the user role owner of the page. On one hand, this did not leave room for misconducts in the solution behavior. On the other hand, a missing behavior could not be quickly added. A real case was, for example, after a supplier declined terms and conditions of a bid, they could not reverse their decision hence were out of the bidding opportunity.

The business team believed the implementation was simply missing. Upon inspecting the business model, it was clear that the intention was not recorded and probably never discussed. We noticed many cases such as this one occasionally in the process. This is one of those areas where, we believe, requirements for exceptions, as obvious as they may seem, often fall through the crack. The business model was a recorded agreement and a means to make it rather simple to find the missing requirements.

Our current approach is not free from errors such as communication breakdown. However, as seen over this project, the approach succeeded in limiting solution deviations in many directions as well as providing evidence of logical connections even if they were wrong. We continue to conduct research to better capture requirements, especially dynamic ones, in the context of artifact-centric modeling.

4.2 Iterations Must Work Well in Model-Driven Environments

A model-driven environment means iterations are done in high-level models. Iterations must be well supported and must be designed in the tooling from the beginning. With our approach, we have observed iterations in operation all year long. It worked well in many areas and it ceased to work in some areas.

The UI modeling tool from the beginning had support for business model changes designed into the tool [30.] Since it must always be in synchronous with the business model, we made explicit decisions on how the UI tool would deal with different types of business model changes. We considered the tool working rather well. Throughout the development, the modelers and developers went through 2-3 iterations a day per UI model for a period of several months. The iteration was not as fast as we wished. Each version of the generated code needed to be recompiled and reassembled. With the industrial software, the type of server we deployed to, and the sheer size of the application, the time it took was unavoidable.

On the contrary, iterations on the business modeling did not fare as well towards the end. The intermediate UML solution model generated from the business model required additional enhancements before the code could be generated from it. The problem we had was two-fold. The tooling support for making such enhancements was not usable hence ignored by developers. Secondly, change management was not accounted for. This turned out to be a backlash later on in the development process. Every time the business model changed, the generated solution would wipe out the enhancements. Putting changes back would go from 1 day when the solution was still small to eventually more than 5 days as the implementation became larger. When the project schedule became tight at the end, we were forced to stop changing the business model, thus erasing the iteration benefits of a model-driven approach. Some key application problems cannot be resolved. The lesson learned here was that change management must be designed for and not treated as an accidental feature.

4.3 Business Users Cannot Express All That Are Required

As mentioned earlier, our intention is to have business analysts and the business team owning the business process. Though our business modeling approach has been proven for a few years, business users in previous engagements were not active in handling the business process model with WBM. We attempted to achieve this goal early on in this project. Two observations we made in this endeavor. The ownership of the model slowly shifted towards the research team as the project got deeper in the implementation. This is understandable as making a business model deliver the correct specifications of business logic, and completing the information model details needed at run time is a rather daunting task. For example, in addition to basic data attributes given by business analysts, the information model was enhanced with unique keys, extraneous attributes and flags for system functioning. Implementing UI also required additional attributes for the overall usability that were not called for by business analysts. When the appearance of the data model started to vary significantly from the spreadsheets, the business team ceased their interest in the data model in WBM and only referred to the spreadsheets as the data reference.

One concept that was difficult for business users was data structure and modularity. Business users can understand data multiplicity (and some modularity) as in "there needs to be one or more milestones in the request before sending bids to suppliers" and optional/mandatory conditions as in "there can be multiple resource profiles but none is required." Translating these requirements into 1..n or 0..n is not a big hurdle. However, structuring information into a module that can be reused was not natural to them. This is understandable. However, our approach mixes the need to optimize the database tables and the need for business users to enter and update the information model in the same tool. These two goals cannot be achieved by business users alone.

The business team also used less of flow model later on. When decisions to change model were made, they totally entrusted the research team to execute the changes and only occasionally checked the flow during discussions. We want to believe that they at this point were very familiar with the high-level model. Interestingly, the UI mock-ups became a more often used reference by the business team as we got deeper into the project. And that did not change even towards the end of the project.

4.4 Lack of a Visual Tool Changes the Nature of the Development Process

When we embarked on this research, we made a conscious decision not to build a visual editor. Our first priority was to prove the modeling concept and also we could not afford the time or programming efforts. This turned out to be a bad decision. First, the user experience designer did not want to and did not build the UI design directly in the tool; hence the design precision was not captured in the model as intended. We missed our goal in this project.

Lack of a visual tool had more ripple effects as it has proven to cause serious time consumption as well. UI modelers missed out on design precision as any developers would. Lack of a visual tool forced the modeler to work with text such as page names, UI element names, etc. Visually searching for a page was hard as we reached over a hundred pages. Page names were hard to remember even for the modelers themselves. Textual names also created personal workspace effects that made it hard for others to view the UI model.

Lastly, the IT consulting team viewed a tool without a visual editor as a programming tool; developers were assigned to own the UI model. 3 developers later on in the process took over the UI modeling as part of the technology transfer. We started seeing design deviations as in the traditional method. Most of all, we have experienced mismatches in the tool design concepts that were intended for other types of users.

4.5 UI Modelers Need to Understand Business Requirements and the Data

In our project, the user experience designer was involved with the design process from the beginning. However, the UI modeler was not. When UI pages were given to the UI modeler, the experience designer assumed it would be obvious to see the matching from the field labels to the attribute names in the data model. That was true perhaps about 80% of the time. Minor questions arose on data clarifications between the UI modeler and the experience designer. However, we had many pages that appeared similar using similar field labels but in fact needed data attributes from different sections of the model. We had many of those mistakes that were not detected earlier on but once detected, were already widespread and tedious to correct.

The connections from correct business data to correct pages had proven to be essential in creating UI pages that worked. However, using the data model with mixed initiatives (as mentioned in 4.3) can be quite technical and this task may no longer be appropriate to user experience designers. We need to further investigate whether the solution to this problem lies in a better data modeling tool that can support mixed purposes or in a better UI modeling tool.

4.6 Monolithic Model Prevents Parallel Development

When we designed the UI modeling tool, we did not think about multiple modelers in the same project. We were shortsighted. The outsourcing application reached over 200 pages across 4 user roles. At the beginning we decided to use two UI models to support two different style templates (internal procurement users vs. external users.) The separation did not suffice. As we transferred the UI modeling task to the consulting team, the task was assigned such that a modeler owned a particular user role. By storing the UI model in one big file, only one modeler could work on the model at a time. (Versioning tools could not resolve conflicts at the UI modeling level.) We temporarily separated the model further along the owner of the model but this solution was far from ideal. We learned that collaboration must be designed into the tool from the beginning, as it has many impacts on how the model is structured, the choice of modeling language platform, and the methodology of how team members would collaborate.

4.7 Who Should Be UI Modelers?

When we started the project, we targeted user experience designers as the users of the UI modeling tool. However, lack of a visual tool took us off our intention. The year of experience had informed us however that our vision may be questionable. UI design skills are rare in IT consulting team and a project may go without such skills. What do we do? Connecting UI to the data model requires data modeling skills. Connection to business flow and the overall big picture requires business savvies. Designing quality UI requires visual and design sensibility. The lessons learned here is no one possesses such a range of skills. The modeling tool should look for opportunities to take advantage of the available skills while increasing automation to compensate for missing skills. Alternatively, a modeling tool should facilitate handing off from one user to the next, enabling multiple users with different skills to work together even in a single domain of UI modeling.

4.8 Generate What Matches Market Skills

Lastly, one of the complaints we received from the consulting team was the fact the UI code was generated as Websphere Portlet Factory (WPF) XML. This is not the skills that prevail in the market. Developers were not comfortable with the generated code which they did not understand. Moreover, when the generated code needed to be enhanced for debugging, developers were forced to learn the WPF programming environment. The consulting team needed to learn new skills to own the tooling. We have been requested to consider generating UI code such as JSPs and HTML to better match developer skills both as users and future developers of the tool.

5 Summary of Analysis

While we have no quantitative data, our IT consulting team had agreed that the project used much less developers and could deliver much more functionality within the same amount of time, in comparison with a traditional development method. We continued to allow business requirement changes very late in the development process, though not within the final months. We have seen positive attitude from business users who benefited from the ability to fine-tune the design after they had seen the running prototype. The design phase definitely had gone much longer in the development cycle. Had the business modeling iterations were possible through the end, business model changes would continue into the last months of delivery. This would not happen in the traditional development method.

Tooling design had great effects on the success of a model-driven project. We have seen concepts that worked well such as support for iterations, ease of technology transfer to practitioners with the UI tool, convergent to the same business vocabulary by team members, reduction in development efforts and time. We also had seen concepts that did not get sufficient considerations upfront that caused issues later, for example, sample data simulation, collaboration among modelers, and generated code that does not match the user skill sets. Throughout the year, our tool and the meta-model have gone through significant changes within the boundary that the project can afford. Major conceptual changes cannot be done during the project if we could not continue with the existing model. Lots of progress was made during the year to make the tool more usable both for the development process and for what it generates. One area, for example, which we did not discuss, was generating UI code that meets the accessibility requirements.

We have learned a great deal from the success and failure of our choices in the research tool design. Our model-driven development process is very promising. It provides users, with or without programming skills, the ability to look into the business model and/or the UI model to grasp the solution. It enables them to have a direct control over the solution. This has become a very positive experience from our engagement. A year of experience has changed the way we see the impact a model-driven approach have on practitioners. The dual modeling approach has impacts on a variety of practitioners ranging from business users, UI designers, to developers. Many of these issues were not in the scope of academic interests at the beginning. However, when one considers the absence of model-driven development tools in the marketplace, especially in the UI area, some of these lessons we learned we hope to inspire the readers to think differently on the impact of their work to the real world.

Acknowledgements. We would like to thank the extended teams who have been working with us for almost 2 years – the procurement business team Mark Kemp, Jim Carver, Cathy Ng-Pepe, Amelia Carruth, and Russell Parks; the IT consulting team Gautam Majumdar, Glenn Godoy, Marco Gomez, Sachin Popli, Manish Singh, and Raul Laprida. We would like to thank our own development team – Thejaswini Ramachandra, Ajoy Acharyya, Arindum Dutta, Sourav Chakrabarty, and Kalyani Deshpande for the untiring efforts; and the MDBT architect team Prabir Nandi, Terry Heath, and Florian Pinel. Finally, we would like to thank Anil Nigam for his consultant on artifact-centric modeling throughout the project.

References

1. Kumaran, S.: Model-Driven Enterprise. In: Proceedings of the Global EAI (Enterprise Architecture Integration) Summit, pp. 166–180 (2004)
2. Kumaran, S., Nandi, P.: Adaptive Business Objects: A New Component Model for Business Integration. In: Proceedings of ICEIS 2005: 7th International Conference on Enterprise Information Systems (2005)
3. Bhattacharya, K., Guttman, R., Lyman, K., Heath III, F.F., Kumaran, S., Nandi, P., Wu, F.Y., Athman, P., Freiberg, C., Johannse, L., Staudt, A.: A Model-Driven Approach to Industrializing Discovery Processes in Pharmaceutical Research. IBM System Journal 44(1), 145–162 (2005)
4. Bhattacharya, K., Caswell, N.S., Kumaran, S., Nigam, A., Wu, F.Y.: Artifact-centered operational modeling: Lessons Learned from Customer Engagements. IBM Systems Journal 46(4) (2007)
5. Liu, R., Bhattacharya, K., Wu, F.Y.: Modeling Business Contexture and Behavior Using Business Artifacts. In: Krogstie, J., Opdahl, A.L., Sindre, G. (eds.) CAiSE 2007 and WES 2007. LNCS, vol. 4495, pp. 324–339. Springer, Heidelberg (2007)
6. Nigam, A., Caswell, N.: Business Artifacts: An Appraoch to Operational Specification. IBM Systems Journal 42(3), 428–445 (2003)
7. Sukaviriya, N., Sinha, V., Ramachandra, T., Mani, S., Stolze, M.: User-Centered Design and Business Process Modeling: Cross Road in Rapid Prototyping Tools. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4662, pp. 165–178. Springer, Heidelberg (2007)
8. Unified Modeling Language (UML), Version 2.1.2, `http://www.uml.org/`
9. Business Process Execution Language for Web Services, `http://www.ibm.com/developerworks/library/specification/ws-bpel/`
10. COGNIam, `http://www.pna-consulting.nl/`
11. IDEF0: Function Modeling Method, `http://www.idef.com/idef0.html`
12. XPDL, `http://www.wfmc.org/XPDL`

13. Object Management Group's Information on Business Process Modeling Notations, http://www.bpmn.org/
14. Paterno, F., Mancini, C.: Model-based Design of Interactive Applications. ACM Intelligence Magazine Winter, 26–37 (2000)
15. Paterno, F., Santoro, C.: One Model, many Interfaces. In: Proceedings of 4th International Conference on Computer-Aided Design of user Interfaces CADUI 2002, pp. 143–154 (2002)
16. Paterno, F.: Tools for Task Modeling: Where We are, Where We are Headed. In: Proceedings of International Workshop on TAsk MOdels and DIAgrams for user interface design TAMODIA 2002, pp. 10–17 (2002)
17. Puerta, A., Maulsby, D.: Management of Interface Design Knowledge in MOBI-D. In: Proceedings of 2nd International Conference on Intelligent User Interfaces, pp. 249–252 (1997)
18. Puerta, A., Cheng, E., Ou, T., Min, J.M.: User-centered Interface Building. In: Proceedings of CHI 1999: ACM Conference on Human Factors in Computing Systems, pp. 426–433 (1999)
19. Mayhew, D.: The Usability Engineering Lifecycle: A Practitioner's Handbook for User Interface Design, 1st edn. Morgan Kauffmann, San Francisco (1999)
20. Puerta, A., Micheletti, M., Mak, A.: The UI Pilot: A Model-based Tool to Guide Early Interface Design. In: Proceedings of the ACM Intelligent User Interface 2005, pp. 215–222 (2005)
21. Nunes, N.J., Cunha, J.F.: Towards a UML Profile for Interaction Design: the Wisdom Approach. In: Evans, A., Kent, S., Selic, B. (eds.) UML 2000. LNCS, vol. 1939, pp. 101–116. Springer, Heidelberg (2000)
22. Griffiths, T., Barclay, P.J., Paton, M.J., Paton, N.W., Gray, P.D., Kennedy, J., Cooper, R., Goble, C.A., West, A., Smyth, M.: Teallach: A Model-Based User Interface Development Environment for Object Databases. Interacting with Computers 14(1), 31–68 (2001)
23. Griffiths, T., Barclay, P.J., Paton, M.J., Paton, N.W., Gray, P.D., Kennedy, J., Cooper, R., Goble, C.A., West, A., Smyth, M.: Teallach: A Model-Based User Interface Development Environment for Object Databases. Interacting with Computers 14(1), 31–68 (2001)
24. Bouillon, L., Vanderdonckt, J., Chow, K.C.: Flexible Re-engineering of Web Sties. In: Proceedings of ACM Conference on Intelligent User Interfaces IUI 2004, pp. 132–139 (2004)
25. de Baar, D., Foley, J.D., Mullet, K.E.: Coupling Application Design and User Interface Design. In: Proceedings of CHI 1992: ACM Conference on Human Factors in Computing Systems, pp. 259–266 (1992)
26. Kim, W.C., Foley, J.D.: Providing High-level Control and Expert Assistance in the User Interface Presentation Design. In: Proceedings of CHI 1993: ACM Conference on Human Factors in Computing Systems, pp. 430–473 (1993)
27. Penicht, V.M.R., Paterno, F., Gallud, J.A., Lozano, M.D.: Collaborative Social Structures and Task Modeling Integration. In: Doherty, G., Blandford, A. (eds.) DSVIS 2006. LNCS, vol. 4323, pp. 67–80. Springer, Heidelberg (2007)
28. Sousa, K., Mendonca, H., Vanderdonckt, J.: User Interface Derivation from Business Processes: A Model-Driven Approach for Organizational Engineering. In: Proceedings of ACM Symposium on Applied Computing, pp. 553–560 (2008)
29. Sousa, K., Mendonca, H., Vanderdonckt, J.: Addressing the Impact of Business Process Changes on Software User Interfaces. In: Proceedings of the 3rd IEEE/IFIP International Workshop on Business-Driven IT Management BDIM 2008, pp. 11–20 (2008)
30. Sukaviriya, N., Sinha, V., Ramachandra, T., Mani, S.: Model-Driven Approach for Managing Human Interface Design Life Cycle. In: Engels, G., Opdyke, B., Schmidt, D.C., Weil, F. (eds.) MODELS 2007. LNCS, vol. 4735, pp. 226–240. Springer, Heidelberg (2007)

Designing Tools for Supporting User Decision-Making in e-Commerce

Alistair Sutcliffe and Faisal Al-Qaed

Manchester Business School
University of Manchester,
Booth Street West
Manchester M15 3 BP
UK

Abstract. The paper describes a set of tools designed to support a variety of user decision-making strategies. The tools are complemented by an online advisor so they can be adapted to different domains and users can be guided to adopt appropriate tools for different choices in e-commerce, e.g. purchasing high-value products, exploring product fit to users' needs, or selecting products which satisfy requirements. The tools range from simple recommenders to decision support by interactive querying and comparison matrices. They were evaluated in a scenario-based experiment which varied the users' task and motivation, with and without an advisor agent. The results show the tools and advisor were effective in supporting users and agreed with the predictions of ADM (adaptive decision making) theory, on which the design of the tools was based.

Keywords: Decision support, e-commerce, recommenders, advisor agents.

1 Introduction

Users face many different choices and tasks when shopping on the Internet, although support tools tend to fall into only two categories: ranked list matching tools as exemplified by the Google search tool, and recommenders as found in Amazon.com's website. Unfortunately few systems provide more than one tool so users' choice is restricted; furthermore, the available tools rarely match users' natural strategies. Since one of the principles of HCI design is to create interfaces which match the user's task and natural way of working [1, 2], design of e-commerce decision-support tools appears to be somewhat limited.

Decision-support tools range from recommenders [3] to comparison matrix tools and interactive visualisation-based tools for exploring choices [4]. Each tool assumes that a particular user strategy will be adopted in decision-making and search tasks; however, users employ multiple decision-making strategies which they mix in a flexible manner [5]. The conventional user-centred design approach is to specify a task model and then base the interface and dialogue design on the task model. When faced with multiple strategies the usual response is to design configurable or adaptable

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 763–776, 2009.

interfaces [2, 6, 7]. However, when user strategies become nearly as complex as the task itself, the adaptation/adaptive design approach becomes difficult to sustain. An alternative response is to adopt a toolbox approach to design, as found in graphical drawing toolkits which have discrete tools for lines, shapes, textures, etc. that can be combined in a flexible manner. In previous work we developed an advisor agent to help users make decisions with a single tool based on Shneiderman's alpha slider design, and demonstrated that the advice was effective [8]. In this paper we explore the multiple tool paradigm with advisor agents for interactive decision support in e-commerce tasks.

A secondary motivation for this research is to explore the effectiveness of task-based advice. We investigate decision-making advice based on strategies from the Adaptive Decision Making (ADM) theory [9] with online advisors to help users not only to follow the appropriate cognitive strategy but also select effective tool support for their strategies. Since users tend to access help and explanation systems when they perceive the task to be difficult, we investigate the effectiveness of avatar agents for delivering advice.

The paper is organised as follows: the next section provides a brief background of related research, then the design of the decision-support toolkit and advisor agent is described. This is followed by explanation of the design, methods and results of the evaluation experiment which compared use of the toolkit with and without the advisor agent. The discussion reviews the experimental findings and concludes with plans for future research.

2 Related Work

Personalisation and recommender systems have tailored choices to users by modelling individual consumers' preferences [7, 10-12], one of the most popular being Amazon.com [13] which provides personalised recommendations according to consumers' profiles, ratings of particular products and purchase histories. The SETA toolkit supports creation of e-shops and recommends products best fitting users' needs as well as adapting the layout of store catalogues to user preferences and expertise [14-15], while SYSKILL and WEBERT [16] recommend web pages from users' feedback on viewed pages which the system uses as training data for learning user preferences. Recommenders have been a popular approach to help consumer choice [17], although Shneiderman has regarded decision making as an interactive search process supported by iterative querying and visualisation [4, 18]. Haubl and Trifts [19] argue that people follow a two-stage decision-making process, noting that while recommenders can support the initial screening of alternatives, comparison matrix tools are necessary for in-depth evaluation of options to help users make actual decisions. Pereira [20] investigated the interaction between the search strategy and users' product knowledge, using two strategies from ADM theory, Elimination By Aspects (EBA) for initial filtering and the Weighted ADDition method (WADD) for comparison and trade-off analysis, supported by personalised recommenders and hypertext for browsing alternatives. Subjects with little product knowledge preferred recommenders whereas subjects with high product knowledge preferred WADD and EBA strategies. Although user needs and decision strategies have been considered, few adaptable decision support tools have been produced for e-commerce.

3 Toolkit and System Architecture

A two-tier client server architecture was adopted (see figure 1) with the client containing eight tools supporting different browsing and search functions, a configuration user interface which captures data for the tool recommender, and the advisor agent.

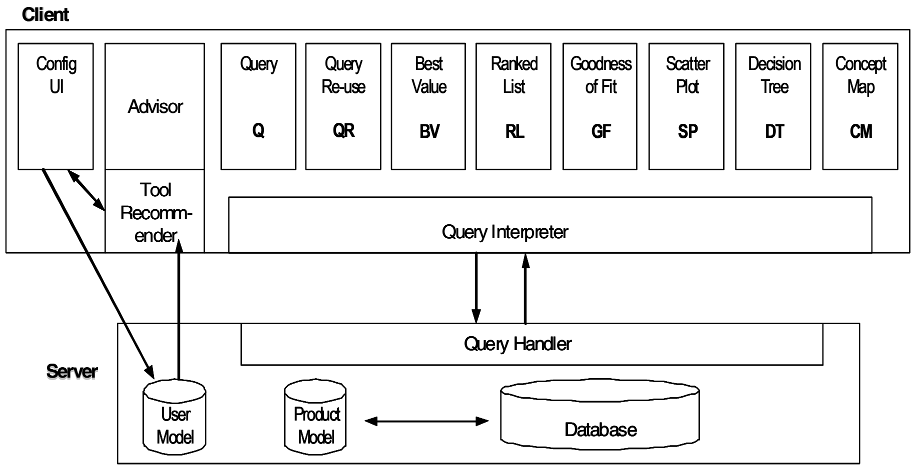


Fig. 1. System architecture of the Adaptive Decision Support System (ADSS)

The system design was informed by Adaptive Decision Making Theory [5] which considers the decision-making strategies that people should adopt according to the decision-making task, their knowledge of the domain, and contextual constraints such as time and motivation. In more complex tasks and domains compensatory (trade-off) strategies are predicted, while in less complex domains non-compensatory strategies are adopted, e.g. filter on single attribute, rely on memory of previous choices, etc. Three tools support non-compensatory search strategies with minimal user interfaces so these are not described in further detail: a template-based keyword search for products, a query reuse tool which allows previous queries to be saved and reused by picking from a menu list, and a recommender which selects the best value options from a range of pre-set queries. Non-compensatory strategies focus decisions on a single attribute choice (e.g. cost) without any trade-off decisions. Three tools support compensatory strategies with slider controls so users can interactively assess different range-value setting in queries with different feedback displays (ranked list, goodness-of-fit and scatterplot); and the final two tools support browsing-style exploration (decision tree and concept maps). Examples of compensatory strategies in ADM theory are filtering choices (Elimination by Aspects) and WADD weighted addition decisions which trade off different attribute values. We implemented the eight tools to cover the diversity of decision-making strategies described in ADM theory and with the application scenario that ADSS would act as an Advisor plug that could be configured for several different e-commerce sites, so it had to adapt to different product domains as well as user constraints and knowledge. All the tools, the query interpreter and the tool recommender were implemented as JavaAplets and HTML. The advisor agent was developed using the Guile 3D studio agent tool to script the advice dialogue.

The server side houses the application database, the user model populated by the configuration dialogue and the product model which contains meta-data on the application database. A query handler communicates with the client-side query interpreters, calling database search functions implemented as a MySQL server. Server-side functions were implemented as Java Servlets hosted on Tomcat Server 3.5.

3.1 Tool Descriptions

Three tools share the slider-based query controls to support more compensatory, value-range exploration of choices, based on the alpha slider paradigm [4]. The ranked list tool illustrated in figure 2 implements slider-based querying with user controls for ranking the results for each variable.

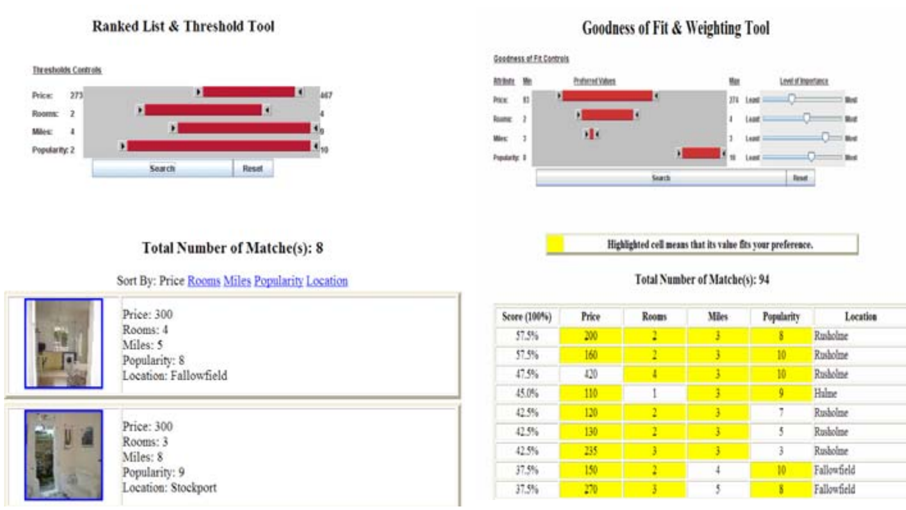


Fig. 2. Ranked list tool (left) and Goodness-of-fit tool (right)

The tool can be configured with sliders for any numeric attributes in the dataset with a simple palette picking list.

The goodness-of-fit tool (figure 2) augments the ranked list design by post-processing exact-match and near-miss query results according to how closely they fitted the search criteria. The search criteria can be biased by user-configurable weights so different rankings can be produced reflecting search-variable importance as well as closeness-of-fit to search parameters.

The scatterplot tool adds a dynamically updated display to the functionality of the basic ranked list tool (see figure 3) so the location of the search results can be inspected on a map. Results are also displayed in table format and can be sorted by clicking on

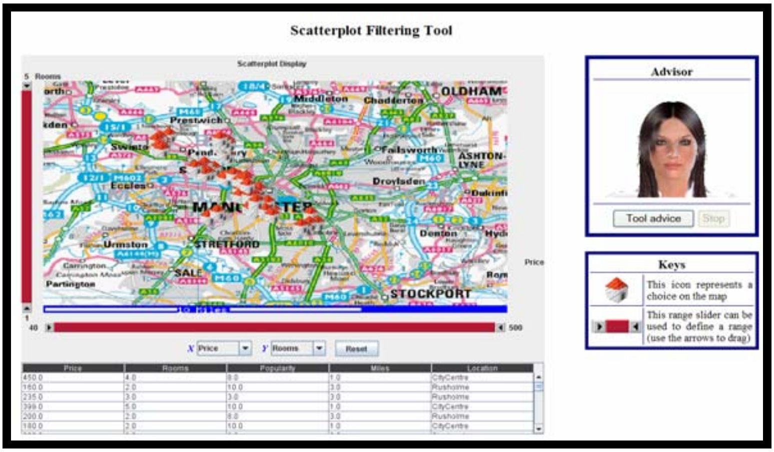


Fig. 3. Scatterplot tool showing the map configuration

the column headings. As with the other tools, the scatterplot is configurable, so any attributes can be selected for the sliders. If non-spatial attributes are chosen a scatterplot graph display is used to show the results within value coordinate space, following Shneiderman's starfield design.

Two tools support exploratory search: first, the decision tree, illustrated in figure 4, displays queries in tree format with pre-set value-range splits, allowing the user to walk the tree to reach different queries and their results in terminal branches. The tree can be reconfigured to choose query attributes, range splits and the query order.

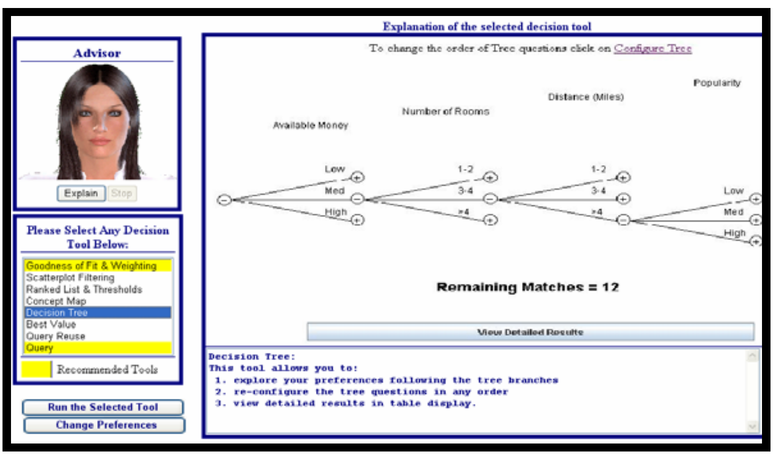


Fig. 4. Decision tree tool, also showing the tool recommender and advice

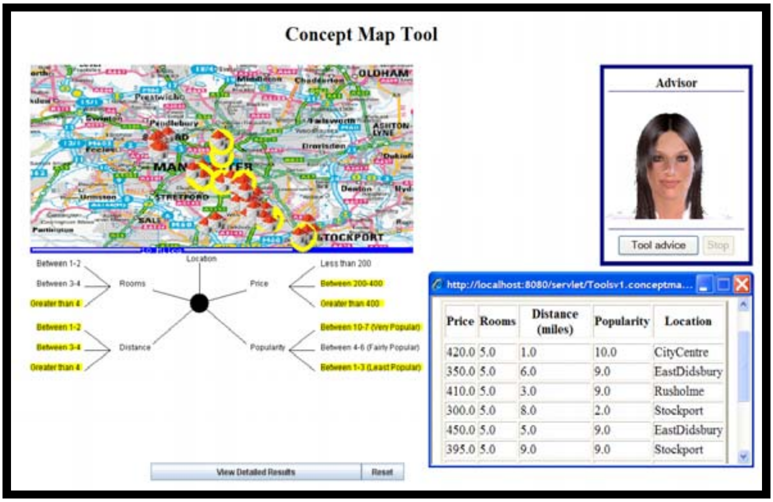


Fig. 5. Conceptual map tool

The second exploration tool is a conceptual map (see figure 5), which displays a flattened decision tree with range-split leaves. Queries can be formed by picking any combination of leaf nodes and selecting areas on the map. Results are displayed on the map as well as in a table.

3.2 Advisor Agent and Tool Recommender

The advisor agent illustrated in figures 2-5 provides advice on operating the appropriate tool, general search strategy advice and recommendations for selecting tools, tailored according to the user's answers in the configuration dialogue. The configuration dialogue asks the user to rate their motivation on a three-point scale and select either direct search or browsing exploration preference for their task. This information is combined with the product model which categorises the configured database on a 3-point scale for product value (unit cost), complexity and volume of products in the database. The tool recommender highlights one or more tools determined by a set of rules, based on ADM theory. An example rule to illustrate the format is:

> IF User <Motive = High, Task = Explore> AND Product <Value = High, Complexity = Medium, Volume = High> THEN Recommend <Concept-Map, Decision Tree, Scatterplot>

In this case compensatory exploration-support tools are recommended; for similar products with search tasks, ranked list-related tools are suggested; when user motivation is low, non-compensatory tools are preferred (best value, query, query reuse), etc. Speech advice is generated from text scripts which can be replayed or cancelled by the user.

4 Experimental Evaluation

The multi-tool environment was experimentally evaluated to assess the effect of advice and user tool/strategy choice. 24 subjects (14 males and 10 females, age range 20-39) from the University of Manchester participated in the experiment. Most subjects were postgraduate students with considerable Internet search experience. The subjects were paid £10 for their participation. A between-groups design was used to assess the effect of task and tool-use advice; the Advisor group were provided with tool recommendations, task-related advice and tool-use help, while the Control group had no advice or help. The experimental procedure was (i) brief training on using the eight decision tools and the advisor (in the Advisor group); (ii) experimental task with four scenario variations; (iii) post-test questionnaire; and (iv) debriefing interview. The task was to find apartments near the University of Manchester with the following scenarios, counterbalanced between subjects in both conditions:

Scenario 1: Low Motivation/Time and General Exploration

Assume that you do not have clear preferences, you have little time and you need to quickly find an apartment taking into account that the deadline for signing the apartment contract is closing soon. Find any apartment that would be acceptable.

Scenario 2: High Motivation/Time and General Exploration

Assume that you do not have clear preferences; you have plenty of time and you need to make sure that you find the most suitable apartment. Find an apartment that you really like.

Scenario 3: Low Motivation/Time and Search

Assume that you have little time and you need to quickly find an apartment taking into account that the deadline for signing the apartment contract is closing soon. Your preferences are for an apartment that has at least 2 bedrooms, its weekly price is less than £200 and is as close as possible to the university.

Scenario 4: High Motivation/Time and Search

Assume that you have plenty of time and you need to make sure that you find the best apartment you can. Find an apartment that has at least 3 bedrooms, its weekly price is less than £200 and is as close as possible to the university. Also, the most popular apartment is preferred. You should do more search evaluations and make sure the apartment is really suitable by checking the results.

No time limit was set for task completion and all subjects produced queries to complete the task. In the Advisor group, the subjects had to answer two configuration questions on their perceived motivation level (high/low) and the expected task (search/explore) after viewing the test scenario. Based on the configuration answers, the Advisor recommended one or more tools using rules was based on ADM theory [5]. In the Control group, the subjects had to select tools without any recommendations. The independent variables, manipulated in the four scenarios, were perceived motivation based on available time (high/low) and task type (explore/search). The dependent variables were number of tools used, performance times, tools selected, queries generated, usability errors and tool usage patterns.

5 Results

5.1 Number of Tools Used

Overall, more tools were used in the Control than in the Advisor condition (Mann-Whitney U=18, df=1, p=0.001); and this difference was also found when High-Motivation (2,4; U=19, df=1, p=0.001) and Low-Motivation (1,3; U=33, df=1, p<0.05) were compared with data for explore and search tasks. When analysed by task (motivations merged) more tools were used in the control condition for both the explore (1,2) - (U=34, df=1, p<0.05), and search (3,4) - (U=24.5, df=1, p=0.005), tasks. When differences were compared between groups for each scenario, more tools were used in the Control condition for both high motivation scenarios (4 Search, Mann-Whitney U=35, df=1, p<0.05) and (2 Explore, Mann-Whitney U=35, df=1, p=0.052), but not in the Low-Motivation scenarios (1 and 3); see figure 6. Within both Control and Advisor groups, more tools were used in the High-Motivation (2,4) than in the Low-Motivation scenarios (1,3; Wilcoxon test, p<0.005). However, no task (explore/search scenarios) differences were apparent in either condition, so it appears that system recommendation narrows the choice of tools, while motivation increases the diversity of tools used in both conditions.

When individual scenarios were compared within the Control group, more tools were used with higher motivations in both tasks (1 & 3) v. (2 & 4) p<0.01, Wilcoxon matched pairs test, and the same differences were found in the Advisor condition (p<0.01, Wilcoxon). However, no differences were apparent between scenarios with different tasks (explore/search) at the same motivation level in both groups, so users appear to follow similar strategies and tool use for search and explore oriented decisions.

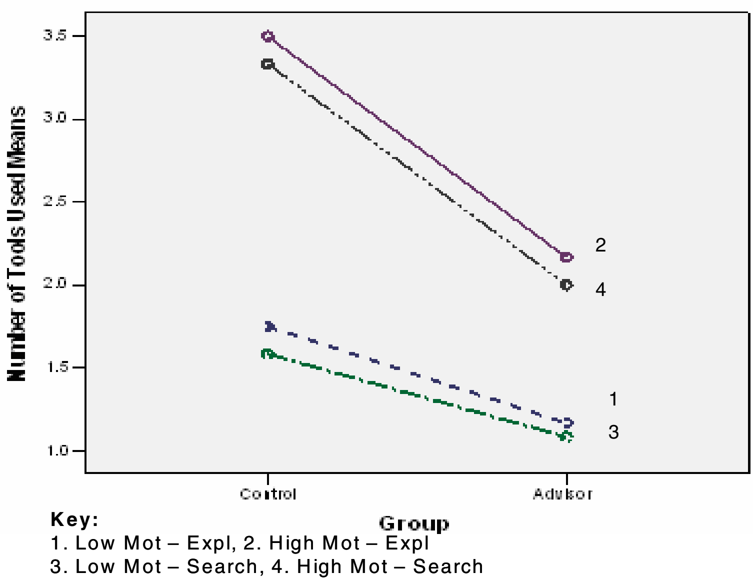
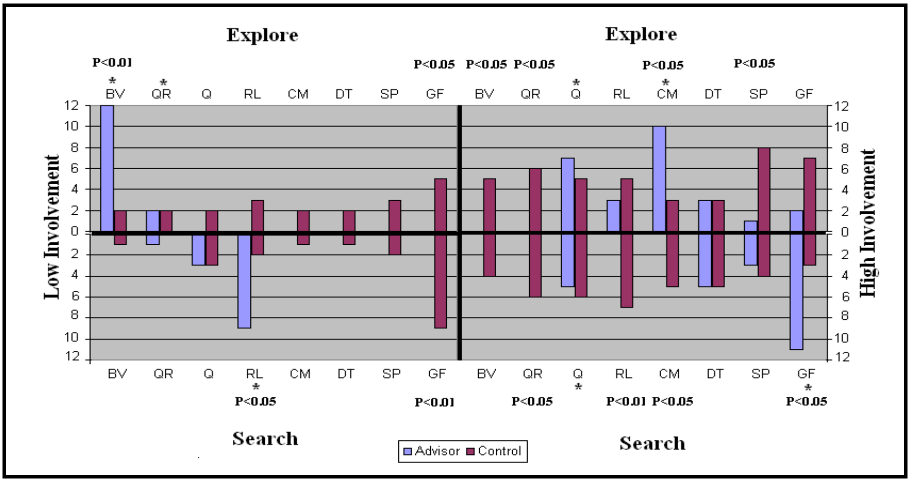


Fig. 6. Number of tools used



(Note: (*) indicates the recommended decision tools; only significant differences are reported, using Binomial test)

Fig. 7. Tool usage frequencies in different scenarios, for tool codes see figure 1 or Table 2

Figure 7 illustrates the subjects' tool selections in the four scenarios. Dark bars in the figure represent subjects' tool usage in the Control group (no advice) and light bars represent subjects' tool usage in the Advisor group (with advice). The left/right quadrants map to the subjects' level of motivation (low and high) while the upper/lower quadrants map to the subjects' task (explore and search). The (*) indicates the recommended tools in each scenario.

The left-hand (low motivation) quadrants illustrate little effect of the task with a similar tool use distribution in the Control condition; however, users followed the system advice and selected the recommended tools (best value (BV) for search and ranked list (RL) for explore), with no use of goodness-of-fit (GF), in spite of their preference in the Control condition.. However, the query reuse tool recommendation was partially ignored, possibly because the BV tool appeared first in the recommendation list. The high-motivation scenarios in the right-hand quadrants also show the impact of system recommendations, although a wider selection of tools were used even in the Advisor condition. No particular differences were apparent in the Control condition apart from a slight preference for scatterplot (SP) and goodness-of-fit in explore. When users were not guided they tended to use more compensatory and exploratory support tools (GF, RL, SP, CM), especially in the high-motivation scenarios. Decision tree and scatterplot tools were used in spite of the system recommendations. In debriefing interviews subjects reported they found the scatterplot easy to use for querying and the decision tree gave a good overview for both tasks.

5.2 Queries Generated

More queries were submitted by the Control group than by the Advisor group ($p<0.05$, Binomial test on totals), and following the pattern of previous results, high-motivation scenarios produced more queries than the low-motivation scenarios

Table 1. Queries Submitted by Scenario in both conditions

	Scenario 1: L-E	Scenario 2: H-E	Scenario 3: L-S	Scenario 4: H-S	Total
Control	28	86	22	63	199
Advisor	2	68	16	54	152

L, H = Low, High Motivation; E, S = Explore, Search.

(p<0.05), in both groups; see table 1. The Control group created more queries than the Advisor group, especially in Scenario 1 (p<0.0001) and a few more in Scenario 2 (p<0.05), but differences were not significant in the other two scenarios. There were no differences in query totals between the tasks for either group.

Advice focused tool usage and encouraged users to match tools to the tasks more effectively. Although lower query volumes in the Advisor group might be taken as evidence of more efficient usage in the search scenarios (2, 4), the totals were lower, and inter-group differences just failed to reach significance: p=0.52 in scenario 4, ns in scenario 3. The lower query volumes in the exploration scenarios (1 and 2) might be seen as less productive; however, the CM and DT tools afforded inspection-based exploration in both conditions, so fewer queries may have been necessary.

In the Control group more queries were generated with the SP tools, possibly reflecting the ease of slider-based query formulation (see table 2). The Advisor group focused tool choice in both tasks, generally following system recommendations, although the subsequent rank-order query generation frequency was GF>RL>CM>QR. In the Control condition, queries were evenly distributed among the tools apart from decision tree (DT), which may have been used more for inspection in both conditions.

Table 2. Queries submitted by tool type in each condition

	Control	Advisor	Sig
Best Value	12	12	ns
Query Reuse	25	4	**
Query	22	26	ns
Ranked List	29	18	*
Concept Map	26	30	ns
Decision Tree	16	16	ns
Scatter Plot	43	22	*
Goodness of Fit	38	24	*
Total	199	152	*

*Sig = Binomial test on totals * p <0.05, ** <0.01.*

Within the Advisor group, more frequent queries were created with the concept map (CM) tool, followed by Q>GF>SP, with more marked inter-task differences in use. The effect of advice appears to be to reduce query volumes for query reuse (QR), RL, SP and GF tools, reflecting more focused tool use ($p<0.05$ Binomial test). Thus, when the users were guided they showed more focused patterns of tool use, closely matching the scenario tasks (e.g. CM in scenario 2 explore, GF in scenario 4, search).

The main effect of the advice was to shift use from the SP, GF, RL and Q tools, which supported compensatory decision making, to the CM for an exploration overview.

5.3 Search Times and Performance

All subjects completed the search tasks correctly and there were no significant differences in the subjects' search times between the groups in all scenarios (ANOVA), even though subjects in the Advisor group had an additional task in configuring the system and listening to/reading advice for tool recommendations and tool-use help.

When configuration and advice access times were subtracted from total completion times, the Advisor condition was significantly faster than the Control group ($F=14.73$, $df=1,23$, $p=0.001$) and for each scenario ($p<0.05$); see figure 8. Following the pattern of tool usage, the motivation manipulation also produced significant differences in completion times between scenarios 1, 3 and 2, 4 for both conditions (t tests, $p<0.5$). This shows that task-related and operational help was effective in improving the subjects' search performance.

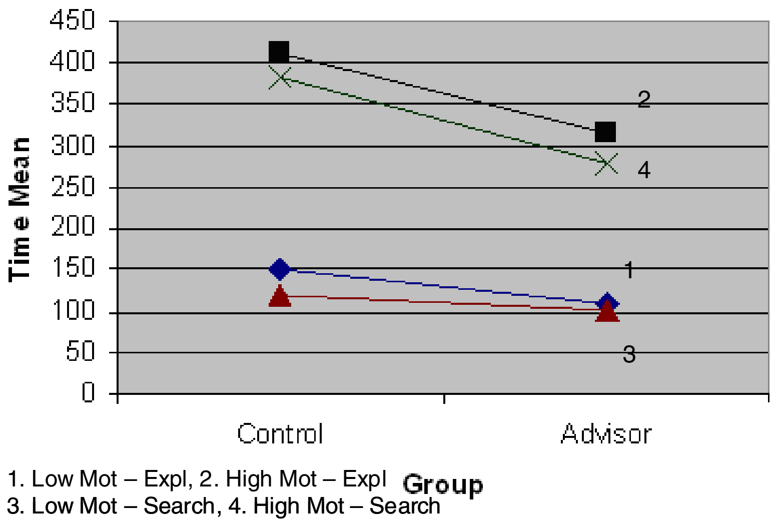


Fig. 8. Search performance (Advisor vs. Control) – time in seconds, excluding configuration and advice accesses time

5.4 Usability Problems

There were few usability problems overall, although the Control group did experience 14 errors compared with 8 in the Advisor group. The errors were either problems in using the slider controls or failing to find the sort/rank functions on the table headings. All errors were eventually overcome in both groups. In the Advisor group 15 critical incidents were reported when users complained of difficulty in mapping the scenarios to the answers in the configuration dialogue.

5.5 Debriefing Interviews

In debriefing interviews, positive comments were made about the system providing different ways of searching that matched different situations (Advisor group: 83.3%, Control group: 75%), being better than the current e-commerce tools (A: 75%, C: 66.7%). Most subjects reported that the tools were easy to use (100%, 83.3%) and that the tools complemented each other (75%, 58.3%). The Advisor group reported that the highlighted (recommended) choices helped them to decide which tool to use in different scenarios (83.3%). Use of sliders for queries rather than entering values into text boxes also received many favourable comments (75%, 58.3%).

Negative comments came mainly from the Control group, highlighting the need for tool recommenders; for example, too many tools, confused over which one to select (91.7%); some tools require help (e.g. DT, CM, SP filtering tools: 75%); and advice on which tool is appropriate for each task would be helpful (91.7%). A few subjects in the Advisor group complained that the configuration questions were not clear (33.3%); however, they liked the agent appearance and noted that it drew their attention to the advice (91.7%), and preferred speech to text in the configuration advice (58.3%), even though the speech sounded artificial (66.7%).

6 Discussion

The environment described in this paper illustrates theory-led design, in that the design of the tools and the recommender was based on Adaptive Decision Making, sound cognitive theory describing user strategies for making choices. Our design approach follows in the toolkit tradition of HCI design in which tool support is targeted at user strategies within a complex task, hence building on research into adaptable and adaptive systems [6]. The toolkit can be exploited in two ways, first using the recommender as illustrated in this paper, in which the assumed implementation environment is adaptable decision support which could be portable across domains. This would require configuration of product models and database querying modules for several domains, which could be provided with modest development effort. The second route follows a toolbox paradigm in which the recommender indicates appropriate tools for a specific application and a sub-set of the tools are configured for that domain.

The experimental evaluation demonstrated that the system advice and tool recommendations were effective in shaping user behaviour and improving the efficiency of decision making. Post-experiment questionnaires and interviews also demonstrated that the recommender and toolkit design was well received. While evidence for improved performance is limited, we can point to the reduced number of queries submitted for task completion and the shorter completion times, once allowance was made for the configuration dialogue. Since the configuration setup dialogue is a one-off investment for each application domain, over time the performance improvement payoff should be considerable. The tool recommendations influenced user tool choice following the strategies in ADM theory which were embedded in the recommender rules. The scenario manipulations also demonstrated that users follow the advice according to their motivation and the task context. Observed tool choice matched the system recommendations; however, occasionally users seemed to prefer using the scatterplot and decision tree in the high motivation-search tasks in spite of the system advice. This may reflect a sequential ordering in users' strategies, in which they explore the search space to check alternatives even when undertaking well specified searches. In our future work we will investigate recommendations for sequential and integrated use of the tools in both exploration and search tasks. Active training in trade-off strategies improves user exploration of options and satisfaction [3, 21], so another direction is to augment our advisor with more directed tuition for compensatory tools.

The limitations of the study lie in the possible suggestibility of our participants who may have followed the system recommendations from an experimental task conformance bias; however, we point to the positive comments and questionnaire rating which the participants made about the recommenders and their reports that, while provision of several tools was useful, they were more effective with a recommender. The same criticism can be levelled at the scenario manipulations; although in this case we point to the clear effect of the manipulations in tool choice, and the credibility checks we carried out in de-briefing interviews. The subjects' choice of decision-making strategies elicited from interviews not connected to the experimental task demonstrated that the selected strategies agreed with the observations in the experimental tasks.

In conclusion, the system proved to be an effective means of improving users' performance in complex decision making where multiple strategies will be necessary as rapid change in domains can influence task strategies. The toolkit recommender architecture has the potential for exploitation as both a configurable toolbox and a toolkit for flexible decision support across domains. From our results it appears that user-driven decision making with exploration as well as search is an important part of e-commerce functionality which needs to be balanced against the current trend for developing system initiative recommender systems.

References

[1] Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Interaction. Addison-Wesley, Reading (2004)

[2] Benyon, D., Turner, P., Turner, S.: Designing Interactive Systems: People, Activities, Contexts, Technologies. Addison Wesley, Reading (2004)

[3] Pu, P., Kumar, P.: Evaluating Example-based Search Tools. In: Proceedings of the 5th ACM Conference on Electronic Commerce. ACM Press, New York (2004)

[4] Ahlberg, C., Shneiderman, B.: Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In: Readings in Information Visualization: Using Vision to Think. Morgan Kaufmann, San Francisco (1999)

[5] Payne, J.W., Bettman, J.R., Johnson, E.J.: The Adaptive Decision Maker. Cambridge University Press, Cambridge (1993)
[6] Fischer, G.: User Modeling in Human-Computer Interaction. User Modeling and User-Adapted Interaction 11(1/2), 65–86 (2001)
[7] Flor, G.D.L.: User Modeling & Adaptive User Interfaces, ILRT Research, University of Bristol, UK (2004)
[8] Al-Qaed, F., Sutcliffe, A.G.: Adaptive Decision Support System (ADSS) for B2C E-Commerce. In: Proceedings 8th International Conference on Electronic Commerce, pp. 492–503. ACM Press, New York (2006)
[9] Corbin, R.M.: Decisions that Might not get Made. In: Cognitive Processes in Choice and Decision Behavior. Lawrence Erlbaum Associates, Hillsdale (1980)
[10] Hong, W., Thong, J.Y.L., Tam, K.Y.: The Effects of Information Format and Shopping Task on Consumers' Online Shopping Behavior: A Cognitive Fit Perspective. Journal of Management Information Systems 21(3), 149–184 (2005)
[11] Hung, L.-P.: A Personalized Recommendation System Based on Product Taxonomy for One-to-One Marketing Online. Expert Systems with Applications 29(2), 383–392 (2005)
[12] Langley, P.: User Modeling in Adaptive Interfaces. In: Proceedings of the Seventh International Conference on User Modeling. Springer, Banff (1999)
[13] Hildebrand, C.: One to a Customer: Customer Relationship Management. CIO Enterprise Magazine (1999)
[14] Ardissono, L., Goy, A., Petrone, G., Segnan, M.: Personalization in Business-to-Customer Interaction. Communications of the ACM 45(5), 52–53 (2002)
[15] Ardissono, L., Goy, A., Petrone, G., Segnan, M.: A Multi-Agent Infrastructure for Developing Personalized Web-Based Systems. ACM Transactions on Internet Technology 5(1), 47–69 (2005)
[16] Pazzani, M., Muramatsu, J., Billsus, D.: SYSKILL & WEBERT: Identifying Interesting Web Sites. In: Proceedings of the Thirteenth National Conference on Artificial Intelligence. AAAI Press, Portland (1996)
[17] Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.: Evaluating Collaborative Filtering Recommender Systems. ACM Transactions on Information Systems 22(1), 5–53 (2004)
[18] Shneiderman, B.: Creating Creativity: User Interfaces for Supporting Innovation. ACM Transactions on Computer-Human Interaction 7(1), 114–138 (2000)
[19] Haubl, G., Trifts, V.: Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids. Marketing Science 19(1), 4–21 (2000)
[20] Pereira, R.E.: Optimizing Human-Computer Interaction for the Electronic Commerce Environment. Journal of Electronic Commerce Research 1(1), 23–44 (2000)
[21] Pu, P., Chen, L.: Integrating Tradeoff Support in Product Search Tools for E-Commerce Sites. In: Proceedings of the 6th ACM conference on Electronic Commerce. ACM Press, New York (2005)

Designing for Culturally Contextualized Learning Activity Planning: Matching Learning Theories and Practice

Aparecido Fabiano Pinatti de Carvalho[1,2], Junia Coutinho Anacleto[1], and Vania Paula de Almeida Neris[1,3]

[1] Advanced Interaction Lab, Computing Department, Federal University of São Carlos, São Carlos, Brazil
[2] Interaction Design Centre, Department of Computer Science and Information System, University of Limerick, Limerick, Republic of Ireland
[3] Computing Institute, State University of Campinas, Campinas, Brazil
fabiano.pinatti@ul.ie, junia@dc.ufscar.br, neris@ic.unicamp.br

Abstract. Helping teachers in their activities has been an issue more and more explored in Computer Science. However, in order to support teachers effectively, it is necessary to understand their needs and to design tools that they can easily manage. One of those needs is undoubtedly to put in practice pedagogical principles. This paper presents the design of PACO-T, a tool for helping teachers in planning learning activities (LAs) supported by common sense knowledge, based on PACO, a seven-step textual framework for planning pedagogically suitable LAs. The design was based on the results of a case study carried out to investigate how teachers can plan LAs following PACO steps, using common sense knowledge from a common sense knowledge base collaboratively built through the web. Moreover, the interface design was ruled by a Web Design Pattern Language, attempting to improve the usability of the tool. PACO-T aims to help teachers to put in practice the recommendation for contextualizing LAs to the target group, found in several learning theories.

Keywords: e-learning, learning activity, contextualization, common sense knowledge, computer tool, design, Web Design Pattern, pedagogical issues.

1 Introduction

Technology can be an important support to education. Due to such support, e-learning practices are becoming more and more usual nowadays. Despite of the potential use of technology for education, understanding the needs of the actors involved in the process, i.e. teachers and students, is very important to allow the development of tools that can successfully support the educational process. For instance, in order to effectively support education, it is important to take into account pedagogical issues surrounding it in the development of learning technology, so that the developed solution can help teachers to put in practice such issues [1, 2].

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 777–790, 2009.

One of the main pedagogical issues discussed in relevant Learning Theories like those proposed by Freire [3], Freinet [4], Ausubel [5] and Gagné [6], is to contextualize the learning process to the target group's cultural background. A way of doing that, widely discussed by Freire [3], is to use common sense knowledge, i.e. the knowledge shared and accepted as true within a social community, spanning several knowledge categories such as temporal, spatial, social, cultural and so forth [7], for leveraging discussion and promoting effective learning.

Though a well-known pedagogical issue, until early 2006 it was not possible to find in the literature how teachers could easily access the common sense knowledge of their target group in computers, especially because the amount of information comprising common sense knowledge is huge – around 1,000,000,000 bits, according to Landauer [8]. However, since Lenat started working in Cyc, aiming to build a large scale common sense knowledge base [7], there have been advances in this area through the use of computational technology [9]. Making use of those advances, recent researches [10-13] have shown that common sense knowledge collected collaboratively through the web can be used to state the target group's previous knowledge and to identify its needs.

Therefore, this paper focus on the design of a computational tool that makes teachers able to access and to explore common sense knowledge, collected collaboratively through the web and stored in the knowledge base of the project OMCS-Br (Brazilian Open Mind Common Sense) [14], during the planning of learning activities (LAs). PACO-T is a common sense-aided **T**ool which guides teachers towards the planning of culturally contextualized LAs. It is based on PACO [15], a seven-step textual framework for planning pedagogically suitable LAs.

It is worth pointing out that the design of PACO-T was concerned about HCI (Human-Computer Interaction) issues, with the purpose of developing a tool under usability criteria, so that teachers can easily interact with it. For that, a case study was conducted to identify *(1)* how the process of teachers' planning LAs using PACO could be mapped into a computational tool; and *(2)* how common sense knowledge can help teachers answer questions brought up along the steps of PACO, so that they can contextualize their LAs to their target group, which can be either the students who are going to participate of a LA or the members of a social group with whom the students are going to interact after the LA in order to apply the knowledge acquired [10].

Still related to HCI issues, the design of the interfaces was ruled by a Pattern Language for designing web application proposed by Montero et al. [16], since the tool is a web application. Once Patterns are successful solutions for recurrent problems [17], the design of interface ruled by Patterns can lead to solutions already well disseminated among users that they will probably know and be able to deal with easily, which will consequently improve the system usability and accessibility.

The paper is organized as follows: section 2 discusses how common sense knowledge can be used for supporting teachers in planning LAs according to the proposal of the framework PACO, taking into account the results of the case study previously mentioned; section 3 presents the design of PACO-T, discusses some decisions for providing usability for the tool and goes over some details of its implementation; finally, section 4 presents some conclusion remarks and points to future works.

2 The Framework PACO and Common Sense Knowledge – A Case Study

PACO is a textual framework designed to support teachers in **P**lanning LAs supported by **CO**mputers, which is composed by seven steps [15]:

1. To define the LA theme, target public and general goal;
2. To organize the LA topics;
3. To choose a pedagogical/methodological reference;
4. To plan the learning tasks;
5. To choose computer tools to support the activities execution;
6. To edit the learning objects which are going to be used in the LA ; and
7. To test pedagogical and technological issues.

The essence of PACO relies on the fact that the definition of the learning tasks and the selection of the computer tools to support their performance should be addressed by pedagogical issues and by characteristics of the target group. PACO has already been used by teachers from different areas such as computer science, nursing and occupational therapy, who have no previous knowledge about planning LAs supported by computers. The feedback received from these case studies showed that, even though someone does not have experience on planning LAs, s/he can do it by following the framework steps [15].

One of the LA planning made by nursing teachers was specially proposed to *(i)* analyze the possibility of using common sense knowledge during the planning of LAs, in order to contextualize them to the target group's needs, and *(ii)* design a computational tool to support teachers with this task [10, 13, 18]. In the case study, two nursing teachers planned a LA to prepare their students on how to orient caregivers in the community from which the common sense knowledge was collected. Therefore, the educational context was composed by three actors [10]:

1. Nursing teachers, who had to plan a learning activity following PACO steps, taking into account the information stored in a common sense knowledge base;
2. Second year nursing students, who were going to be thought how to orient caregivers;
3. The population, whose members could become a caregiver, from which the common sense knowledge was collected.

The common sense knowledge was used to call the students' attention to the way which the population talked about requirements to be a caregiver or about procedures which might be taken while home caring a sick person. The students were presented points which they should emphasize during the orientation. Thus, the target group was members of the community where the common sense knowledge were collected from, who were going to take care of a sick person at home.

Through the case study, it was possible to identify how common sense knowledge could support teachers in answering questions brought up along the framework steps, as it is explained in the following. In that occasion, the common sense knowledge in the OMCS-Br knowledge base was semi-automatically organized in an on-line common sense matrix, taking into account specific parameters defined by the teachers involved in the LA planning – see [10] for more details. All the planning

Table 1. Possible support offered by common sense knowledge in each step of PACO[10]

Step	Support
1	To define the LA theme.
	To compose the LA justification.
	To define the LA general goal.
	To define the LA specific goal.
2	To decide on topics that should be approached during the LA, so that the LA fits the students' needs.
	To decide the degree of detail with which each topic should be approached.
3	To reach pedagogical issues addressed in Freire's, Freinet's, Ausubel's and Gagné's Learning Theories.
4	To fit the LA tasks to the pedagogical/ methodological references adopted.
	To know how the target group usually study.
5	To know with which computer tools the target group is familiar.
6	To compose the learning material.
7	-

was followed and registered in field notes by a researcher who used to interact with the teachers in some moments in order to illuminate issues related to what he had observed. In those interactions the teachers explained their reasons for having done specific decisions for the planning, allowing the researcher to understand what they were thinking of. The teachers were also asked to think aloud, verbalizing their thinking during the planning. Everything was audio recorded and both the field notes and the audio were analyzed afterwards. Through the data analysis it was confirmed that common sense knowledge can support teachers during the planning of LAs using PACO in several different ways, as Table 1 summarizes.

Concerning the framework Step 1, it turned out that teachers could think of defining the LA theme taking into account the needs that they observed in the common sense knowledge available [18]. In the case study, although the LA theme had been previously well defined – how to orient people to home care a sick person –, teachers identified themes which could originate other LAs. One example is the theme "getting basic knowledge about nursing practices". By the analysis of the available OMCS-Br common sense knowledge, the teachers noticed that the population usually mentioned that, in order to home care a sick person, it is necessary to learn about basic nursing practices. From this evidence, the idea of planning a LA in the future about that theme came out.

Still related to the Step 1, it was also found out that common sense can help teachers to compose the LA justification, another task of PACO Step 1. For instance, in some situations teachers realized that some topics were not mentioned in the knowledge base, when they should be, or that the members of the target group misunderstood related topics, or even that the target group was very interested in the topic. Based on those evidences, teachers were able to compose the justification of promoting a LA on that theme.

In the same way, considering what was observed in the knowledge base, the teachers could define the general and the specific goals of their LA. For example, still considering the theme "getting basic knowledge about nursing practices", teachers told that they could propose as general goal of the new LA "to teach the population about basic nursing practices". The comment came out when the researcher who was observing the LA planning asked the teachers how they thought that the knowledge they were analyzing could help them define the LA goals. Moreover, according to the items they found in the knowledge base, they could propose specific goals such as "to teach the population about techniques to control a fever", or "to teach the population techniques to change the clothes of a sick person that cannot move", and so on.

Regarding the framework Step 2, where teachers have to define the topics they are going to approach in the LA, to organize these topics, and to define how specific they will be in each topic, the teachers could also collect evidences from the common sense knowledge base to help them with these tasks. In the case study, based on what they verified in the common sense knowledge from the OMCS-Br knowledge base, they selected topics such as "the importance of caregivers' having leisure" and "the diet of sick people" to be approached in the LA.

The first topic was proposed because the teachers realized that people failed to mention that a caregiver needs help to take care of a sick person at home when they were exploring the statements available in the common sense knowledge base. According to them, the common sense knowledge showed that people from the target group considered that, in order to home care a sick person, it is necessary to dedicate all their time to the task, forgetting to mention their own leisure, resting and comfort. This fact should be considered, according to the teachers, because the caregiver needs to rest in order not to become sick. Hence, this topic was emphasized during the LA, so that the nursing students remind the caregiver about the need of having someone else to share the responsibilities of home caring the sick person so that s/he can rest and can be well to perform the home care.

The second topic was chosen because people from the target group always mentioned that sick people like or need to eat soup. According to the teachers, this is not always true. Depending on the illness or its seriousness, the sick person can eat and sometimes has to eat other things besides soup. Thus they decided to add a topic in order to teach their students about which kind of diet is suitable for which situation and to make their students aware about the need for giving this orientation to the caregiver, since the caregiver possibly share the opinion that a sick person should eat only soup, once this is the common sense from her/his community.

About evidences that can help teachers to decide the degree of detail which they should give to specific topics during the LA planning, the teachers tended to propose several tasks to work on subjects that the target group failed to mention or misunderstood, such as the ones previously addressed, while they tended just to mention, in abstract terms, the subjects which they considered that the target group already knew. For example, the teachers found out in the knowledge base that the target group knew very well the main requirements that one should have in order to be a caregiver. In this case, they propose a simple activity to remind the students that they should check whether the person who is going to home care a sick person fits to the requirements for being a caregiver. They decided to keep this topic in the LA because they were concerned that, although the target group's members knew about the main requirements for being a caregiver, sometimes a person who does not fit to the requirements may have to home care a sick person.

In relation to Step 3, the case study showed that common sense knowledge can support teachers to plan LAs according to four Learning Theories [18]: Freire's [3], Freinet's [4], Ausubel [5], and Gagné [6]. As in the Step 3 teachers have to choose pedagogical/methodological references to plan the LA tasks, teachers are advised to select one of those Learning Theories, so that they can use common sense to plan the tasks that the students will perform to reach the LA goals.

Depending on the pedagogical/methodological references chosen for the task in Step 3, the planning of the tasks in Step 4 can also be supported by common sense knowledge. For example, Freire's Learning Theory suggests that teachers should make students build the new knowledge on their common sense. According to this theory, teachers should discuss with the students the reasons why they believe that some pieces of common sense knowledge are correct and then introduce the new knowledge correcting or complementing the previous knowledge [3]. Thus, adopting this pedagogical/ methodological reference, teachers can search on the common sense knowledge the pieces of knowledge they can use in the learning tasks. Ausubel's Theory defends a similar approach. According to Ausubel, in order to promote meaningful learning, teachers should make the students attach the new piece of knowledge to the knowledge which is already in their cognitive structure. Considering that common sense knowledge is part of the cognitive structure of all students, the planning of the learning tasks can be directed so that meaningful learning can take place. More details of how common sense can support these two Learning Theories and the others previously mentioned can be found in [2].

Besides the support to fit the LA tasks to the chosen pedagogical/methodological references, teachers can also search in the common sense knowledge base for information on how people from their target group usually study. For example, if teachers find out, in the knowledge base, statements such as "in order to get informed, one has to read" and "in order to learn, one has to study several books", they can interpret that people from that target group are used to reading and, therefore, propose some readings to discuss the topic. On the contrary, if teachers find out statements such as "to get informed, one has to watch television" and "one can learn a lot watching a movie", they can try to find some movies related to the LA and propose other kind of activities.

In the same way, teachers can search the common sense knowledge base for information about computer tools and try to figure out which tools are frequently used by people from their target group. In this manner, they can try to choose computer tools which the target group members are familiar with, avoiding that students have to spend lots of time to learn how to interact with a tool before starting to perform a task. This is the way that common sense can support teachers in PACO Step 5.

Finally, regarding the support that common sense can give in Step 6, a recent research has shown how common sense knowledge can be used for helping teachers to edit learning objects and fulfill their metadata in order to facilitate their discovery and reuse. See [19] for details.

3 Designing PACO-T

The case study previously presented allowed a requirements elicitation on issues that a computational tool should address for supporting teachers in planning contextualized LA in the same way that the teachers of the case study were able to do. From the elicitation, it was possible the computational design of PACO-T, which is presented along this section.

The requirements elicitation showed that a solution to effectively support teachers in planning common sense contextualized LAs according to the proposal of PACO, would be develop a question-based system to guide teachers in providing the necessary information for the LA planning. Therefore, it was decided to organize the questions that the tool would present to teachers in seven steps, as in the original proposal of the framework, since previous studies showed that this organization makes easier the teachers' task of planning LAs [15]. During these steps, teachers should be offered access to the common sense knowledge, so that they can use it in the same way that the nursing teachers could use during the case study LA planning.

In so doing, it was decided to use the resources of the OMCS-Br project, since it provides a framework for collecting and using common sense knowledge in computer applications [2, 14]. It was also decided to develop the tool for the Web platform, considering the facilities that the Web provides for accessing computer systems all over the world and the wide spread dissemination of Web technology. Nowadays more and more people get used to web technology in their work and personal life, using e-mails, electronic agendas, on-line shopping system and so forth, since they offer flexibility and fast solutions in a society where time is very restricted.

The tool was designed using the paper prototyping technique, since it is considered a low cost and suitable technique for eliciting and refining computer system requirements [20]. Thinking of achieving usability criteria in the tool, its interfaces were sketched following the Patterns of the Web Design Pattern Language proposed by Montero et al. [16], since that Pattern Language was proposed envisioning to improve usability in sites designed trough it [16] and PACO-T would basically be a web site where teachers would be able to plan LAs. The set of usability guidelines proposed by Nielsen for designing web systems [21] were also considered.

Figure 1 shows the prototype of the tool home page. Specifically for this paper, *the paper prototype* was freely translated into English, because the system is only available in Portuguese. The application of ten Patterns from the considered Pattern Language can be observed in that interface as it is explained in the following. To begin with the Patterns related to the whole interface, labeled with number 1. This interface implements the Patterns *Welcome* and *Homepage.* According to [16], every single web system must provide a place that allows users to identify where they are. Providing a reception place where use and access conditions can be evaluated is the proposal of the Pattern *Welcome.* The other referred Pattern, *Homepage,* goes towards the need of providing a point of reference, which allows users to answer what they can get from that web site, how and when. The authors of the Pattern Language mention that usually the implementation of the Patterns *Welcome* and *Homepage* is the same many times. This is what happens in PACO-T.

Therefore, the interface in Figure 1 is proposed to allow teachers to identify where they are and what they can do in that place. The implementation of other Patterns of the Pattern Language gives support for these goals. For instance, the area in Figure 1 labeled with number 2 corresponds to the solution used in the system for the Pattern *Tagline,* which suggests the presentation of a statement summarizing the site purpose [16]. The text in area 7 also informs the user about the purpose of the site. That text was composed having in mind the suggestions of the Patterns *Polyglot* and *Polite.*

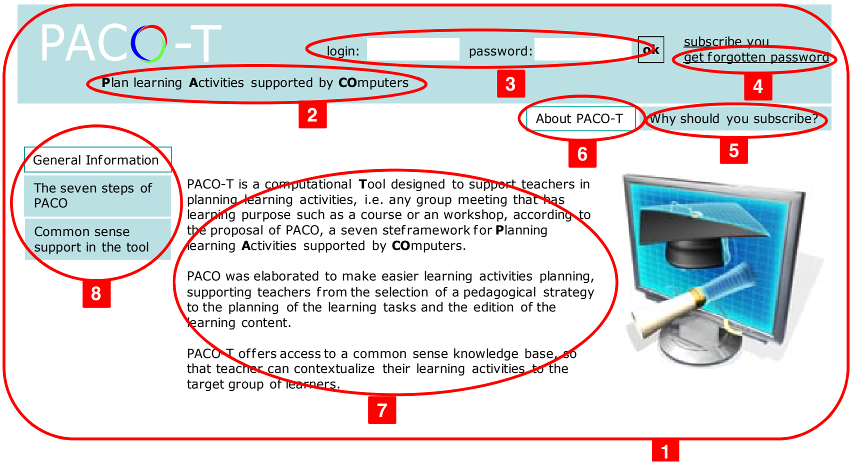


Fig. 1. PACO-T Home page

The first one suggests the use of a language that can be understood for the largest variety of people who can access the site. The second one advises to care about the visitors. Thinking about those suggestions, a short text was composed, explaining the purpose of the tool, using a simple vocabulary and suggestive examples to illustrate terms that could not be understood by ordinary people who visited the site, such as the concept of LA. The *Polite* can also be noticed in area 2, since the system cares about explaining to the users why they should register themselves on the site.

Other Patterns present in the interface showed in Figure 1 are: *Form* (area 3), *Second Chance* (area 4), *About us* (area 6), *Indication* (areas 4, 5, 6 and all other links available in the interface) and *Location* (in general, the whole interface that allows teachers to recognize that they are in the site/tool PACO-T and, specifically, the areas 6 and 8 that allows teachers to know that they are in the page "General information" of the section "About PACO-T"). The tool also applies the Pattern *Danger*, which says that a site should be designed to need as few plug-ins as possible for working [16]. PACO-T does not require any specific plug-in to work.

Regarding the use of usability guidelines in the design, the anchors of the links in the pages were designed to be as meaningful as possible, i.e. the text of the links should allow teachers to know what information or functionality they are going to get if they click on it. Another guideline followed was to use the imperative form of the verbs in links that corresponds to functionalities of the system. For example, the verbs of the links "subscribe you" and "get forgotten password" are in imperative form.

Thinking of usability issues, each step was split in several sub-steps in order to reduce the amount of information with which teachers should deal in each interface. The left navigation tree presented in Figure 2 allows seeing such organization in sub-steps. As Figure 2 shows, there are five sub-steps in PACO-T Step 1: *(i)* Define the LA title; *(ii)* Define the target group's profile; *(iii)* Define the theme, *(iv)* Define the objectives and *(v)* Add complementary information.

Note that the interface was designed for guiding teachers along the interaction. So, it is worth calling the attention for the usability guidelines applied in the interface at

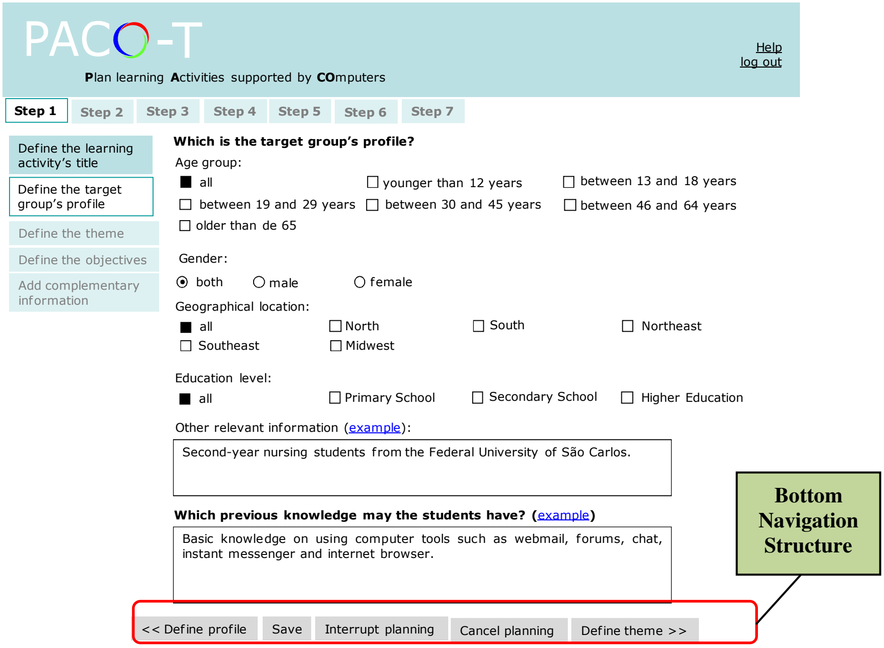


Fig. 2. PACO-T Step 1 – Define the target group's profile

Figure 2. The interface tells exactly where the teacher is (Step1/Sub-step 2 – Define the target group's profile). Moreover, it tells which Step/sub-steps have been already concluded. In this example, no step has been concluded yet, as it can be noticed by the weaker color of the tabs correspondent to Steps 2 to 7. Nonetheless, regarding the Step 1, it can be noticed that sub-step 1, "Define the learning activity title" has been already done. The other sub-steps have not been done yet, as the weaker colors indicate. This is a usability guideline for designing web systems proposed by Nielsen [21] and also a solution for the Patterns *Location* and *Indication* [16].

In addition to that, the interfaces were designed to prevent user errors. For example, links and tabs that cannot be clicked yet (the ones in weaker colors) are deactivated. They are kept in the interface so that the user can have an idea in which point of the planning they are, according to the solution proposed by the Pattern Indication[16]. However, as PACO defines that the planning should happen sequentially, since a step depends on information given in previous steps, teachers can only proceed when they finished filling the form in and clicking on the button to go to the next step, available in the bottom navigation structure. These buttons always have the symbol ">>" indicating the action of going forth to the next step.

Teachers also have total control of the system, one of the main usability issues. They can go back to a previous step they have already concluded and modifies what they want; they can interrupt the planning and continue later; they can cancel the planning and so forth. Furthermore, teachers can ask for information about the step where they are by clicking on the link "Help" or get sample information about how they should answer each question on the interface by clicking on the link "example" available after the question. All of those design concerns were explored in the interface taking into account the orientation of the Pattern Language used to guide the design and the usability guidelines presented in [21].

Along the system, teachers face with questions such as the ones shown in Figure 2. These questions support them in planning the desired LA and the questions were defined taking into account the information required by PACO.

The interface in Figure 2 precedes the first sub-step in PACO-T where teachers can use the support offered by common sense knowledge. It is an important interface, because it is where teachers define the LA target group's profile and, consequently, allows the system to filter the common sense knowledge base that will be used so that the tool presents to teachers only common sense knowledge collected from people with the same profile of their target group. Regarding implementation issues, that interface allows the system to connect to a specific semantic network of the OMCS-Br project. Semantic networks are the knowledge representation adopted in OMCS-Br (for details about the decision of representing the knowledge as a semantic network, see [2, 14, 22]). In OMCS-Br, there are several semantics networks representing the knowledge collected in the project, which are called ConceptNets. This is because OMCS-Br offers the possibility of generating a ConceptNet to each combination of the profile parameters established for the project (age group, gender or geographical location and education level). Figure 3 presents an interface where common sense knowledge can be helpful for teachers.

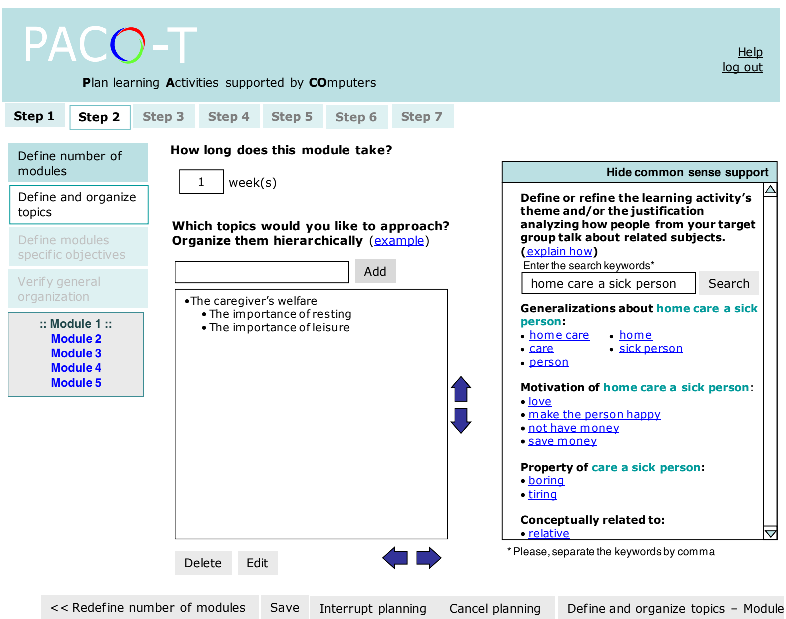


Fig. 3. Computational Representation of PACO – Step 2 – Define and Organize Topics

In this interface, teachers should define the topics that they want to approach in the LA. In order to analyze how people talk about subjects of interest and decide for the topics of the LA, teachers have to provide some keywords in the search text field and click on the button “Search” or to add a new topic in the list of topics. In doing so, the system sends the keywords to a function in the API called *GetContext()*. The *GetContext()* retrieves the context related to the provided keywords. It tries to find nodes in the ConceptNet equals to the keywords and then performs spreading activation radiating outward from the nodes found [22], i.e. it performs a semantic expansion on the keywords provided. After finding words related to the context, the system sends the words which relevance to the context is greater than 30% to the *DisplayNode()* function. This function receives a word or a phrase and returns the relations in the ConceptNet which have a node equal to the provided entry. In this way, the words returned by the *GetContext()* function are sent one by one to the *DisplayNode()* and five relation at most of each relation type returned are considered.

For instance, consider the example in Figure 3. The teacher provided the keyword “home care a sick person”. The system sends to the *GetContext()* the following keywords: “home care a sick person”, “home”, “care”, “sick”, “person”. As a rule, composed keywords are split before being sent to *GetContext()* in order to increase the possibility of return. It is possible to do this without impacting in the desired context, because *GetContext()* retrieves the contextual intersection of multiple concepts [22]. In the example previously described, some words and expressions in the context returned by *GetContext()* that have relevance greater than 30% were: “home care a sick person”, “home care”, “care a sick person”, “love”, “save money”, “relatives”, “have time” and “be prepared”. Each word was sent to the *DisplayNode()* function. For example, to the expression “home care a sick person” some relations returned by *DisplayNode()* were:

```
(SuperThematicKLine "home care sick person" "care sick person")
(SuperThematicKLine "home care sick person" "care")
(SuperThematicKLine "home care sick person" "care")
(MotivationOf "home care sick person" "love")
(ConceptuallyRelatedTo "home care sick person" "relative")
(MotivationOf "home care sick person" "make the person happy")
…
```

SuperThematicKLine, ConceptuallyRelatedTo and *MotivationOf* are some of the relation types in ConceptNet. All those relation types are based on Minsky’s Theory about the mind [23].

After retrieving the relations from ConceptNet, PACO-T maps them to natural language and, then, presents it to teachers. Figure 4 exemplifies the mapping performed by PACO-T before showing the common sense relations to teachers.

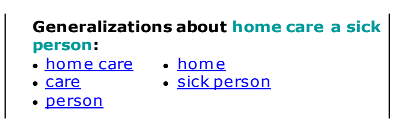


Fig. 4. SuperThematicKLine mapped to Natural Language

The relation type mapped in Figure 4 is the *SuperThematicKLine*. That relation type refers to the generalization/specialization principle and, therefore, the more abstract concepts are presented as generalizations for the more specific concept. In the example, the concepts "home care", "care", "person", "home" and "sick person" can be considered in some way more abstract that the concept "home care a sick person".

Note that the items in Figure 4 are presented as links. Clicking on the link, the system performs a new search in the ConceptNet, using as keyword the text of the link and updates the information in the common sense support box. This allows teachers to navigate among the concepts previously presented and to analyze how people with the same profile of their target group talk about related subjects. After the analysis, they can use the information gotten in one of the ways discussed in section 2. The common sense support box is always presented when teachers are asked to decide something based on this kind of knowledge.

By following the steps of the computational representation of PACO, answering the questions presented in each step and exploring the available common sense knowledge, teachers can plan their LAs taking into account pedagogical issues and fit it to their target group needs. It is worth pointing out that teachers can hide the common sense support box in any moment. At the end of the planning, teachers can export the LA plan to text format so that they can print it.

Figure 5, though in Portuguese, presents the same interface showed in Figure 3 after implementation. It can be observed that the implementation is in conformance with the design of the interfaces previous discussed. The interfaces are currently in Portuguese but it is intended to make possible their internationalization. Having this in mind, Java, JSF, Spring, Ajax and Hibernate have been used for the implementation, since they offer resources for easily internationalize interfaces.

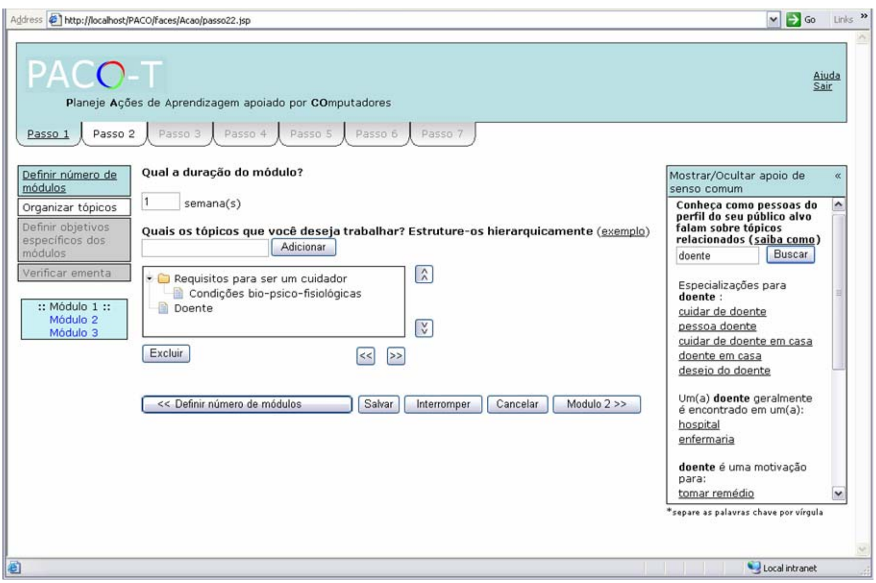


Fig. 5. PACO-T Functional Interface – Step 2 – Define and organize topics

4 Conclusions and Future Works

Contextualizing learning is one of the main pedagogical issues found in the literature and common sense knowledge can be used for this purpose. This paper presented the design process of a tool to support teachers in addressing such issue during the planning of LAs. The design was made considering requirements elicited from teachers by the conduction of a case study.

In the case study, PACO framework was used as the theoretical reference for supporting LA planning and common sense knowledge from the OMCS-Br knowledge base was made available for teachers to help them answering questions brought during the steps of PACO, so that the final LA was culturally contextualized to its target group.

The tool can make the task of planning LAs easier to teachers and to give them conditions to plan LAs that fit to their target group's needs and, therefore, to promote effective learning, putting in practice principle from renowned Learning Theories. In addition to that, the availability of common sense knowledge collected from people with the same profile of the teachers' target group is a useful tool for contextualization that still cannot be found in any other tools for planning LAs. Furthermore, interfaces designed were performed following the Web Design Pattern Language of Montero et al. [16], aiming to address usability. Making the tool available on the web, implementing it using technologies such JSF and keeping a simple language were an attempt of addressing accessibility issues.

As future work, it is proposed usability and accessibility tests using the tool in order to improve it, refining its interfaces to make it even more simple and intuitive to use. Still, it is proposed to implement accessibility issues to make the site the most compatible with assistive technology as possible.

Acknowledgements. We thank FAPESP and CAPES for partially support this project, Prof. Silvia Zem-Mascarenhas, Prof. Rosely M. de Figueiredo and all students from DEnf/UFSCar, who have participated on the learning activity planning and execution, the undergraduates Raphael da Silva Santos and Paulo Papotti, who has worked hard on the implementation of PACO-T, and Henry Lieberman from MIT MediaLab for the initial collaboration with the OMCS-Br project.

References

1. Conole, G., Fill, K.: A Learning Design Toolkit to Create Pedagogically Effective Learning Activities. Journal of Interactive Media in Education. Special Issue on Portable Learning, 1–16 (2005)
2. Anacleto, J.C., de Carvalho, A.F.P., Ferreira, A.M., Pereira, E.N., Carlos, A.J.F.: Common Sense-based Applications to Advance Personalized Learning. In: The 2008 International Conference on System, Man and Cybernetics (IEEE SMC 2008), pp. 1–10. IEEE Computer Society, Los Alamitos (2008)
3. Freire, P.: Pedagogia da Autonomia: Saberes Necessários à Prática Educativa, 31th edn. Paz e Terra, São Paulo (1996) (in Portuguese)
4. Freinet, C.: Education Through Work: A Model for Child Centered Learning. Edwin Mellen Press, New York (1993)
5. Ausubel, D.P.: Educational Psychology: A Cognitive View. Holt, Rinehart and Winston, New York (1968)

6. Gagné, R.M.: The Conditions of Learning, 3rd edn. Holt, Rinehart and Winston, New York (1974)
7. Lenat, D.B., Guha, R.V., Pittman, K., Pratt, D., Shepherd, M.: Cyc: Toward Programs with Common Sense. Communications of the ACM 33(8), 30–49 (1990)
8. Landauer, T.K.: How much do people remember? Some estimates of the quantity of learned information in long-term memory. Cognitive Science 10, 477–493 (1986)
9. Panton, K., Matuszek, C., Lenat, D., Schneider, D., Witbrock, M., Siegel, N., Shepard, B.: Common Sense Reasoning – From Cyc to Intelligent Assistant. In: Cai, Y., Abascal, J. (eds.) Ambient Intelligence in Everyday Life. LNCS, vol. 3864, pp. 1–31. Springer, Heidelberg (2006)
10. de Carvalho, A.F.P., Anacleto, J.C., Zem-Mascarenhas, S.H.: Learning Activities on Health Care Supported by Common Sense Knowledge. In: The 23rd Annual ACM Symposium on Applied Computing (SAC 2008), pp. 1385–1389. ACM Press, New York (2008)
11. de Carvalho, A.F.P., Anacleto, J.C., Neris, V.P.A.: Supporting Teachers to Plan Culturally Contextualized Learning Activities. In: Kendall, M., Samways, B. (eds.) Learning to Live in the Knowledge Society, IFIP International Federation for Information Processing, vol. 281, pp. 171–174. Springer, Boston (2008)
12. Anacleto, J.C., Godoi, M.S., de Carvalho, A.F.P., Lieberman, H.: A Common Sense-Based On-Line Assistant for Training Employees. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4662, pp. 243–254. Springer, Heidelberg (2007)
13. Anacleto, J.C., de Carvalho, A.F.P., Neris, V.P.A., Godoi, M.S., Talarico Neto, A.: How Can Common Sense Support Instructors with Distance Educations? In: The 2006 Brazilian Symposium on Informatics and Education (SBIE 2006), pp. 217–226. Brazilian Computing Society, Porto Alegre (2006)
14. Anacleto, J.C., de Carvalho, A.F.P., Pereira, E.N., Ferreira, A.M., Carlos, A.J.F.: Machines with Good Sense: How Can Computers Become Capable of Sensible Reasoning? In: Bramer, M. (ed.) Artificial Intelligence in Theory and Practice II, IFIP International Federation for Information Processing, vol. 276, pp. 195–204. Springer, Boston (2008)
15. Neris, V.P.A., Anacleto, J.C., Zem-Mascarenhas, S.H., De Carvalho, A.F.P.: PACO - A Framework for Planning Learning Activities Supported by Computers. In: The 18th Brazilian Symposium on Informatics in Education (SBIE 2007), pp. 597–606. Brazilian Computing Society, Porto Alegre (2007)
16. Montero, F., Lozano, M., González, P., Ramos, I.: A First Approach to Design Web Sites by Using Patterns. In: The 2002 VikingPLoP Conference (2002)
17. Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., Angel, S.A.: A Pattern Language: Towns, Buildings, Construction. Oxford University Press, New York (1978)
18. de Carvalho, A.F.P., Anacleto, J.C., Zem-Mascarenhas, S.H.: Planning Learning Activities Pedagogically Suitable by Using Common Sense Knowledge. In: The 16th International Conference on Computing (CIC 2007), April 11, 2009. IEEE CS Press, New York (2009), http://magno-congreso.cic.ipn.mx/CD-2007/IEEE/index.htm
19. Anacleto, J.C., Carlos, A.J.F., de Carvalho, A.F.P., Godoi, M.S.: Using Common Sense Knowledge to Support Learning Objects Edition and Discovery for Reuse. In: The 13th Brazilian Symposium on Multimedia and the Web (WebMedia 2007), pp. 290–297. Brazilian Computing Society, Porto Alegre (2007)
20. Snyder, C.: Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces. Morgan Kaufmann, San Francisco (2003)
21. Nielsen, J.: Designing Web Usability. Peachpit Press (1999)
22. Liu, H., Singh, P.: ConceptNet: a Practical Commonsense Reasoning Toolkit. BT Technology Journal 22(4), 221–226 (2004)
23. Minsky, M.: The Society of Mind. Simon and Schuster, New York (1988)

WIPDash: Work Item and People Dashboard for Software Development Teams

Mikkel R. Jakobsen[1], Roland Fernandez[2], Mary Czerwinski[2], Kori Inkpen[2], Olga Kulyk[3], and George G. Robertson[2]

[1] Department of Computer Science, University of Copenhagen, Denmark
mikkelrj@diku.dk
[2] Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA
{rfernand,marycz,kori,ggr}@microsoft.com
[3] Human Media Interaction, University of Twente, 7500 AE, Enschede, The Netherlands
okulyk@utwente.nl

Abstract. We present WIPDash, a visualization for software development teams designed to increase group awareness of work items and code base activity. WIPDash was iteratively designed by working with two development teams, using interviews, observations, and focus groups, as well as sketches of the prototype. Based on those observations and feedback, we prototyped WIPDash and deployed it with two software teams for a one week field study. We summarize the lessons learned, and include suggestions for a future version.

Keywords: Information visualization, software development, large display, cooperative work, CSCW, situational awareness, field study.

1 Introduction

Team collaboration and coordination in software development is difficult [14]. First, it may require frequent coordination to plan and review progress of a team. Second, completing a task often involves team collaboration because knowledge is divided between team members who have different roles or own different parts of the system. Team members may work on multiple task items at a time, or belong to more than one team, adding to the challenge of coordination. Thus, team members need to be aware of what others on the team are doing [10,14].

In this paper, we present WIPDash (Work Item and People Dashboard), a visualization of work items in a team's software repository. Our goal is to help software teams be aware of the overall status of a project, and understand ongoing activities related to the team. Initially, we conducted interviews and field observations within a software development organization in order to understand the needs of collocated software teams. We then discussed these results in a series of focus group to iteratively design WIPDash. Finally, we deployed WIPDash with two software teams in an attempt to observe which features and functions the team actually used, and how they used those features. Our findings led us to a number of design lessons, and yet another design iteration, which we introduce at the end of the paper.

The main contributions of this work include (1) detailed findings about how developers maintain team awareness using existing techniques and tools, (2) a novel

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 791–804, 2009.

awareness visualization based on developers' needs, and (3) lessons learned from a deployment with two teams along with a conceptual overview of new design ideas based on that deployment.

2 Supporting Software Team Awareness

Many organizations adopt Agile Software Development methodologies that promote shorter iterations and daily stand-up meetings to improve coordination [15]. Also, collocation of a team in a shared team space may improve productivity [19]. Yet, software teams are still challenged with maintaining awareness of ongoing activity [10], and find collaborative tools that support their development work useful [14].

Software teams typically store information about work items (e.g., tasks and bugs) in software repositories, such as Microsoft Team Foundation Server (TFS), to support team coordination. However, such repository systems are not designed to give an overview of the state of a project or to keep team members aware of the team's current activities. In addition, it is not easy to see changes to work items in a software repository. Developers may not feel that they get a proper return on the time they invest on updating work items and often the status of work items is not up to date.

One way to improve team awareness is to show data from a team's repository on a large display in a shared workspace [e.g., 1,8]. FASTDash [1] showed developers' current activities in a code base. A field study showed that FASTDash increased communication within the team by 200%, and helped participants know who had which files checked out, who was blocked and needed assistance, and helped resolve conflicts with checked out code. Improvements were suggested based on the study, such as using metrics other than file size to allocate screen space, and to add support for people to track work items that are assigned to them. Still, it was not clear which types of information were most useful or how visualizations could be best designed to get awareness information at a glance. O'Reilly et al. [12] visualized checked-in code changes on a multi-monitor display. The authors concluded that the visualization helped to inform developers about progress and overall effort of the team. However, it is not clear how participants used the display. De Souza, Froelich and Dourish [16] have shown that source code could be mined to visualize both social and technical relationships of projects. We were inspired by these findings and focused our visualization on support for work item awareness.

Fitzpatrick et al. [6] described a long-term study of a software team using a tickertape tool where messages from CVS, a revision control system, were displayed. The authors found that the tickertape tool stimulated more focused discussion about source code changes, reduced the number of empty check-in messages, and helped coordinate and negotiate work within the team. The authors mentioned the modest screen real estate requirements of the tickertape tool as an important benefit.

Hill and Holland [9] describe the concept of showing the history of a user's interactions with files as part of the representation of the files. Recent research has empirically studied use of interaction history to help software development teams [3,7]. TeamTracks [3] directs the attention of a programmer to important parts of the source code based on the history of programmers' interactions with the code. Augur [7] combines information about code activity with a line-oriented source code visualization similar to SeeSoft [4]. Froehlich and Dourish [7] present case studies of four developers who used Augur to gain insight into their code and their development activities. Their findings support the idea of combining information about activity and code in one view that is based on spatial organization of the code. However, Augur's potential usefulness as an awareness tool for a collocated team remains unknown.

The research efforts mentioned so far involved representations of source code, check-ins, and the use of code files. In contrast, Ellis et al. [5] aimed at helping large distributed software teams to coordinate their work on change requests by visualizing bugs. They presented SHO, a visualization with bugs shown as circles ordered, colored, and sized by different importance metrics. Participants in an experiment were more successful at completing tasks using SHO than using Bugzilla. Another recent study by Sarma et al. [14] presented a desktop awareness system based on code activity and check-ins. Their Palantír tool addressed mainly artifact changes in order to prevent potential conflicts. The visualizations were useful for exploring databases of bugs to identify areas of concern. However, it is not clear how useful these types of visualizations are for maintaining awareness of work items activity and project progress in collocated software teams, which is the focus of this paper.

3 User-Centered Design

The research presented in this paper follows a user-center design approach. We began by gathering observations and conducting semi-structured interviews with software developers to gain insight into their work practices and needs in collocated team workspaces. Our goal was to understand how collocated teams coordinate their work and to get input on what visualization features and views would be most useful for them. We then sketched an initial design of a visualization to support team awareness, which we presented to a focus group for feedback.

3.1 Interviews and Field Observations

In situ observations were performed with two Agile teams (see Figure 1). The teams were observed for five hours during one week. Observations were carried out at different times of the day and included morning stand-up meetings and iteration planning meetings. One team typically had five to ten members present, while the other team typically had eight to fifteen members present. We also carried out semi-structured interviews with ten individuals from these teams (eight males). Each interview lasted 35-40 minutes and was audio recorded with the participants' permission. Interviewees received a free lunch coupon as gratuity for their participation. Each interviewee had two to fifteen years of experience in software development and ranged in age from 20 to 46. Based on questions from related studies [10,17] and questions motivated by our in situ observations, the interviews focused on the following aspects:

- experiences working in a collocated team workspace;
- what tools and alerts are currently used to keep track of what other team members are working on and what is missing in existing tools;
- how work items and tasks are currently managed;

Fig. 1. Shared team rooms for the two software teams we observed

- types of meetings and the use of a projector in the team room;
- how progress and project health in general is monitored;
- wishes on what to display on the large shared screen and how to support work flow.

3.2 Interview and Field Observation Results

The interviews and observations showed that the teams work in iterations, which are blocks of time typically one to two weeks long. Daily stand-up meetings in the morning help the team to keep track of who works on what, and they are considered an important time for team bonding. Work items are created in a repository.

The teams used many software tools to coordinate their activities including Team Foundation Server (TFS), email, instant messaging, Live Meeting, and SharePoint. The teams varied in work style, size, and physical workspace arrangement. Both teams adopted a seating arrangement that corresponded to individual roles on the team: developers, testers, writers, and support (including program managers).

In addition to regular stand-up meetings, iteration planning, and bug triages (where resolved bugs are discussed and new bugs are assigned to team members), ad-hoc conversations frequently occurred in the team rooms. Some team members used chat or email, but often team members just shouted out a question or rolled their chairs over to talk to each other. Moreover, other people came in and out of the team spaces.

Although shared team rooms can be noisy and distracting, and offer less privacy than private offices, most team members felt that the team room was more effective for team work. An exception to this is documentation writers (two out of ten interviewees) who said that they preferred to work in a private office or from home; they needed to concentrate and only came to the team room for meetings or when they needed to speak to a team member. Since they were frequently absent from the team room they tended to be less aware of activities that were going on within the team.

Both teams used shared whiteboards and sticky notes on the walls. Team members defined, categorized, and prioritized work items during iteration planning meetings and used sticky notes to represent tasks or work items. The teams also used a projector on the wall to display information for various tasks such as work items during iteration planning, when assigning new tasks, or for code reviews.

The dynamic nature of stand-up meetings requires a quick, glanceable overview of recent activity. Teams currently do not have a suitable tool for displaying important information. TFS gives no overview of unassigned tasks, and does not allow more than one person to be assigned to a task. Also, team members have to make a burn down chart or a task list for each meeting, which is time consuming. It is possible to export charts and work items from TFS, but this often results in a long Excel table with no way of synchronizing changes back with work items in TFS.

3.3 Focus Group Feedback

Based on what we learned from the interviews and observations, we sketched an initial design for a team awareness visualization. This initial design sketch, based on the current iteration of one of the team's work items repository, was presented to a focus group on a large projection screen in a meeting room to help participants imagine how the visualization would look when deployed in their own team room. Eight participants from two Agile development teams took part, including an architect, program managers, developers, lead developers and testers ranging in age from 30-48. The focus group session was video recorded with the teams' permissions.

Participants found that the awareness display presented information in a new perspective they had not seen before and they liked being able to see an overview of the whole project in one view. Team members expressed the need for different filters and view modes, since some work items might be irrelevant for their role.

We derived three key requirements for our design based on this feedback. The awareness visualization should (1) give an overview of iteration progress, with the ability to summarize over the last day, week, month, or version, (2) give details on individual work items and the people these are assigned to, and (3) list current and recent activities, either of people or on work items.

4 Work Item and People Visualization

Based on all of these results, we developed WIPDash, a visualization suitable for a large shared display in a collocated software team space. The visualization was implemented as a Windows application that reads data about work items from TFS. Our intention was that team members could glance at the shared display to see the overall status of the project and the recent changes made (especially from the past 24 hours). We also wanted team members to be able to use WIPDash on their individual machines, where they could switch between different views and filters, and get details on demand.

The WIPDash window consists of two parts (see Figure 2). The left part of the window contains a spatial representation of areas of the project and the work items in those areas. For instance, the area labeled 'Docs' contains items related to project documentation. The right part of the window consists of a list of view modes, a team panel, and drop-down lists of iterations, work item states and types. These lists can be used to filter and highlight work items in the left view.

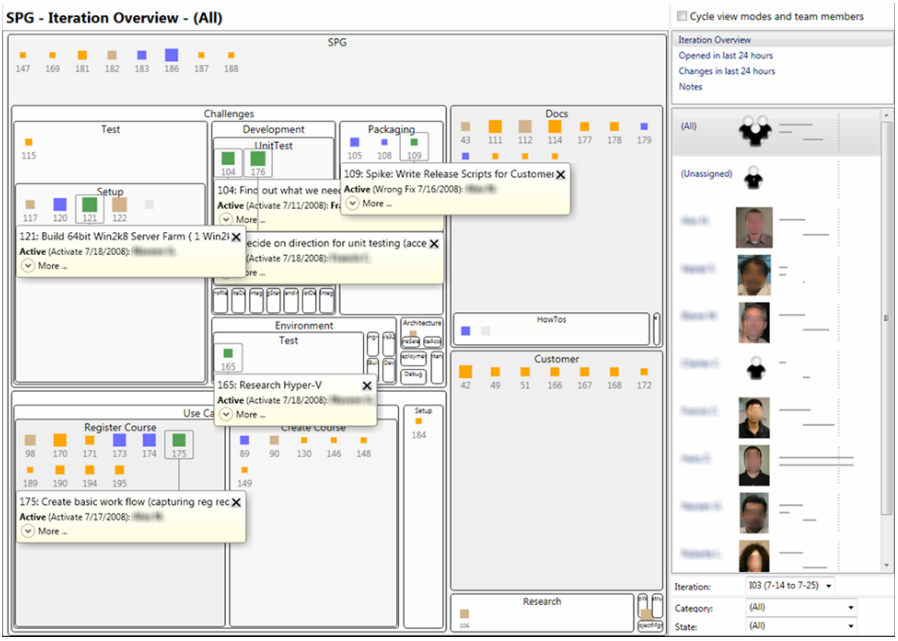


Fig. 2. WIPDash showing the iteration overview for the current iteration of a project

4.1 Work Item Treemap

WIPDash uses a squarified treemap for laying out project areas as rectangles [2]. The treemap is a scalable approach to spatially organizing hierarchically structured data such as a hierarchy of project areas. Each rectangle is sized proportionally to the number of open work items in the area. A minimum threshold is used to ensure that areas that do not contain any open items are shown in the map. Rectangles are labeled with the name of the project area.

We wanted to preserve the spatial layout of project areas and work items to make it easier for users to remember where areas are located in the visualization. However, treemap algorithms can cause the spatial layout to change considerably when the data changes. Since WIPDash would be shown on both a large display and on individual team members' displays (which may have different screen dimensions), the layout had to vary across instances of WIPDash. In order to keep the layout consistent for the purpose of the field study, the treemap was fixed and then shared by all instances of WIPDash. This layout could be explicitly updated and the treemap would render again. Since the relative size of an area does not change dynamically to reflect the number of open work items, we color a rectangle darker as more open work items are associated with the area.

Each rectangle in the treemap contains icons that represent the work items associated with that area. The icons in a rectangle are evenly spaced in a grid, placed in the order they were created starting from the top-left corner. Space between icons is reduced to fit all icons within the rectangle, and if there is enough space between icons, the ID number of the item is shown below the icon. The color of an icon indicates the state of the work item (e.g., proposed, active, resolved, or closed) and the shape of an icon indicates the type of the work item (e.g., feature, bug, or task). Icon size represents either priority level or estimated hours of the work item, with larger sized icons representing items of higher priority or higher estimate of work hours, as designated by the user or team.

Moving the mouse cursor over a work item icon shows a tooltip with details about the item. Clicking on an icon shows a popup window with details about the work item. An 'Add note' section in the detail window can be expand to add a sticky note to the work item. A small yellow sticky note symbol is displayed on the work item to indicate that it has a note attached.

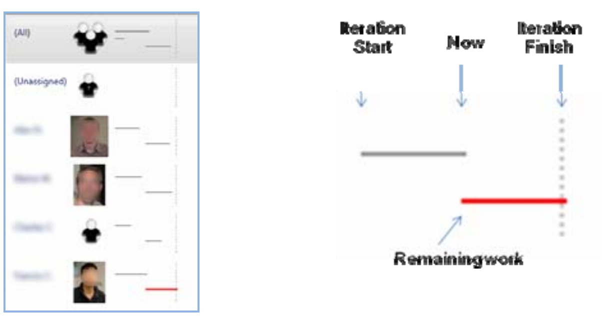
4.2 Iteration Filtering and Highlighting

Selecting an iteration in the "iteration list" shows all work items that are assigned to that iteration and highlights them on the treemap. Since teams are usually only interested in closed items for the current iteration, we removed closed items that were not assigned to the selected iteration in order to avoid clutter.

4.3 Team Panel

The team panel (see Figure 3a) contains the names and pictures of the team members. Clicking on a team member shows the icons for all work items assigned to or closed by the selected team member. The team panel contains two options in addition to the team members: (1) "all" which is used to select all items regardless of whom they are assigned to, and (2) "unassigned", which is used to select all unassigned work items.

For each person, WIPDash shows horizontal lines that represent the total amount of work, the amount of completed work, and the work remaining in the iteration that is assigned to that person (Figure 3b). The x-axis measures work hours and a dotted vertical line represents the end of the iteration, corresponding to the total number of work hours in the iteration. The remaining work line is colored red if the estimated hours of remaining work exceeds the time left in the iteration.



(a) (b)

Fig. 3. (a) Team panel showing names and pictures of team members, and (b) a graphical representation of the amount of work each member has assigned, completed and remaining

4.4 View Modes

One goal of our design was to make the information on the display glanceable. Thus, to avoid cluttering the display by showing too much information in one view, users can choose between four different view modes.

The *Iteration overview* mode highlights all work items in the current iteration. This view aims to provide an overview of the iteration status. Team members can see how much work has been done and how many work items remain in the iteration. Details are automatically displayed for work items that are currently being working on, including who is working on the item. More details about an item can be shown by clicking on '*More*' (see Figure 2).

The *Opened in last 24 hours* view is designed to keep the team aware of incoming tasks and issues. Work items opened within the last 24 hours are highlighted with a yellow border and background. The opacity of the border and background varies to distinguish recently opened items from items that were opened less recently. Also, similar to the iteration overview, details are shown for recently opened items, including when an item was opened and by whom.

The *Changed in last 24 hours* view aims to keep the team aware of recent changes to work items. Similar to the *Opened* view mode, a yellow border and background is shown around icons of work items that changed within the last 24 hours. Again, details are automatically displayed for recently changed work items, including the change made and who made the change. For simplicity, the same color coding is used for both *Opened* and *Changed* view modes in order to draw users' attention to work items with any recent activity.

The *Notes* view calls up the sticky notes for all work items with a note attached. The Notes view allows users to spot work items that need attention, for example if a team member makes a request to pair up on a specific task.

In order to provide continuous awareness and to allow passive use, WIPDash allows cycling through view modes and through team members within each view mode. A person remains selected for ten seconds before the next person on the team panel is selected. The next view mode is selected after cycling through all team panel selections. We were interested to see if this cycling behavior was useful or distracting to the teams we studied. The visualization updates with new or changed work items by querying Team Foundation Server (TFS) once per minute. All of the information used in our visualization came from the teams' data entries in TFS.

5 In Situ Deployment

We deployed WIPDash with two collocated Agile software teams and observed its use for one week. Our aim was to understand the usefulness of the WIPDash and the effect it had on team members' situational awareness and on group processes. Team A had eight members (seven male), with an age range of 22-46. Team B had 16 members (14 male), with an age range of 27-48. Individual roles on both teams included lead developer, developer, tester, test lead, program manager, writer and group manager. WIPDash was installed on a large display in the team rooms and on team members' individual workstations. Data were automatically collected in WIPDash in order to describe how participants interacted with the visualization throughout the study. WIPDash was installed on a Thursday and an orientation session was given the following day. We observed the teams the following week (Monday through Friday). Afterwards, we met with each team for post-usage discussions. Each participant received a $50 gratuity coupon for their participation. In the Team A room, the awareness visualization ran on a projected wall display on the most accessible wall to the whole team. In the Team B room, the awareness display was installed on a 52-inch plasma touch screen toward the front and right side of the team room. Some members of Team B had their backs to this display.

5.1 Supporting Daily Stand-Up Meetings

Team A used WIPDash daily on their large display during stand-up meetings to coordinate meetings. Specifically, they found the *Iteration overview* useful, both in terms of status and also to jog their memories about work items from the previous day. During stand-up meetings, one of the team members selected each person from the team panel to display his information, and that person then talked about his work. The team considered details about the active items assigned to a person especially useful. Team members said they would have found it beneficial to have team members displayed in a random order during stand-up meetings—just to make it more fun. They also referred to the display to view the status of a remote team member when she called in for the standup meeting. This suggests that WIPDash could be useful for supporting collaboration with distant team members. Some team members commented that they liked to look at the awareness display first thing in the morning to see what team members in Argentina had been doing for the last ten hours. Some team members who were on vacation for most of the time during our study also said that they used the awareness visualization when they got back to get a sense of what the team had been doing for the past week and "*Where are we now*?" in the iteration.

We observed that the available information in WIPDash was not completely sufficient for reviewing what had been worked on during the previous day. Specifically, if a developer had completed his work on a work item and then reassigned the item to somebody else for testing, that item no longer showed in the *Changed* view for that developer. This was discussed during a stand-up meeting, and the team suggested a view where all work items that a person had worked on would be highlighted, even if they had been reassigned. The team further elaborated on an idea of one, concise overview containing all the information they would need for their standup meetings, including recently resolved items that were reassigned to other members and items that members had worked on yesterday. Finally, the team expressed a wish for extending the shared wall display with an additional projector to show project information like spreadsheets or code next to WIPDash.

5.2 Automatic Cycling

The automatic cycling between views was found to be problematic. WIPDash cycled between all members in the team panel, including people with no items assigned or without any recent activity. Thus, nothing of interest is shown in parts of the cycle. Members of Team A suggested that cycling would make sense if done only between views that contain recent changes. Team B had only one work item assigned per person at a time and therefore cycling through each view for each team member was not useful.

Team members commented that notifications needed to be more assertive when a change or an update happened, such as an audio herald combined with a fisheye notification message about the change on the awareness display. Also, team members wanted to configure which views were displayed on their personal displays and when notifications should appear. An RSS-style feed would probably be a useful option for the personal workstations. We are pursuing that idea in the next iteration.

5.3 Use on Large Display and Individual Displays

We analyzed data logged by WIPDash to see how the two teams used WIPDash on the large display and on their individual machines. In all, we collected log data from ten personal machines, five members on Team B and five members on Team A, in addition to the two machines running the large shared displays. When a user started interacting with WIPDash, the default cycling between views was suspended. The log data showed that Team A used the large display (showing the iteration overview) during their stand-up meeting every morning around 9:30 am. Apart from the stand-up meetings, Team A physically interacted with the large display only twice during the study. In contrast, Team B interacted with the large display on average five times per day. In Team B some of the participants did not have WIPDash installed on their own machines. For this team WIPDash was running on a new touch screen display, which had a lot of appeal. One possible reason why Team A did not interact much with the large display was that it was projected high up on one of the walls, making personal display interaction more reasonable than going to the laptop in the corner of the room that controlled the view in order to interact with the group view.

5.4 Types of Interaction

In all, 596 selections were made in the right panel of WIPDash. Selections were primarily made in the team panel (66%). The view mode panel (17%), and iteration panel (15%) were also used to change the view. *Iteration overview* was the most frequently selected view mode, selected more than 50% of the time. The *Changes* view and the *Opened* view were each selected between three and eight times, whereas the *Notes* view was only selected once.

Participants from Team A clicked on items to call up details 21 times while participants from Team B brought up additional details 22 times. Only two sticky notes were created in WIPDash, both by a lead developer on Team A. Follow-up discussions related to sticky notes revealed that users would rather use the existing ‘comments’ data structure in TFS to view and add notes to work items in WIPDash.

5.5 Questionnaires

Satisfaction with WIPDash was assessed using a questionnaire administered to team members at the post-usage meeting. The questionnaire contained nine questions from [13] and eight questions to address distraction and awareness, using five-point Likert-scale with lower scores reflecting negative responses. We balanced the valence of our

Table 1. User Satisfaction

	Question	Average Rating (SD)
1.	I have difficulty understanding WIPDash.*	4.10 (0.7)
2.	WIPDash is easy to use.	3.63 (0.8)
3.	WIPDash is reliable.	2.63 (1.1)
4.	I have confidence in the information provided by WIPDash.	3.90 (0.7)
5.	I need more training to understand WIPDash.*	4.05 (0.9)
6.	WIPDash is informative.	3.32 (0.9)
7.	WIPDash is comprehensible.	3.37 (0.8)
8.	Overall, I am satisfied with WIPDash.	3.00 (1.1)
9.	I would be happy to use WIPDash in the future.	3.11 (1.1)
10.	I find WIPDash distracting.*	3.95 (0.9)
11.	WIPDash grabs my attention at the right time.	2.84 (0.9)
12.	It's worth giving up the screen space to run WIPDash on my PC.	2.74 (0.9)
13.	WIPDash helps me stay aware of information that's critical.	2.72 (1.0)
14.	I like being notified when a work item gets reassigned.	3.22 (1.2)
15.	WIPDash's notifications often distract me.*	3.83 (0.8)
16.	Having WIPDash displayed in front of the team is embarrassing.*	4.17 (1.0)
17.	I would rather have WIPDash displayed only privately.*	4.18 (1.1)

satisfaction questions. For negatively phrased questions (marked with an asterisk in Table 1), we reversed the rating so that higher was always positive.

Ratings on usefulness and satisfaction with the system were mostly neutral to positive (see Table 1) and there were no significant differences between the two teams (paired t-test). Team members said that they had confidence that the information was displayed correctly and on time. From the ratings, it was clear that the notifications of changes were not grabbing attention well enough, and that WIPDash was seen as less reliable than we would have liked. However, the teams were not embarrassed to show their personal work item information and they leaned towards using a future version of WIPDash.

A questionnaire on situational awareness and group satisfaction was administered to team members before in situ deployment and at the post-usage meeting. The questionnaire included questions from [11,18] and used five-point Likert-scale with lower scores reflecting negative responses. Paired t-tests showed no significant differences between the Before and After conditions. The results could be affected by a particular iteration stage or the day of the week the questionnaire was completed.

6 Discussion, Lessons Learned and Conceptual Redesign

Several factors may have affected the use and adoption of WIPDash. First, size and location of the shared display affected how team members made use of the display by glancing at it or physically interacting with it. During observations of Team A, we saw that team members often looked at the display when entering or leaving the room. This was not the case in the Team B room. One reason may be that the Team A room had a large display, projected high up on a wall, that was visible to everyone in the room. In contrast, Team B's shared display was smaller, located in a corner of the room, and was not directly visible to all team members.

Second, the two teams in our study organized their work differently. Team A assigned work items to team members during an iteration planning meeting, and had daily stand-up meetings to follow up on progress of the team. Using the shared display during stand-up meetings may have influenced Team A's familiarity with the display, and consequently increased their use of the display. In contrast, Team B did not have daily stand-up meetings; instead, they talked with each other about progress and status throughout the day. Also, Team B assigned only one work item to each person and viewed work in terms of releases that span several iterations, not single iterations. This suggests that Team B could make even better use of an overview of activities, but using different time spans.

Third, Team B had many proposed items in their repository, but most items were not scheduled to be worked on. Also, many of the project areas, which were shown in WIPDash because they contained proposed work items, were simply not considered relevant by the team at the time of our observations. Thus, as Team B had many more items to track than did Team A, and since WIPDash showed all work items in 'proposed' state, the display for Team B was more cluttered.

Our study has limitations that should be considered when interpreting the results. Although very different work styles were observed, the two teams were from the same organization. Therefore, our results may not be generalizable to software teams everywhere. Also, we only observed the teams for one week. While we learned much from the initial feedback, it would be interesting to see the long term effects of WIPDash. Another concern is that we do not know how the collocated Agile teams that we have focused on compare with larger, distributed teams. It could be that a focus on collocated, Agile software teams may reveal some issues in team coordination that also apply to distributed software teams. For example, Gutwin et al. [8] suggest a need for awareness of areas of expertise within a distributed open source project team because developers work on all parts of the code. This might relate to the information needs we seek to provide with our visualization (e.g., what people are working on and have been working on). For example, a glance at WIPDash first thing in the morning to see what team members in Argentina have been doing for the last ten hours could be very useful in terms of setting daily priorities or offering more timely assistance. We intend to extend our focus to distributed team awareness in our future work.

Informed by the insights we have gained from our study of WIPDash, the design of the next version of the dashboard will include the following:

- A more glanceable display;
- sound cues for users to look at the dashboard when information changes;
- on a user's display, it will only show notifications (e.g., 'toaster' type alerts in the system tray) as information changes;
- a wide diversity of project data field definitions and usages;
- support for add-on visualizations developed by a dashboard community;
- data caching to reduce the load on the repositories from dashboard clients.

7 Conclusion and Future Work

This study on WIPDash suggests benefits from providing awareness of work item status. Earlier work on FASTDash [1] showed benefits from providing a team situation awareness display based on code activity. An interesting perspective for future work is to combine information about work items with information about code activity in one visualization [14]. A potential disadvantage of such an approach is that the display gets too busy. However, linking code activity and the state of work items could give the team a solid shared context if it focused simply on what is or has just recently changed. That is the main goal for our next iteration. We intend to create a simple list of recently changed work items, and the people who are actively related to them. In addition, the new WIPDash will have integrated chat and RSS feeds for notifications to the desktop. The ethnographic study of Souza and Redmiles [17] confirms our observations: (a) a need for awareness of who made the changes; (b) a need to peripherally integrate awareness notifications into an existing work items repository, and to link them to code changes in order to keep work items' status better up to date and thus coordinate work more proactively.

In this paper, we have reported on a human-centered design approach to developing a situational awareness dashboard visualization to help software development teams track people and work items. From observations and interviews of development teams we learned about their current work practice and what might be provided to improve situation awareness. We then developed a dashboard for supporting awareness in teams, called WIPDash (Work Item and People Dashboard). We gathered feedback on an initial design from a focus group, which drove the detailed design and the implementation of WIPDash. Finally, we have studied the use of WIPDash in situ with two development teams, and reflected on the observations and data we gathered. While questions remain to be answered, the results from our study provide initial insights about use of a shared display to support team awareness of work item data in software repositories.

References

1. Biehl, J.T., Czerwinski, M., Smith, G., Robertson, G.G.: FASTDash: a visual dashboard for fostering awareness in software teams. In: Proc. CHI 2007, pp. 1313–1322. ACM, New York (2007)
2. Bruls, M., Huizing, K., van Wijk, J.J.: Squarified Treemaps. In: Proc. TCVG 2000, pp. 33–42. IEEE Press, Los Alamitos (2000)
3. DeLine, R., Czerwinski, M., Robertson, G.: Easing program comprehension by sharing navigation data. In: Proc. VL/HCC 2005, pp. 241–248. IEEE Computer Society, Los Alamitos (2005)
4. Eick, S.C., Steffen, J.L., Sumner Jr., E.E.: Seesoft - A Tool for Visualizing Line Oriented Software Statistics. IEEE Trans. on Software Engineering 18(11), 957–968 (1992)
5. Ellis, J.B., Wahid, S., Danis, C., Kellogg, W.A.: Task and social visualization in software development: evaluation of a prototype. In: Proc. CHI 2007, pp. 577–586. ACM, New York (2007)
6. Fitzpatrick, G., Marshall, P., Phillips, A.: CVS integration with notification and chat: lightweight software team collaboration. In: Proc. CSCW 2006, pp. 49–58. ACM, New York (2006)
7. Froehlich, J., Dourish, P.: Unifying artifacts and activities in a visual tool for distributed software development teams. In: Proc. ICSE 2004, pp. 387–396. IEEE, Los Alamitos (2004)

8. Gutwin, C., Penner, R., Schneider, K.: Group Awareness in Distributed Software Development. In: Proc. CSCW 2004, pp. 72–81. ACM, New York (2004)
9. Hill, C.W., Hollan, J.D., Wroblewski, D., McCandless, T.: Edit wear and read wear. In: Proc. CHI 1992, pp. 3–9. ACM Press, New York (1992)
10. Ko, J., DeLine, R., Venolia, G.: Information Needs in Collocated Software Development Teams. In: Proc. ICSE 2007, pp. 344–353. IEEE Computer Society, Los Alamitos (2007)
11. Olaniran, A.: A Model of Group Satisfaction in Computer Mediated Communication and Face-to-Face Meetings. Behaviour and Information Technology 15(1), 24–36 (1996)
12. O'Reilly, C., Bustard, D., Morrow, P.: The war room command console: shared visualizations for inclusive team coordination. In: Proc. SoftVis 2005, pp. 57–65. ACM, New York (2005)
13. Paul, S., Seetharaman, P., Ramamurthy, K.: User Satisfaction with System, Decision Process, and Outcome in GDSS Based Meeting: An Experimental Investigation. In: Proc. HICSS 2004, vol. 1. IEEE Press, Los Alamitos (2004)
14. Sarma, A., Redmiles, D., van der Hoek, A.: Empirical evidence of the benefits of workspace awareness in software configuration management. In: Proc. SIGSOFT 2008/FSE-16, pp. 113–123 (2008)
15. Schwaber, K., Beedle, M.: Agile Software Development with Scrum. Prentice Hall, Englewood Cliffs (2002)
16. de Souza, C., Froehlich, J., Dourish, P.: Seeking the source: Software source code as a social and technical artifact. In: Proc. GROUP 2005, pp. 197–206. ACM, New York (2005)
17. de Souza, C.R., Redmiles, D.F.: An empirical study of software developers' management of dependencies and changes. In: Proc. ICSE 2008, pp. 241–250. ACM, New York (2008)
18. Taylor, R.M.: Situational Awareness Rating Technique (SART) The Development of a Tool for Aircrew System Design. In: Proceedings of the Symposium on Situational Awareness in Aerospace Operations, AGARD-CP-478 (1989)
19. Teasley, S., Covi, L., Krishnan, M.S., Olson, J.S.: How does radical collocation help a team succeed? In: Proc. CSCW 2000, pp. 339–346. ACM, New York (2000)

CGD – A New Algorithm to Optimize Space Occupation in Ellimaps

Benoît Otjacques[1], Maël Cornil[1], Monique Noirhomme[2], and Fernand Feltz[1]

[1] Public Research Center – Gabriel Lippmann
Department ISC – Informatics, Systems and Collaboration
41, Rue du Brill
L-4422 Belvaux, Luxembourg
otjacque@lippmann.lu, cornil@lippmann.lu, feltz@lippmann.lu
[2] University of Namur (FUNDP)
Computer Science Institute
21, Rue Grangagnage
B-5000 Namur, Belgium
monique.noirhomme@info.fundp.ac.be

Abstract. How to visualize datasets hierarchically structured is a basic issue in information visualization. Compared to the common diagrams based on the nodes-links paradigm (e.g. trees), the enclosure-based methods have shown high potential to represent simultaneously the structure of the hierarchy and the weight of nodes. In addition, these methods often support scalability up to sizes where trees become very complicated to understand. Several approaches belong to this class of visualization methods such as treemaps, ellimaps, circular treemaps or Voronoi treemaps. This paper focuses on the specific case of ellimaps in which the nodes are represented by ellipses nested one into each other. A controlled experiment has previously shown that the initial version of the ellimaps was efficient to support the perception of the dataset structure and was reasonably acceptable for the perception of the node weights. However it suffers from a major drawback in terms of display space occupation. We have tackled this issue and the paper proposes a new algorithm to draw ellimaps. It is based on successive distortions and relocations of the ellipses in order to occupy a larger proportion of the display space than the initial algorithm. A Monte-Carlo simulation has been used to evaluate the filling ratio of the display space in this new approach. The results show a significant improvement of this factor.

Keywords: Information Visualization, Ellimaps, Hierarchies Visualization.

1 Introduction

Generic data structures are often taken as a basis to classify visualization techniques and hierarchies are considered as one of these reference structures. They are encountered in various domains like botany, management or computer science. In some cases we can observe the importance to visualize not only the data structure itself but also some attributes of the nodes. For instance, the organization chart of a company can also show the number of employees in each division and service. As another example,

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 805–818, 2009.

the administrators of a large data storage system can be equally interested by the structure of the repository, the size of each folder and the last access date of each file.

This paper explores how to visualize weighted hierarchies and tries to find a reasonable balance between competing constraints: a good perception of the data structure, a visual representation of the node weights and an efficient occupation of the display space. It is structured as follows. After the introduction, the state-of-art is discussed and some weaknesses of current propositions are pointed out. Next a new algorithm aiming to tackle a pending issue (i.e. low space occupation) is described. Then we discuss the results of an empirical evaluation of this new approach. Finally some conclusions are drawn and some paths for further research are highlighted.

2 State of Art

Grouping is acknowledged as a central issue in perception since the seminal work of Gestaltist psychologists in the 1920's. More recent works have added new principles of grouping to the initial list. Among those principles, two are especially relevant in our study. *Connectedness* expresses the fact that drawing lines among the items of a set is a powerful mean to show that some relationships exist among them [11]. *Common region* is the tendency to group together the elements that lie within the same bounded area [10]. This principle explains, for instance, the success of Venn and Euler diagrams to represent a set of elements (cf. [17] pp. 194-196).

It appears that most of the techniques to visualize hierarchies rely on the *common region* and / or the *connectedness* principles.

In this specific context, *connectedness* means representing nodes by punctual visual objects (e.g. points, icons) and the hierarchical relationship by lines (e.g. straights, curves). The various versions of trees are typical examples of this approach. This class of techniques highlights well the structure of the dataset and makes it very easy to understand. Unfortunately, they do not support very well scalability (cf. [2] p. 149). If the number of nodes grows their layout rapidly becomes very difficult to understand and the navigation tasks become challenging. However, considering their power to visualize the dataset structure, they cannot be neglected and numerous researchers have proposed some improvement to the basic diagrams, such as using polar coordinates [3], hyperbolic geometry [4] or patterns based on circles [6].

For visualizing hierarchies, the *common region* principle materializes by nesting shapes into each other. The nodes are represented by rectangles [13], ellipses [8], complex shapes [1] or even distorted shapes [16]. The hierarchical relationship is represented by the successive inclusion of these shapes. The treemaps [13] are a well-known example of this paradigm and have been by far the most investigated (see [14] for an historical review). They support very well scalability and they have proven to be able to display thousands of nodes in a single screen. They also show the relative weight of the nodes, which allows to rapidly identify the preeminent ones. Unfortunately, treemaps must also acknowledge some limitations. Lee et al. [5] explain that '*treemaps are appropriate when showing the attribute value distributions is more important than showing the graph structure*'. In order to tackle the issues associated to the initial treemap algorithm, several researchers have proposed new strategies to dimension and position the nested rectangles (e.g. [15] for improving the perception of the data structure, [18] for the ordering consistency).

Ellimaps [8] are more recent and have been significantly less studied (e.g. [9] for an example of use for monitoring web-based platforms). Basically they are founded on layout algorithms similar to the ones used in treemaps. Nevertheless, they use nested ellipses instead of nested rectangles and this offers some new perspectives as well as new challenges. A previous study [8] has shown that (under the conditions summarized hereafter) the ellimaps provide a better support to the perception of the hierarchical structure than squarified treemaps. In addition, they received better score for the subjective perception of the test users. These findings are based on a controlled experiment with 34 subjects who were asked to answer questions like identifying the sibling nodes of a given node N, comparing the weight of several nodes or finding the most weighted node among the children of a node N (cf. *objective measures: time to complete, success rate*) and to fill in a satisfaction questionnaire (cf. *subjective measures: rating of items on a Likert-scale*). The experiment was carried out with sample hierarchies having around 500 nodes and 8 levels of depth.

The authors of the ellimaps acknowledge, however, that the rather poor occupation of the display space is a major drawback of their technique. This limitation raises some issues concerning the scalability of the initial version of this technique. We have therefore decided to tackle this issue in order to reach a better balance between the perception of the dataset structure and the occupation of the display space.

At the end of our literature review we must also point out that beside the techniques that clearly rely on one of the above-mentioned Gestalt principles, we can also find some hybrid approaches. The Space-Optimized Tree [7] uses a rule based on *common region* to divide the display space but uses nodes and links to show the items and their relationships. Another original approach is proposed by Schultz et al. [12] who obtain a space-filling effect with a point-based rendering technique.

3 New Algorithm: Combined Geometrical Distortions

3.1 Formalization of the Problem

Our research problem consists in finding how to draw the n ellipses E_i corresponding to the n children N_i into the ellipse E* representing their parent node. It can be formalized by three conditions.

- Including the n ellipses E_i into the ellipse E* $(1 \leq i \leq n)$;
- Keeping the ratio of the areas of the ellipses E_i equal to the ratio of their corresponding node weights N_i;
- Occupying a larger proportion of the display space than the initial ellimap algorithm.

Considering the third condition we first need to compute the occupation space of the initial algorithm. It is easy to show that the aspect ratio of the inserted rectangle R that maximizes Area(R) is equal to the ratio (*length of semi major axis / length of semi minor axis*) of the parent ellipse E*. Our subsequent developments are based on this configuration.

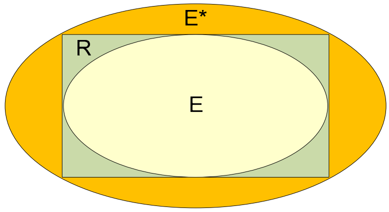


Fig. 1. Computation of display space occupation in initial ellimap algorithm

We need to compute the ratio ρ = Area(E) / Area(E*) which can be expressed as $\rho = \rho_1 . \rho_2$ with ρ_1 = Area(E) / Area(R) and ρ_2 = Area(R) / Area(E*).

This ratio ρ does not depend on the number of ellipses E_i (i.e. children nodes) that are included into E* (i.e. parent node). Indeed, in case of several children the rectangle R is divided into smaller rectangles R_i having a cumulated area equal to Area(R). Each ellipse E_i is inserted into the corresponding R_i. For every ellipse E_i the ratio Area(E_i) / Area(R_i) is identical and consequently is also equal to ρ_1. It is trivial to show that $\rho_1 = \pi/4 \approx 0.7854$. We can also easily compute that $\rho_2 = 2/\pi$. Finally we obtain $\rho = 0.5$. In other words, the initial ellimap algorithm uses only 50% of the available display space of the parent node (cf. ellipse E*) to display its children (cf. ellipses E_i). This obviously appears to be insufficient and it must be increased.

3.2 New Algorithm

Our new algorithm is called Combined Geometrical Distortions (CGD) because it successively applies translations and distortions of the nested ellipses in order to increase their areas while conserving their surface ratio at each hierarchical level.

The process is initialized (cf. **Fig. 2**) with the former ellimap algorithm with size-depending division (i.e. a rectangle R is inserted into the parent ellipse E*, then R is divided along its largest dimension in *n* smaller rectangles R_i in which the ellipses E_i corresponding to the children nodes are inserted).

In the first step (cf. **Fig. 3**) of CGD we identify the ellipse E_1 of which the center is the closest to the center of E* along the division axis (in the example: x axis). If two

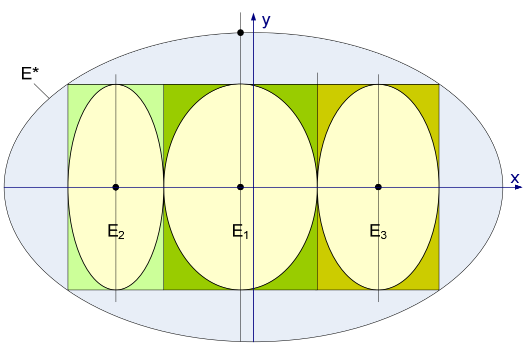


Fig. 2. Initialization step by size-depending ellimap algorithm

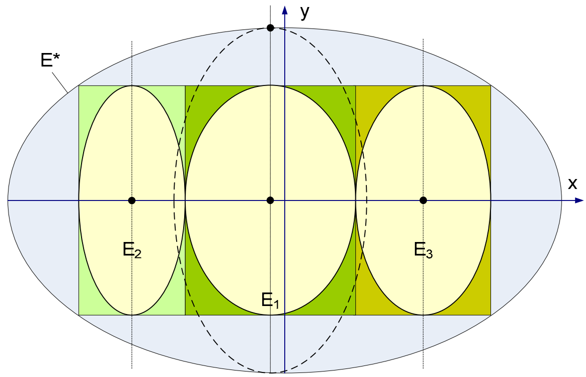


Fig. 3. CGD algorithm: step 1

ellipses E_i are equally distant from this point we randomly choose one of them. Then we stretch E_1 along the other axis (i.e y axis) to make it tangent to E*. The distortion factor is called D. Next the ellipse E_1 is stretched with a factor k/D along the x axis ($\sqrt{2}$ seems to be a good value for k and it is used in the evaluation, cf. section 4).

In the second step the ellipses E_i ($i \neq 1$) are translated along the x axis in order to be tangent again (cf. **Fig. 4**).

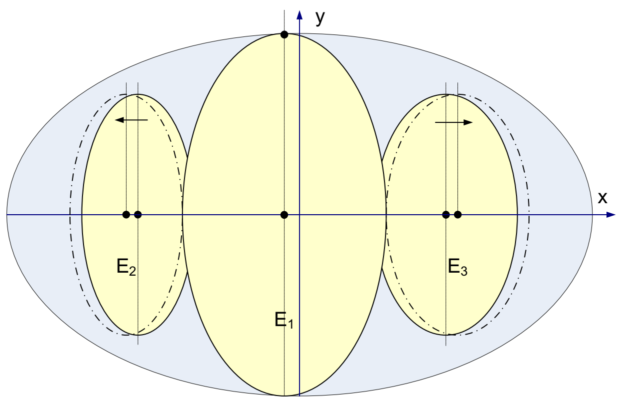


Fig. 4. CGD algorithm: step 2

In the third step the ellipses E_i ($i \neq 1$) are distorted similarly as E_1 has been in the first step: stretched along the y axis to be tangent to E* and stretched along the x axis. Note that these steps increase the area of the ellipses E_i by a factor of k. However, their ratio is kept constant.

Then (step 4) the ellipses E_i ($i \neq 1$) that have just been distorted are translated along the x axis to be tangent again.

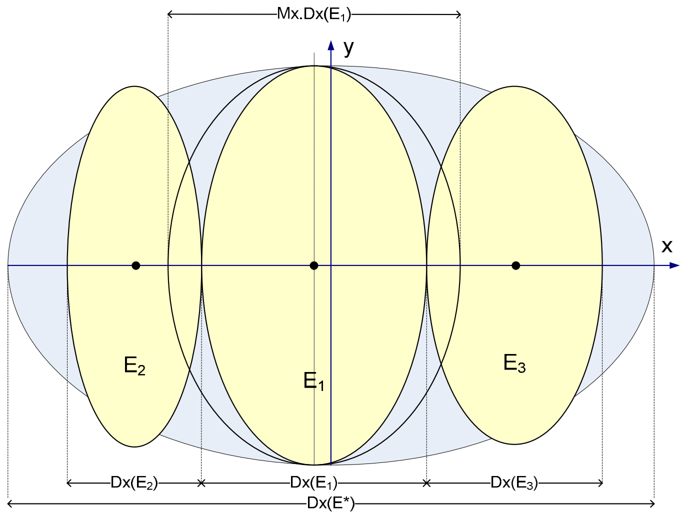


Fig. 5. CGD algorithm: step 5

We stretch then (step 5) the E_1 ellipse along the x axis with a magnifying factor Mx that is computed in order to use the remaining space along the x axis:

$Mx = Dx(E^*) / [Dx(E_1) + Dx(E_2) + Dx(E_3)]$

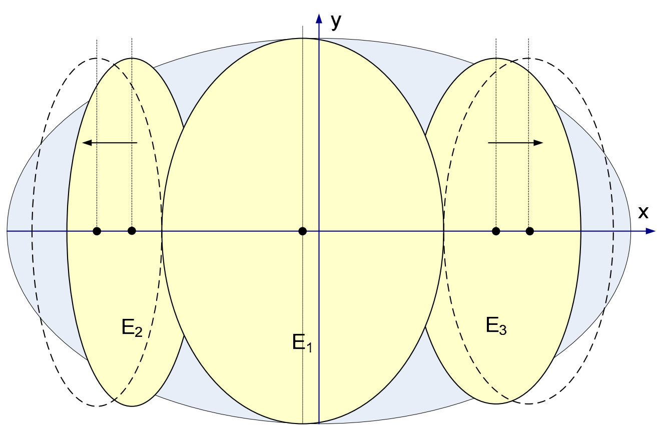


Fig. 6. CGD algorithm: step 6

In the sixth step the ellipses E_i ($i \neq 1$) are translated to keep the tangency property. Then (step 7) they are stretched along the x axis with the magnifying factor Mx.

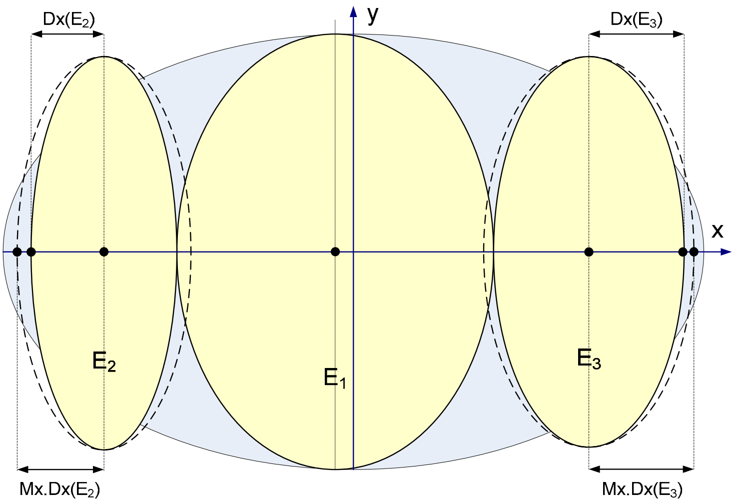


Fig. 7. CGD algorithm: step 7

The ellipse E_i ($i \neq 1$) are then translated again along the x axis to keep the tangency property (step 8, cf. **Fig. 8**).

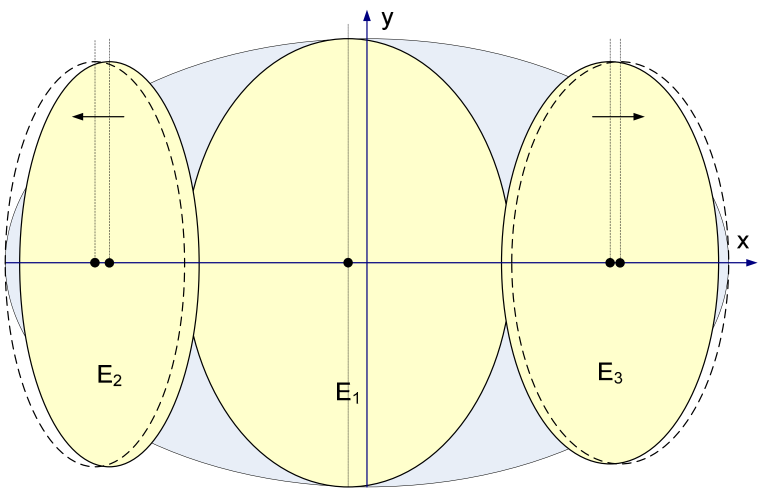


Fig. 8. CGD algorithm: step 8

In the ninth step, an empirical approach is used to decrease the size of the ellipses E_i (cf. **Fig. 9**). This iterative strategy (steps 9 to 11) was found to be the most effective to reach a well balanced result between display space occupation and computation time. For each ellipse E_i we compute the distance $Sy(E_i)$ defined as the difference between $Dy(E_i)$ (i.e. semi-axis of E_i along the y axis) and the y coordinate of the intersection point I_i between the ellipse E* and the vertical symmetry axis of E_i.

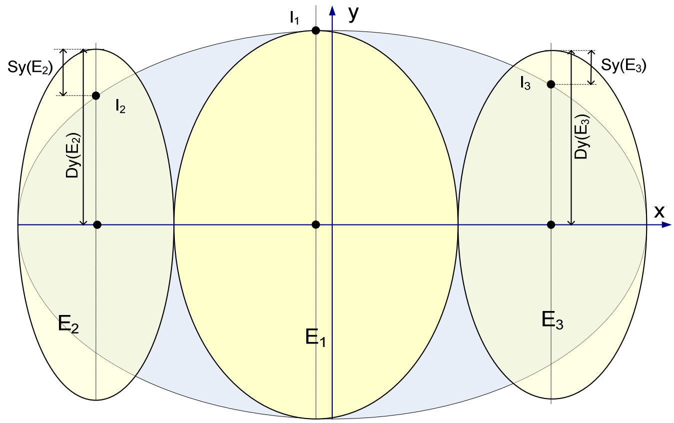


Fig. 9. CGD algorithm: step 9

Next we compute $Syr(E_i) = Dy(E_i) / [Dy(E_i) – Sy(E_i)]$ and

Syr-max = max { $Syr(E_i)$} with i = 1 to n.

If Syr-max > 1, we must go to the tenth step to decrease the size of the ellipses E_i .

If Syr-max ≤ 1, we go to the eleventh step.

Note that (Syr-max < 1) does not mean that every ellipse E_i is fully located inside the ellipse E*. However, it is almost always true and if it is false our algorithm is still valid because the eleventh step will correct any inaccuracy from this view point.

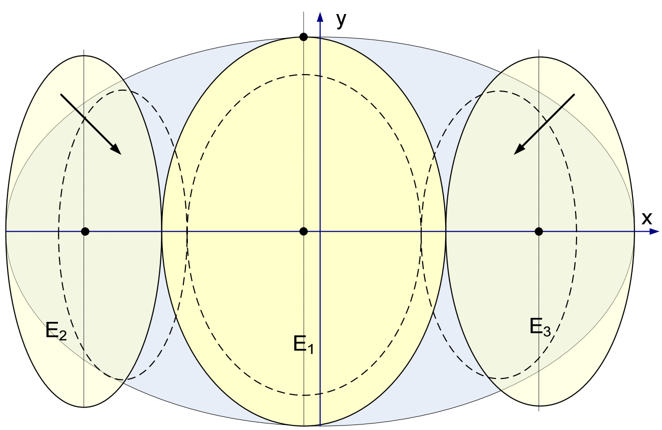


Fig. 10. CGD algorithm: step 10

The tenth step aims to decrease the size of every ellipse E_i with an identical factor R along the x and y axis. Having explored several strategies to compute R, it has appeared that a good option was to choose a constant function (e.g. R = 0.95). After the reduction step, some translations along the x axis are applied to the ellipses E_i in order to keep them tangent. Then we go back to step 9 to evaluate whether all the ellipses E_i are (approximately) located inside E*.

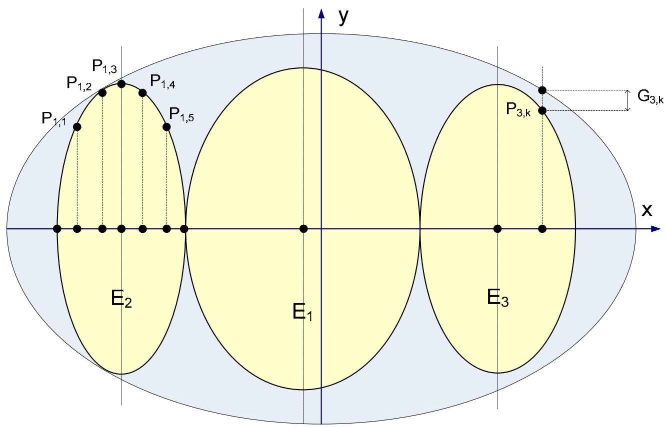


Fig. 11. CGD algorithm: step 11

Our purpose at the end of the tenth step is to have the best layout possible because it will be used in the eleventh step as the initial point of a process that is more intensive in computation than the previously described operations.

The eleventh step checks whether the ellipses E_i are located inside E* by computing the coordinates of *m* points $P_{i,j}$ of their perimeter. If at least one of the points $P_{i,j}$ for one of the ellipses E_i is located outside the ellipse E*, we apply again the reduction factor R to every ellipse E_i. When all points $P_{i,j}$ are inside the ellipse E*, we iteratively increase the size of all ellipses E_i with a magnifying factor M_k (e.g. 1.1) until one of the points $P_{i,j}$ is either tangent to the perimeter of E* (tangency for $P_{i,k}$ is defined by the expression $G_{i,k} < \varepsilon$, see **Fig. 11**) (case A) or outside E (case B). This moment of the iterative process is called the critical step. In the case A, the iterative process is stopped and the ellipses E_i are the final ones. In the case B, a new factor M_{k+1} (with $M_{k+1}<M_k$) is chosen (e.g. 1.01). The iterative process is then resumed at the step just before the critical one and the ellipses are iteratively magnified with the new ratio M_{k+1}.

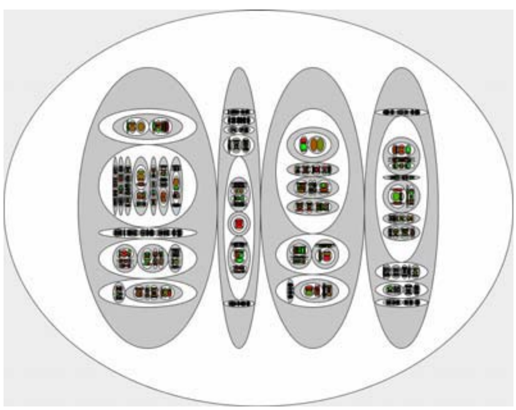


Fig. 12. Initial size-depending algorithm for a simulated hierarchy with 4477 nodes

The final result is illustrated below. The CGD algorithm (cf. **Fig. 13**) can advantageously be compared to the initial size-depending one (cf. **Fig. 12**).

We can also mention at this point that our Java implementation of CGD running on a classic PC (Windows XP, Intel Core 2 CPU at 2.40 GHz, 1 GB RAM) produces graphics allowing real time interaction up to several thousands of nodes (tested with datasets up to 10,000 nodes).

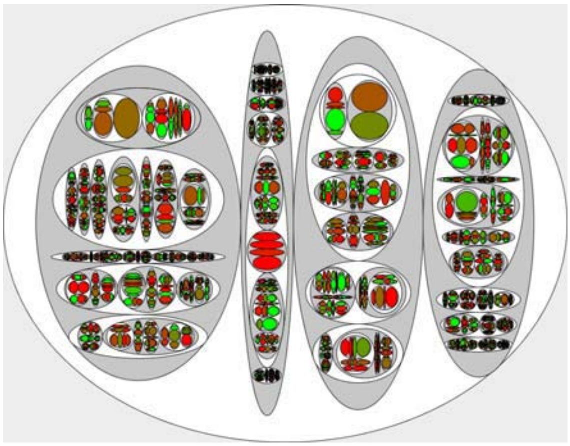


Fig. 13. CGD algorithm with the same hierarchy as Fig. 12

It is important to point out at this stage that any percent of gain in display space occupation can have a significant influence on the final result because it applies to every level of the hierarchy.

In fact, the occupation of the display space decreases at an exponential rate with the level of depth in the hierarchy.

With the hypothesis of a mean occupation rate γ at each level and considering that the ratio between the area of the root ellipse and the rectangular display space is known (= $\pi/4$), we can approximate the cumulated area of all ellipses E_i of level n with (Eq1):

$$\text{occupied space(level n)} = \gamma^n \,.\, (0.25 \,.\, \pi \,.\, \text{display space}) \text{ with } \gamma \in]0, 1] \qquad \text{(Eq.1)}$$

4 Evaluation

4.1 Mean Rate of Display Space Occupation

In order to evaluate the gain in space usage of the CGD algorithm we have identified two parameters having a potential influence on the occupation of the display space: the mean number of children by node and the ratio of the node weights. This is due to the fact that for each node having some children a new instance of CGD is executed to define the most appropriate ellipses to be drawn. Therefore neither the total number of nodes in the hierarchy nor the number of depth levels influence the space occupation for each instance of CGD. Nevertheless, it is obvious that those parameters influence the global picture. In this first sub-section we evaluate the γ factor of the equation (Eq.1) and in the second one we study its influence on the global rate of display space occupation.

Our evaluation is based on a Monte Carlo approach. A large number of hierarchical datasets H_k were generated. They are hypothesized to be a significant sample of the datasets encountered in reality. Considering the preceding discussion, we only need to generate datasets having one level of depth with varying values for the number of children nodes of the root and for the distribution of the weights of those children nodes.

For each hierarchy H_k we compute for the root N_k the value of ODS_k (Occupation of the Display Space) as the ratio of the cumulated area of the *n* ellipses representing its children $C_{i,k}$ and the ellipse representing N_k : $ODS_k = [\Sigma_{i=1..n} \text{Area}(C_{i,k})] / \text{Area}(N_k)$.

Next we compute the mean value μ_{ODS} and the standard deviation σ_{ODS} of the ODS parameter for the complete sample of hierarchies $\{H_k\}$.

Our random datasets generation module permits to set the value of the mean and standard deviation of the number of children by node (according to a lognormal law because negative values does not make sense) as well as the mean and the standard deviation of the node weight (according to a lognormal law for the same reason as above). The complete set of hierarchies $\{H_k\}$ is composed by aggregating some subsets $\{H_{k,s}\}$. Each subset $\{H_{k,s}\}$ is randomly generated by setting a value for the four parameters mentioned just above. In our evaluation the values set are given in Table 1. A random set $\{H_{k,s}\}$ was generated for each combination of the boxes, which means that 25 subsets $\{H_{k,s}\}$ were produced. Each subset $\{H_{k,s}\}$ contains 100 hierarchies.

At this stage it is important to mention that the random draw can generate some hierarchies where the root only has one child. In this case, the CGD algorithm makes the single child ellipse identical to the parent one. This behavior can be considered as a favorable bias that artificially increases the value of μ_{ODS}. Therefore, to strengthen the validity of our results, we have excluded the hierarchies corresponding to this case from the complete sample set $\{H_k\}$. We have also removed from $\{H_k\}$ the hierarchies H_k where the root has no children nodes. Finally the resulting set $\{H_k\}_{\text{filtered}}$ contains 2015 hierarchies. The global results of this simulation are displayed in Table 2. Note that the hierarchies H_k were not sorted (e.g. according to the node weights).

First we mention the distribution parameters of two main characteristics of $\{H_k\}_{\text{filtered}}$: the number of children of the root (NCh) and the ratio between the largest and smallest weight of these children (WR). For a node N having *p* children N_m, WR is defined by the expression: $WR = \max_{m=1..p}(\text{weight}(N_m)) / \min_{m=1..p}(\text{weight}(N_m))$.

Table 1. Parameters settings for the random generation of hierarchies

Number of children by node	$\mu_{set} = 2$ $\sigma_{set} = 2$	$\mu_{set} = 4$ $\sigma_{set} = 4$	$\mu_{set} = 6$ $\sigma_{set} = 6$	$\mu_{set} = 8$ $\sigma_{set} = 8$	$\mu_{set} = 10$ $\sigma_{set} = 10$

Children weight	$\mu_{set} = 100$ $\sigma_{set} = 50$	$\mu_{set} = 100$ $\sigma_{set} = 75$	$\mu_{set} = 100$ $\sigma_{set} = 100$	$\mu_{set} = 100$ $\sigma_{set} = 125$	$\mu_{set} = 100$ $\sigma_{set} = 150$

Table 2. Global results

	Min	Max	Mean	Std. Dev.
NCh	2	119	7.06	6.78
WR	1.01	619.08	13.69	30
ODS - CGD	0.578	0.757	0.682	0.038
ODS – initial algorithm	0.5	0.5	0.5	0

We can observe that the mean rate of display space occupation (μ_{ODS} = 0.682) is above the one of the initial ellimap algorithm (0.50). A t-test and a non-parametric Wilcoxon rank test showed that the observed difference between the mean ODS in both cases is highly significant ($p < 0.001$). Those figures confirm the impression given by comparing Fig. 12 and Fig. 13. Therefore it is clear that the CGD represents a step forward compared to the initial proposition. In our simulation the worst ODS for CGD (0.578) is still better than the ODS for the initial ellimap algorithm.

The low standard deviation of ODS is another notable result because it shows that the performance of CGD is stable for a relatively large range of configurations of the hierarchies to be displayed.

Finally, with the hypothesis of a mean occupation rate (γ_{CGD} = 0.682) at each level, the cumulated area of all ellipses E_i of level *n* can be approximated by (Eq. 2) for the CGD algorithm.

$$\text{occupied space(level n)} = (0.682^{n} \,.\, 0.25 \,.\, \pi \,.\, \text{display space}) \qquad \text{(Eq.2)}$$

The Table 3 concretely illustrates the meaning of (Eq. 2) with numbers. At the second level of depth, the CGD algorithm uses almost the double of the display space than the initial algorithm to draw the corresponding ellipses. At the fifth level, this ratio is nearly five and at the tenth level it reaches more than twenty.

While the absolute values of ODS may appear low for high levels of depth, it must be reminded that the remaining part of the display space is used to show the ellipses of the upper levels and therefore the hierarchical structure itself.

Table 3. Comparison of display space occupation for CGD and initial algorithms

	Level (*n*)	ODS		Ratio
		Initial algo.	CGD algo.	
Display space		100	100	1
Root	0	78.54	78.54	1
Level 1	1	39.27	53.56	1.36
Level 2	2	19.63	36.53	1.86
Level 3	3	9.82	24.91	2.54
Level 4	4	4.91	16.99	3.46
Level 5	5	2.45	11.59	4.72
…	…	…	…	…
Level 10	10	0.08	1.71	22.29

Table 4. Influence of the statistical nature of γ_{CGD}

	Level (n)	ODS		Real Example
		Very lucky draw $\gamma = 0.75$	Very unlucky draw $\gamma = 0.58$	
Display space		100	100	100
Root	0	78.54	78.54	78.54
Level 1	1	58.90	45.55	53.56
Level 2	2	44.18	26.42	39.29
…	…	…	…	…
Level 5	5	18.64	5.16	13.29

4.2 Exploration of the Global Range of Display Space Occupation

While the γ factor is important to evaluate the performance of the algorithms, it may be useful to draw the attention on its statistical nature. Because it has been computed as a mean of observed values, the real value for a given instance of ellimap will most of the time be different. This phenomenon is illustrated in Table 4. In the most favorable cases ($\gamma = 0.75$ at every level), the ODS value is significantly higher and in the worst cases ($\gamma = 0.58$ at every level) it is much lower. Nevertheless, those extreme values have a very little probability to be observed. Therefore we have added in Table 4 the results of a real random draw (number of depth levels: 6; number of nodes: 15176; mean WR: 13.04; mean ODS = 0.70).

5 Conclusion

This paper proposes a new algorithm called CGD to increase the low occupation of the display space in the ellimap visualization technique, which has been identified as its major weakness. The evaluation of CGD shows that the mean occupation of the display space is significantly higher than what was observed with the initial proposition. It may be noted that, to the limit of our knowledge, this evaluation is the first one to quantitatively study the space occupation of a visualization technique of hierarchies. This new approach reaches an interesting balance between the representation of the data set structure, the weights of the nodes and the occupation of the display space. In the future, a user-based evaluation similar to Otjacques et al.'s one [8] is envisaged to assess the influence of this geometrical improvement in terms of data cognition.

References

1. Balzer, M., Deussen, O., Lewerentz, C.: Voronoi Treemaps for the Visualization of Software Metrics. In: ACM Symposium on Software Visualization 2005, pp. 165–172. ACM Press, New York (2005)
2. Card, S., Mackinlay, J., Shneiderman, B.: Readings in Information Visualization, Using Vision to Think. Morgan Kaufmann, San Francisco (1999)

3. Chi, E.H., Pitkow, J., Mackinlay, J., Pirolli, P., Gosswiler, R., Card, S.K.: Visualizing the Evolution of Web Ecologies. In: ACM Conference on Human Factors in Computing Systems 1998 (CHI 1998), pp. 400–407. ACM Press, New York (2005) (1998)
4. Lamping, L., Rao, R.: Laying out and Visualizing Large Trees Using a Hyperbolic Space. In: ACM Symposium on User Interface Software and Technology (UIST 1994), pp. 13–14. ACM Press, New York (1994)
5. Lee, B., Parr, C., Plaisant, C., Bederson, B., Veksler, V., Gray, W., Kotfila, C.: TreePlus: Interactive Exploration of Networks with Enhanced Tree Layouts. IEEE Transactions on Visualization and Computer Graphics 12(6), 1414–1426 (2006)
6. Lin, C., Yen, H.: On Balloon Drawing of Rooted Trees. In: Healy, P., Nikolov, N.S. (eds.) GD 2005. LNCS, vol. 3843, pp. 285–296. Springer, Heidelberg (2006)
7. Nguyen, Q., Huang, M.: A Space-Optimized Tree Visualization. In: IEEE International Symposium on Information Visualization, pp. 85–92. IEEE Press, New York (2002)
8. Otjacques, B., Collin, P., Feltz, F., Noirhomme, M.: Ellimaps: une technique basée sur la loi d'inclusion pour représenter des hiérarchies avec nœuds pondérés. Revue d'Intelligence Artificielle 22(3-4), 301–327 (2008)
9. Otjacques, B., Noirhomme, M., Gobert, X., Collin, P., Feltz, F.: Visualizing the activity of a web-based collaborative platform. In: International Conference on Information Visualization (IV 2007), pp. 251–256. IEEE Computer Society, Los Alamitos (2007)
10. Palmer, S.E.: Common Region: A new principle of perceptual grouping. Cognitive Psychology 24, 436–447 (1992)
11. Palmer, S.E., Rock, I.: Rethinking perceptual organization: The role of uniform connectedness. Psychonomic Bulletin and Review 1(1), 29–55 (1994)
12. Schultz, H.-J., Hadlak, S., Schumann, H.: Point-Based Tree Representation: A new Approach for Large Hierarchies. In: IEEE Pacific Visualization Symposium (PacificVis 2009). IEEE Press, NewYork (2009)
13. Shneiderman, B.: Tree Visualization with Tree-Maps: 2-d Space-Filling Approach. ACM Transactions on Graphics 11(1), 92–99 (1992)
14. Shneiderman, B.: Treemaps for space-constrained visualization of hierarchies, University of Maryland, Human – Computer Interaction Lab. (2009), `http://www.cs.umd.edu/hcil/treemap-history/` (accessed April 7, 2009)
15. van Wijk, J.J., van de Wetering, H.: Cushion TreeMaps. In: IEEE International Symposium on Information Visualization, pp. 73–78. IEEE Press, NewYork (1999)
16. Vliegen, R., van Wijk, J.J., van der Linden, E.: Visualizing Business Data with Generalized Treemaps. IEEE Transactions on Visualization and Computer Graphics 12(5), 789–796 (2006)
17. Ware, C.: Information Visualization, Perception for Design, 2nd edn. Morgan Kaufmann, San Francisco (2004)
18. Wood, J.: Spatially Ordered Treemaps. IEEE Transactions on Visualization and Computer Graphics 14(6), 1348–1355 (2008)

Visual Search Strategies of Tag Clouds - Results from an Eyetracking Study

Johann Schrammel[1], Stephanie Deutsch[1], and Manfred Tscheligi[1,2]

[1] CURE - Center for Usability Research and Engineering
[2] University of Salzburg, ICT&S Center
{schrammel,deutsch,tscheligi}@cure.at

Abstract. Tag clouds have become a frequently used interaction technique in the web in the past couple of years. Research has shown the influence of variables such as tag size and location on the perception of tag clouds. However, several questions remain unclear. First, little is know on how tag clouds are perceived visually and which search strategies users apply when looking for tags in a tag cloud. Second, there are variables, especially tag location, were prior work comes to conflicting results. Third, several approaches to present tag clouds with the tags semantically clustered have been proposed recently. However, it remains unclear which effects these new approaches have on the perception of tag clouds. In this paper we report the results of an extensive study on the perception of tag clouds using eye tracking technology that allows answering these questions.

Keywords: Tag clouds, eye tracking, folksonomy, clustering, semantic grouping, visualization.

1 Introduction

A tag cloud is a visual depiction of words, called tags, which is typically used to describe the contents of a web service. It consists of a visually weighted list of tags arranged in a certain layout. Tags are usually hyperlinks in form of single words that lead to a collection of more elaborated issues. The weight corresponding to their relevance or popularity within a context is characterized by one or more visual features for each word, such as font size, color, intensity, position or a combination of such features. The most used layout strategy is alphabetical sorting of tags, but also other principles such as importance-based or spatially optimized have been proposed and used. Recently semantic layouts, i.e. the clustering and display of tags according to their semantic relatedness, have attracted a lot of attention. Tag clouds have been used for various tasks such as searching, browsing, recognizing and impression formation about a subject [1].

Besides the evaluation of the importance of different visual features only little empirical evidence is available on the effectiveness of tag clouds. Several user-based studies on different aspects of tag clouds such as the interaction between different properties in tag cloud design are available that use traditional measurement methods

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 819–831, 2009.

of human performance (e.g. reaction time, accuracy). However, we are not aware of any investigations focusing on the visual perception of tag clouds that involved measurements with physiological methods, such as the recording of eye movements during stimulus presentation. We think eye tracking applications are particularly interesting as they allow interdisciplinary insights into the visual, cognitive, and attentional aspects of human information processing, for an example of use see [2]. Within HCI and Usability research eye movement recordings can provide an objective source of interface evaluation data.

2 Related Work

2.1 Tag Clouds

The influence and importance of visual features of tags in tag clouds has been the subject of several studies [1,3,4,5]. These studies provide concise evidence that the visual features font size, font weight and intensity seem to be the most important variables in affecting the perception of tag clouds. However, the exact role of tag position is not clear yet. *Bateman et al.* [4] reported no influence of the tag position on performance whereas *Rivadeneira et al.* [1] report that tags in the upper left quadrant of the tag cloud were recalled more frequently by their participants than tags placed in the lower-right quadrant. A similar focus on the upper-left quadrant was found by *Halvey and Keane* [3] for specific search tasks.

Words in tag clouds are typically listed in alphabetical order from upper-left to lower-right. However, recently several proposals for semantic arrangment of tags in a cloud were made. A layout algorithm based on tag similarity was proposed by *Hassan-Montero and Herrero-Solana* [6]. *Fujimura et al.* [7] use the cosine similarity of tag feature vectors (terms and their weight generated from a set of tagged documents) to measure tag similarity and then calculate a tag layout where the semantic relatedness is represented as the distance between tags. *Berlocher et al.* [8] propose a very similar approach.

Whereas the creators of these proposals clearly expect advantages for the interaction it remains questionable whether such improvements actually can be achieved. *Schrammel et al.* [5] could not find any substantial gains in performance for search tasks as well as for tag rememberability when they compared semantic layouts to tag clouds ordered alphabetically or randomly. Only for specific searches semantic layouts could provide an improvement over random organization, but as alphabetical layouts are still better than semantics for this type of task the actual usefulness of this approach remains questionable.

The usage of tag clouds have to consider different tasks and task contexts. For the application of tag clouds in search contexts *Sinclair and Cardew-Hall* [9] showed that people prefer traditional search interfaces over tag clouds to look for specific information. Tag cloud interfaces were preferred for the retrieving of more general information contents, so after all they remain to serve users as supplementary tools in search tasks, not as sole alternative to traditional search interfaces. Similarly the participants in the study of *Kuo et al.* [10] were able to resolve descriptive retrieval tasks better and faster from tag cloud interfaces than from queries of hierarchical ranked data, in this regard they rated tag clouds as more helpful. Concerning the identification of relationships between multiple concepts users gave better answers when using a search engine than with tag clouds. Overall users spent more time and ranked tag cloud interfaces less helpful as they did for a traditional search engine.

These results point out certain inappropriateness of tags clouds for use in specific search tasks which for our experiment led us to include two different types of tasks: *specific search*, i.e. finding a given tag in a tag cloud, and *general search*, i.e. finding a tag within a tag cloud that belongs to a specific topic.

2.2 Basic Characteristics of Gaze Behavior

Within eye tracking research, variables that appear to be valid indicators of visual perceptive processes are fixations, saccades, and scanpaths [11]. Eye fixation is defined as the spatially stable gaze on a specific area of the visual field for approx. 200-300 milliseconds. Although their interpretation depends on the context, fixation represents the critical moment between states of attention and higher cognitive processes e.g. encoding information. Relevant for analysis should be the number of fixation overall, target fixation, fixation duration as indicators of complexity of the visual task, and time interval from stimulus onset to first fixation on the display [12]. Between two fixations a saccade represents the eye movement to the next viewing position and typically lasts for 20-35 milliseconds. During saccades there is no encoding activity, so no indication about visual processing can be earned, only number of saccades and regressive saccades (backtracking eye-movements) can indicate difficulties during encoding processing.

We need to make saccades because of the limitation of the eyes acuity. The visual field is typically divided into three regions: foveal, parafoveal, and peripheral. The acuity is very good in the fovea (the central 2° of vision), is not nearly so good in the parafovea (out to 5° on either side of fixation) and is even poorer in the periphery (the region beyond the parafovea). In our experimental setup (see section 'Apparatus' below for details) 2° is equivalent to about 35 pixel on the screen or 4 letters of a tag in the smallest font size.

Although humans can easily decouple the locus of attention and eye gaze location in simple discrimination tasks [13], in complex information processing tasks such as reading, the link between the two is probably quite tight. For analyzing the perception of tag clouds we therefore can securely work with the hypothesis that the spot of fixation is identical with the spot of attention.

2.3 Gaze Behavior in Related Tasks

Basic Visual Search

Eye movements are influenced by textual and typographical variables e.g. the manipulation of visual features. According to *Duchowski* [14] the classical search theory of Treisman and Gelade from 1980 contains a simple model of the visual system with practical relevance for the examination of eye movements. In this model a pre-attentive subsystem analyses simple visual features over the entire visual field in parallel. Classical simple visual features are brightness, orientation, size (or spatial frequency), color, movement, probably curvature, and maybe more.

Reading
According to Keith Rayner [11] when reading English eye fixations last about 200-250 ms and the mean saccade size corresponds to 7-9 letter spaces. Besides there is evidence that eye movements are influenced by textual and typographical variables, as important to note, fixation duration increases and saccade length decreases when a text becomes more difficult. Also factors such as the quality of print, line length, and letter spacing influence patterns of eye movements, perception respectively. Perceptual span in reading: for readers of alphabetical orthographies (e.g., English, French, and Dutch) the span extends from the beginning of the currently fixated word but no more than 3-4 letters to the left of fixation to about 14-15 letter spaces to the right of fixation.

Looking at graphs
Other studies focus on eye movement analysis when people view statistical graphs, maps, and images [11]. In order to detect the underlying cognitive processes of graph comprehension *Carpenter & Shah* [15] recorded people's eye movements while they examined graphs showing complex interactions and claim for the occurrence of 'chunking' processes during encoding of graphical information.

In the web
Particularly interesting findings offers the usage of eye tracking methodologies within Web usability studies. *Klöckner et al.* [16] observed most of their participants performing a linear strategy while browsing search results in order to evaluate each result one by one before selecting a document to open. Further work on eye movement recordings and Web search is presented by *Cutrell & Guan* [17] and *Pan et al.* [18]. Lastly there should be noted the relevance of using eye tracking methodologies when investigating the impact of advertisements on people in consumer research [19]. Facing the increase of online information platforms and online newspapers new paradigms have been generated where people deal with more complex scenes, e.g. when a web page contains textual and graphical content. For the analysis of eye movements from Web viewers that did not alternate between the text and picture part of an advertisement see *Rayner et al.* [20].

3 Research Questions

The main objective of the current study is to better understand the visual search behavior of users when interacting with tag clouds. We want to find out which search patterns users apply, if and when they change their search strategies and whether they adapt or change these strategies in the different layout conditions. We also want to see whether findings from previous tag cloud studies based on measures of task completion times or item rememberability can be reproduced and confirmed using eye tracking techniques or if new interpretations are needed. We especially want to contribute to clarify the up till now controversial aspect of tag position. The analysis of eye movement metrics could clarify if and when people become aware of how tags are organized in the visualization, concretely if and when people recognize a specific order in a given tag arrangement during a search task, and whether or not they adopt the their search strategies.

4 Method

Apparatus

Eye movements of all participants were recorded with Tobii x-50 Eye Tracker (Tobii Technology AB, Danderyd, Sweden), a standalone eye tracking unit with stimulus setup on a display in front of the subject. The Tobii Eye Tracker uses near infrared diodes to generate reflection patterns on the corneas of the eye of the user. These reflection patterns are collected by a high-resolution camera with a large fieldof- view to finally calculate the gaze point on the screen. Recordings of time stamps and gaze positions relative to stimuli for each eye were collected with ClearView 2.7.1 analysis software. The fixation filter was set to a fixation radius of 30 pixels and minimum fixation duration of 100 milliseconds.

Subjects sat in front of the stimulus monitor at a distance of approximately 60 cm from the eye tracker. Stimuli were displayed on a 17 inch LCD screen with a resolution of 1024 by 768 pixels. Before task initiation eye movements of each subject were calibrated appropriately, respecting calibration settings of 5 points at normal type, medium speed, and medium dot size.

Data manipulation and quality assurance

Eye tracking data are prone to systematic errors, especially for vertical position as small changes in the distance of the participant to the stimulus can have noticeable effects. Therefore we applied a manual control and correction of data. After data acquisition fixation filtered data were exported from ClearView to the statistical software (SPSS v.16.0) in form of a sequential list of fixations and eye movement recordings. Using a gaze pattern tool of ClearView 2.7.1 recordings of each participant were explicitly examined regarding systematic deviations and completeness of tracked data. The task screens were test participants had to click on a 'Next'-button served as reverence points to estimate the quality of the data. In case of systematic deviations of the gaze points were corrected by adjusting the y-coordinates in SPSS.

The exact position on the screen in pixels of all words in the tag clouds within all conditions was measured and these coordinates (X and Y) where used to link the fixation data to an actual word (and its size).

We are very confident that our data is very accurate. A good indicator for this accuracy is that in most cases the target word is listed in as one of the last words fixated before the task was finished.

Sample

15 subjects (7 male, 8 female) were recruited from our internal participant database according to the following criteria: all participants used the internet frequently and had good English skills. Their sights were normal or corrected to normal.

Stimuli

The basic test content for the experiment, adopted from previous work [5], was formed by 304 popular tags taken from different thematic clusters of flickr. Tags were randomly assigned to content groups which formed the basic elements for the tag clouds. Each tag cloud consisted of 76 items arranged in 7 lines with 11 respectively 10 tags per line. Every tag cloud included 6 very big, 11 big, 22 small, and 37 very small tags. Tags were randomly assigned to one of these font size groups.

Three different layouts were used in the experiment:

- *Alphabetic:* Here tags were placed alphabetically starting at the top-left similar to reading text.
- *Random*: Tags were placed randomly on the tag by use of a random number generator.
- *Semantic:* We used the getrelated-function of flickrs API [21] to retrieve a list of the tags most related to each word within the tag cloud. Then based on the number of co-occurring related tags a measure for the relatedness of two tags was calculated. An alternating least-squares algorithm to perform multidimensional scaling (ALSCAL) [22] was used to compute a two dimensional arrangement of the tags. In the third step we used the value on the y-axis to form 7 groups of 11 resp. 10 tags each. Next tags within each group were sorted according to their value on the x-axis. The result provided an 11 times 7 arrangement that was used to generate the tag cloud.

Figure 1 on the next page shows three example tag clouds derived from the same content group i.e. they consist of exactly the same tags with the same size, only the ordering is different.

Procedure

After being instructed adequately, subjects started with the tests. A browser window opened where the participants had to read a short instruction indicating the nature of the task on the upcoming screen. Readiness for stimulus onset was confirmed via mouse click on a ‘Next’-button. Once the target tag was found, participants had to click on it to proceed to the next item.

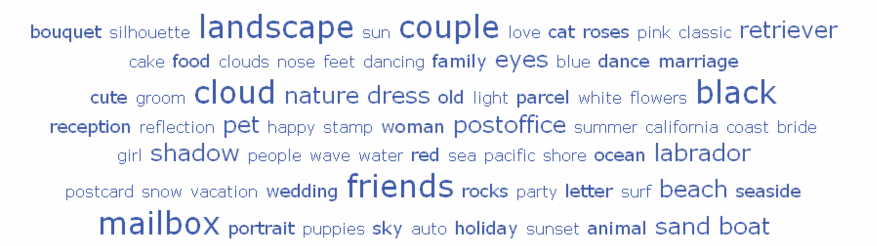
For each of the three layouts 12 search tasks with different targets within the same tag cloud were shown. Target tags were evenly distributed across font sizes and location (quadrants). Subjects had to solve two different types of tasks. In the first task condition participants were demanded to find a specific tag within a tag cloud as fast and accurately as possible. The second task was to find a tag that belongs to a specific topic. In all experiments test conditions where presented in different order using content sets and tag cloud layouts that counterbalanced position and learning effects.

Scanning detection

In order to extract possible scanning patterns from the recorded eye movements of the participants in our experiment an algorithm was implemented that differentiates between a person's serial scanning behavior and chaotic or random search patterns. Our algorithm is based on the reading detecting system of *Campbell and Maglio* [23]. The underlying algorithm is known as pooled evidence and allows reliable and accurate recognition of reading performance. As serial scanning - contrary to reading - also can occur from right to left the algorithm had to be adapted.

a) Alphabetic

b) Random



c) Semantic

Fig. 1. Example content displayed in the three different layouts

Initially the system is in 'chaotic-search' mode until an event assigned to a specific eye movement occurs. Via incrementing or de-incrementing a 'serial-scanning'- evidence variable when the eyes move to the right or to the left, evidence of serial scanning can be accumulated until a certain threshold value is obtained. Once the threshold is reached 'serial-scanning' is detected and the system switches from 'chaotic-search' into 'serial-scanning' mode. In 'serial-scanning' mode the system records every tag until a “scan jump” (long distance move above threshold value) is detected which sends the system back into 'chaotic-search' mode.

In detail we used the following threshold values (difference in pixel): ShortHorizontal: < 180, MediumHorizontal: 180-270, LongHorizontal: > 270, ShortVertical: < 50, MediumVertical: 50-120, LongVertical: > 120.

The algorithm identifies 'serial-scanning' when a short or medium horizontal change occurs three times successively without more than one short horizontal change. The algorithm was tested semi-formally with a small randomly selected subset of data from different test participants. It provided the same estimation regarding 'serial-scanning' or 'chaotic-search' as a human observer of the data identified in about 70 per cent of cases.

5 Results

In the following sections we analyze the observed and recorded gaze behavior of the users with regard to the perception of the tag clouds. We do not focus on the differences in overall search times and user satisfaction, readers interested in these topics are referred to *Schrammel et al.* [5], where these aspects are reported and discussed in detail.

5.1 Visual Search Patterns

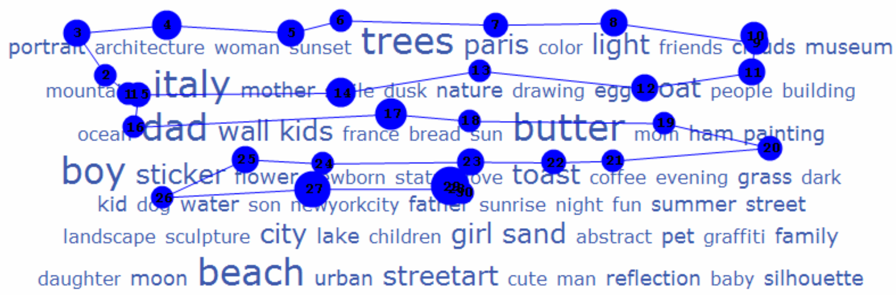
Recognizing Alphabetic Tag Organization
A first relevant question when observing and analyzing the users visual search pattern during interacting with tag clouds is how long it takes the different users to recognize and understand the organization principle of a tag cloud. It is easily possible to identify this point in the eye tracking data for an alphabetic layout and a specific search. Once the user recognizes this organization principle the gaze behavior is much more targeted and heads straightforward to the estimated target position. Unfortunately there is no similar possibility to identify semantic gaze patterns; therefore this section focuses on the alphabetic layout.

We manually analyzed the data provided by the eye tracking system and identified the point at which the users did noticeably change their search behavior. Additionally after the test session we asked the participants if they recognized a layout principle and approximately when. Data from the eye tracking analysis and the self-reported behavior were very stringent, but the data analysis naturally allowed a more finetuned identification of the point in time.

Overall we observed that all participants except one realized the alphabetic order of tags at some time during the experiments. 8 of 15 users recognized the order and adopted the pattern within the first six tags of the target tag search task, 5 of them even within the first three search tasks.

Chaotic and Serial Search Strategies
Another aim was to identify reading patterns or other search strategies during search task. We could identify two typical search patterns that were used by the participants: *chaotic search*, which showed no traceable strategy within the search process, and *serial scanning*, which was typically performed in a characteristic zigzag pattern. These strategies were applied independently from the layout condition i.e. they were equally used for random, semantic and alphabetic layouts. Figure 2 below shows prototypical examples of these two search patterns.

a) Chaotic Search Pattern



b) Serial Search Pattern (Zigzag)

Fig. 2. Typical Search Patterns

A typical behavior of users is to switch from chaotic to systematic search patterns when the chaotic approach does not turn out to provide any results. We could not identify a clear trend after how much time users change their strategies. There seem to be big differences between users, and also individuals seem to not always apply the same strategy. Please also note that users switched back and forth between the two strategies, and did not always consequently apply serial search until the target was found.

With regard to the scanning tag clouds are almost always scanned in a zigzagpattern, i.e. one line is scanned from left to right, the next from right to left and then starting over from left to right. As described in the methods section we modified a reading detection algorithm that allowed us to automatically detect scanning patterns. According to these algorithm users spend about 70 per cent of the time using chaotic search strategies and 30 percent applying scanning strategies. As mentioned above 10 users scan in both directions, but scanning in reading direction not surprisingly is applied slightly more frequently (57 versus 43 percent of scanning time). Statistical analysis with regard to differences caused by stimulus or task on the ratio of chaotic search versus serial scanning does not show any significant differences.

5.2 Tag Size and Attention

Another aspect we were interested in is how the font size of a tag within a tag cloud effects it's perception in the different layout conditions. Common sense and past research [1, 4] suggest that tags in bigger fonts are fixated both more often and for a

Table 1. Average time a single tag of the regarding size group is looked at

Size	**3 - very small**		**4 - small**		**5 - big**		**6 - very big**	
Task	Specific	General	Specific	General	Specific	General	Specific	General
Mean	48,0	35,0	60,2	42,4	85,4	64,5	128,1	113,1

longer time than smaller tags. However, it is not clear whether there is a difference of the relative importance of the tag size in different layout conditions as suspected by [5].

Analysis of Variance (ANOVA) with tag size, task type and layout condition as independent variables shows a significant influence of task ($F_{1,19}$=5,62, p=0,028) and size ($F_{3,61}$=69,06; p<0.000) on the cumulated fixation time per task. The chance for a tag in the biggest font to be fixated by the users gaze is about 2.5 times higher than for a tag in the smallest font. Table 1 below provides an overview on the average time a single tag of the regarding size group was looked at per search task in milliseconds. This trend is very consistent for both types of tasks. Times for the general search tasks are shorter than in specific searches because it is easier to find a tag that qualifies as target.

Schrammel et al. [5] hypothesized that semantic layouts would increase attention towards tags in smaller fonts as they provide valuable information about the semantics of an area of the tag cloud. However, the ANOVA does not show any significant interaction effects ($F_{6,186}$=0,32; p=0,92) for fixation time between tag size and layout condition.

5.3 Influence of Location and Time on Attention

In this section we analyze the importance of the tags location in the tag cloud and the time in the search process on the attention of the user. To do so tag clouds were divided into four quadrants for analyzing the data. Time in the search process was dichotomized into initial (first three seconds) and remaining search time. Three seconds for splitting search time was chosen, as an informal interview with 8 experts in human computer interaction showed that this was their average estimation of the typical duration of initial attention of a user in the context of using tag clouds.

Our data show a clear influence of a tags location on the amount of attention it receives from the user. The attention is strongest in the upper-left quadrant and decreases via upper-right and lower-left towards the lower right quadrant. Figure 3 below provides a graphical representation of this data, both for the initial phase of attention and for the whole search task.

ANOVA shows a highly significant effect of quadrant on the amount of visual attention it receives ($F_{3,58}$=7,64; p<0.000). There is no interaction between quadrant and time in the search process ($F_{3,70}$=0,32; p=0,811). This means the attentional preference towards the top and left is not related to effects of initial attention (e.g. users do start at the top left), but that this is a general phenomenon that occurs at all times during the search.

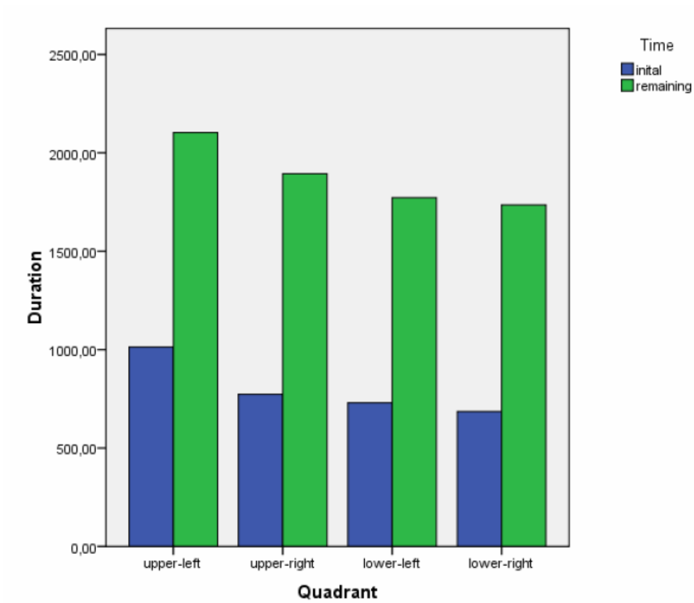


Fig. 3. Attention towards quadrant in initial and remaining time

5.4 Duration of Fixations

The duration of a fixation is an indicator for the complexity of a visual task [11]. Complex tasks typically go hand in hand with longer fixations. ANOVA on fixation times shows a significant influence of the task on average duration of fixations ($F_{1,14}$=5,08; p=0.041): Fixations typically were longer for general search tasks than for specific searches. This is easy to understand, because in the general search the user not only has to identify but also classify the focused tag to decide whether it is a correct choice or not. With regard to the layout conditions we hypothesized the fixation duration to be longest in the semantic condition as the understanding and processing of the semantic layout probably would add additional cognitive load. As expected this is the case for both tasks. However, this difference is not statistically significant ($F_{2,28}$=1,82; p=0,180).

6 Discussion

In this paper we described the results of our analysis of the visual perception of tag clouds. In accordance with prior research tag size was identified as the major influence on the perception. Through our research we were able to clarify conflicting issues with regard to the influence of a tags position on its perception: There is a clear and consistent influence of the quadrant on the amount of attention a tag can expect. The eye tracking analysis showed clearly that the attention decreases from top to bottom as well as left to right.

We did not find any relevant differences in attention between different tag layouts. This was very surprising to us, as for the case of semantic layouts we expected an increased focus on tags in smaller size, as they convey important information about the thematic area of a tag cloud. Also we expect to be able to identify differences in the search patterns related to the organization principle. From the qualitative interviews we know that not recognizing the thematic layout cannot be the sole source of this indifference. Also the reported significant differences in search times between semantic and random layouts suggest that there are different perceptual processes at work. However, these differences do not seem to be related to the overall amount of attention tags in different sizes receive or the application of recognizably different search strategies. We think the explanation might be that users learn the topical organization of a semantic tag cloud implicitly, and therefore no explicit change in the behavior can be observed. Another possible explanation is that users are able to identify the topic of an area without the need to focus on all the different words but by a blurry perception. However, provided that the visual capabilities of humans is only 2° or 35 pixel as described above we do not champion this explanation.

7 Conclusions

In this paper we could identify (and clarify) several important issues with regard to the visual perception of tag clouds. In future work we plan to further explore influences of different aspects such as number of tags in the tag cloud, familiarity with topic area and visual designation of semantic areas on the perception.

Aknowledgements. This work was supported by the Softnet Austria Competence Network. We also would like to thank Rainer Pfister for implementing the scanning detection algorithm.

References

1. Rivadeneira, A.W., Gruen, D.M., Muller, M.J., Millen, D.R.: Getting our head in the clouds: toward evaluation studies of tagclouds. In: Proc. CHI 2007, pp. 995–998. ACM Press, New York (2007)
2. Maughan, L., Gutnikov, S., Stevens, R.: Like more, look more. Look more, like more: The evidence from eye-tracking. Journal of Brand Management 14 (2007)
3. Halvey, M.J., Keane, M.T.: An assessment of tag presentation techniques. In: Proc. WWW 2007, pp. 1313–1314. ACM Press, New York (2007)
4. Bateman, S., Gutwin, C., Nacenta, M.: Seeing things in the clouds: the effect of visual features on tag cloud selections. In: Proc. Hypertext and Hypermedia 2008, pp. 193–202. ACM Press, New York (2008)
5. Schrammel, J., Leitner, M., Tscheligi, M.: Semantically Structured Tag Clouds: An Empirical Evaluation of Clustered Presentation Approaches. In: CHI 2009, Boston, MA, USA, April 3–9 (2009)
6. Hassan-Montero, Y., Herrero-Solana, V.: Improving tagclouds as visual information retrieval interfaces. In: Proc. InfoSciT 2006 (2006)

7. Fujimura, K., Fujimura, S., Matsubayashi, T., Yamada, T., Okuda, H.: Topigraphy: visualization for largescale tag clouds. In: Proc. WWW 2008, pp. 1087–1088. ACM Press, New York (2008)
8. Berlocher, I., Lee, K., Kim, K.: TopicRank: bringing insight to users. In: Proc. SIGIR 2008, pp. 703–704. ACM Press, New York (2008)
9. Sinclair, J., Cardew-Hall, M.: The Folksonomy Tag Cloud: When is it Useful? Journal of Information Sience 34(1), 15–29 (2008)
10. Kuo, B.Y., Hentrich, T., Good, B.M., Wilkinson, M.D.: Tag clouds for summarizing web search results. In: Proc. World Wide Web 2007, pp. 1203–1204. ACM Press, New York (2007)
11. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychological Bulletin 124, 372–422 (1998)
12. Goldberg, H.J., Kotval, X.P.: Computer interface evaluation using eye movements: Methods and constructs. International Journal of Industrial Ergonomics 24, 631–645 (1999)
13. Posner, M.I.: Orienting of attention. Quarterly Journal of Experimental Psychology 32, 3–25 (1980)
14. Duchowski, A.T.: A Breadth-First Survey of Eye Tracking Applications. Behavior Research Methods, Instruments, and Computers 34(4), 455–470 (2002)
15. Carpenter, P.A., Shah, P.: A model of perceptual and conceptual processes in graph comprehension. Journal of Experimental Psychology: Applied 4, 75–100 (1998)
16. Klöckner, K., Wirschum, N., Jameson, A.: Depth and breadth-first processing of search result lists. In: Ext. Abstracts CHI 2004, pp. 1539–1539. ACM Press, New York (2004)
17. Cutrell, E., Guan, Z.: What are you looking for? An eye-tracking study of information usage in Web search. In: CHI 2007- Proceedings - Gaze & Eye Tracking, San Jose, California, USA, April 28 - May 3 (2007)
18. Pan, B., Helmbrooke, H.A., Gay, G.K., Granka, L.A., Feusner, M.K., Newman, J.K.: The Determinants of Web Page Viewing Behavior: An Eye Tracking Study. In: Proceedings of Eye Tracking Research & Applications: Symposium 2004, pp. 147–154 (2004)
19. Lohse, G.L.: Consumer eye movement patterns on Yellow Pages advertising. Journal of Advertising 26, 61–73 (1997)
20. Rayner, K., Rottello, C.M., Stewart, A.J., Keir, J., Duffy, S.A.: Integrating Text and Pictorial Information: Eye Movements when Looking at Print Advertisements. Journal of Experimental Psychology: Applied 7, 219–226 (2001)
21. Flickr API, `http://www.flickr.com/services/api`
22. `http://forrest.psych.unc.edu/research/alscal.html`
23. Campbell, C.S., Maglio, P.P.: A robust algorithm for reading detection. In: Proceedings of the ACM Workshop on Perceptual User Interfaces (2001)

Interactive and Lightweight Mechanisms to Coordinate Interpersonal Privacy in Mediated Communication

Natalia Romero, Laurens Boer, and Panos Markopoulos

Eindhoven University of Technology
{n.a.romero,l.m.boer,p.markopoulos}@tue.nl

Abstract. In this paper we describe three mechanisms that enable people to coordinate their interaction needs with others in their social network. The proposed designs are based on the Privacy Grounding Model [4] that argues the need for lightweight and interactive coordination mechanisms to support the dynamic and dialectic nature of interpersonal privacy coordination.

Keywords: Interpersonal Privacy, Common Ground, Coordination, Computer Mediated Communication, Social Communication.

1 Introduction

Systems supporting mediated communication aim to facilitate networked communities to engage in different types of interaction, which could not be easily done otherwise. By displaying relevant information about a social network, such systems can, for example, help users to engage in casual interaction with their community [2] or to elicit feelings of connectedness within long-distance family members [1]. However, users of such systems often struggle with the practicalities of keeping their interactions at a desired level [3].

The Privacy Grounding Model [4] provides a theoretical account of how people coordinate interpersonal privacy, which is suggestive of the need for communication systems to provide lightweight and interactive coordination mechanisms that support the dynamic and dialectic nature of interpersonal privacy coordination in mediated settings. The model highlights the importance to support collaborative practices that help community members establish a common understanding of their interaction needs. On the one hand such practices should help people to intentionally contribute to a shared understanding of one's need for interaction; on the other hand such contributions should require low effort comparing to the main communication practices, and allow for ambiguity of one's availability state if desired. The use of such mechanisms should minimize uncertainty and increase understanding of people's interpersonal privacy needs.

Three mechanisms were implemented to demonstrate how lightweight interactivity could support the coordination of interpersonal privacy while communicating without incurring additional effort or disclosing too detailed information if undesired. The mechanisms were deployed as extensions of an existing groupware application, Community Bar (CB) [2]. The original CB was modified so the system will assign a distinctive colour to each user, and incoming messages in the chat window will be coloured according to the sender's colour.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 832–833, 2009.

2 Interpersonal Privacy Coordination Mechanisms

Two graphical and one tangible mechanism were implemented in CB to provide representations of one's intention to interact in the chat medium. The *one-click* mechanism offers an option to click on a message received from someone in one's 'buddy list', which serves as an equivalent of a notification stating: "I've seen your message". The *drag-and-drop* mechanism offers a 'palette' with a number of predefined messages (e.g., "Coffee break?", "On the phone"), which the user could select from and drag-and-drop under an existing chat message, representing a light reaction (about one's interaction needs) regarding such a message.

Alternatively, the *Cylinder* provides a tangible interface that uses coloured LEDs to signal incoming messages and alerts sent by others. An alert in CB is similar to the 'nudge' functionality in MS Messenger, a feature used to get someone's attention.

(a) (b) (c)

Fig. 1. The Cylinder: (a) an alert arrives; (b) a time is set; (c) postponing is sent

Once an alert is received, the Cylinder goes up and lights up according to the colour assigned to the particular sender (see Fig 1). The recipient can represent her unavailability in several ways. If she does not react to the movement of the Cylinder, after one minute it automatically gets back to its original position, the light fades away and the sender of the alert receives a text message indicating that there was no reaction to the alert. If the recipient pushes down the tube, a message "I am unavailable" is sent to the sender. Alternatively, the recipient can define a time frame for later contact, and a message "I will be available in about x to y minutes" is sent to the sender. The sender can confirm that message by using the one-click mechanism on the received message, which would make the cylinder of the recipient to quickly go up and down. These mechanisms are fully functional and a long-term evaluation is planned to validate their ability to support lightweight privacy coordination.

References

1. Markopoulos, P., Romero, N., van Baren, J., IJsselsteijn, W., de Ruyter, B., Farschian, B.: Keeping in touch with the family: Home and away with the ASTRA awareness system. In: Proceedings CHI 2004 extended abstracts, pp. 1351–1354 (2004)
2. McEwan, G., Greenberg, S.: Supporting social worlds with the community bar. In: Proceedings of GROUP 2005, pp. 21–30 (2005)
3. Palen, L., Dourish, P.: Unpacking "privacy" for a networked world. In: Proceedings of SIGCHI, pp. 129–136 (2003)
4. Romero, N., Markopoulos, P.: Grounding Interpersonal Privacy in Mediated Settings. In: Proceedings of GROUP 2009 (2009)

Liberating Expression: A Freehand Approach to Business Process Modeling

Nicolas Mangano[1] and Noi Sukaviriya[2]

[1] University of California, Irvine
nmangano@ics.uci.edu
[2] IBM TJ Watson Research Center, Hawthorne, NY
noi@us.ibm.com

Abstract. Tools that support business process modeling are designed for experienced users to draw a process with precision and professional appearance. These tools are not conducive to sketching quick business design ideas.This demo proposal presents Inkus, a non-intrusive business process sketching tool which allows freehand sketches of process ideas and slowly brings the users to the required common business vocabulary. Our goal is to help unleash creativity in business designers and enrich the design process with values beyond drawing.

Keywords: Sketching, business process modeling, use case modeling, low-fidelity UI, creativity support.

1 Overview

Business process modeling (BPM) has become an important part of the business world. BPM helps businesses understand their own processes and increase the efficiency of their organizations. Unfortunately, the tools currently available to make such models are rigid and bulky to use. Although these CAD-based tools provide several advantages in terms of analysis and documentation, they limit BPM designers to a structured language that may not be conductive to the creative process [1]. Instead of relying on these tools to build the initial model, BPM designers typically work out a preliminary design on paper first [2]. With this in mind, we developed Inkus as a non-intrusive environment for creating freehand sketches of business process models. Inkus provides the user with a drawing space, where he or she can create freehand sketches, and a project management space, where a series of visualizations provide feedback on the sketches in the drawing space.

2 Sketching with Inkus

Inkus uses a hybrid approach of informal sketches and formal elements. The user can sketch a crude solution of their business process design, and later annotate their sketch such that Inkus recognizes it as a formal BPM notation. Inkus allows this by first presenting the user with an empty field where he or she can work out the business process design. The user is allowed to sketch freely without any commitments regarding vocabulary or formalized components. Inkus deliberately refrains from

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 834–835, 2009.

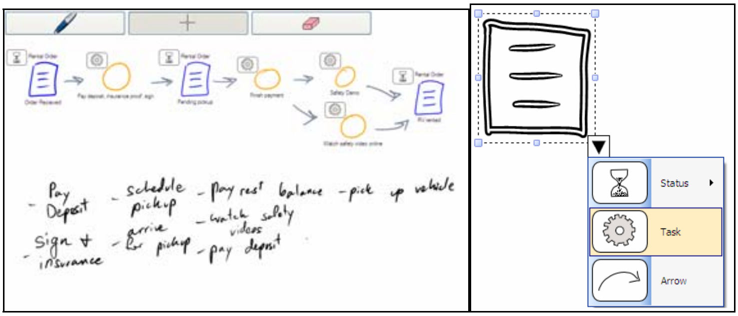


Fig. 1. Elements in the Inkus Sketching Canvas (Left) can be formalized using the lasso menu (Right)

automatically interpreting the drawings because an incorrect interpretation may be disruptive. Once the user is ready to refine their sketch into a more formal state, a selection tool can be used to assign a formal vocabulary to the elements of a sketch (see Fig 1). Inkus was built with the insight that design is not a linear process, and so the components can be left without a name or any other details until when the user is ready to specify them. Currently, Inkus provides the user with the vocabulary to define a rudimentary Artifact-Central Model.

As the user begins to assign meanings to the stroke clusters, Inkus starts providing feedback to the user through the project management space (not pictured). The project management space is dependent on the taxonomic assignments within the drawing space. Inkus uses these to build an internal model, which is then used to present alternative visualizations to the user. Since sketches can become messy and be difficult to manage over a long period of time, alternative visualizations, such as the Use Case View, allow the user with a better understanding of the broader consequences that a change may have on their model.

Drafting business processes using a sketch-based medium may be a promising approach. It offers a quick way to draw diagrams compared to the CAD tools that process designers currently use. Through these quick sketches, Inkus allows process designers to express concerns they would normally not be able to in more rigid media, thereby easing the cognitive burden of holding these concerns in their memory. The simple and lightweight UI of Inkus does not impose a steep learning curve, and the process models could also be shared easily with clients who may not be familiar with the domain. In creating a flexible and rapid medium like Inkus, we hope to allow the designer to maximize their creative potential.

References

1. Chen, Q., Grundy, J., Hosking, J.: An E-Whiteboard Application to Support Early Design Stage Sketching of UML Diagrams. In: 2003 IEEE Symposium on Human Centric Computing Languages and Environments, HCC, pp. 219–226. IEEE Computer Society, Washington (2003)
2. Sukaviriya, N., Sinha, V., Ramachandra, T., Mani, S., Stolze, M.: User-Centered Design and Business Process Modeling: Cross Road in Rapid Prototyping Tools. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4662, pp. 165–178. Springer, Heidelberg (2007)

Multimodal Interaction with Speech, Gestures and Haptic Feedback in a Media Center Application

Markku Turunen[1], Jaakko Hakulinen[1], Juho Hella[1], Juha-Pekka Rajaniemi[1], Aleksi Melto[1], Erno Mäkinen[1], Jussi Rantala[1], Tomi Heimonen[1], Tuuli Laivo[1], Hannu Soronen[2], Mervi Hansen[2], Pellervo Valkama[1], Toni Miettinen[1], and Roope Raisamo[1]

[1] University of Tampere, Tampere Unit for Computer-Human Interaction, Finland
firstname.surname@cs.uta.fi

[2] Tampere University of Technology, The Unit of Human-Centered Technology, Finland
firstname.surname@tut.fi

Abstract. We demonstrate interaction with a multimodal media center application. Mobile phone-based interface includes speech and gesture input and haptic feedback. The setup resembles our long-term public pilot study, where a living room environment containing the application was constructed inside a local media museum allowing visitors to freely test the system.

Keywords: Speech, haptics, gestures, multimodal interaction, media center.

1 Media Center Application and Its Multimodal Interface

This demonstration shows how speech, gestures, and haptic feedback can provide an efficient and accessible interface for a media center for different user groups (regular users, blind users, visually impaired users, and physically impaired users). The media center application [1] is based on a low-cost PC (Athlon X2 3800) running the media center server software, a mobile device (Nokia N95) for interacting with the user and running the client software, a wireless access point to connect these together, and a high-definition 40” digital television displaying the user interface. The server software implements the media center application functionality and includes speech recognition and speech synthesis. The mobile phone has embedded gesture recognizer, speech recognizer, haptic feedback controller, speech synthesizer, and GUI. This setup is illustrated in Figure 1.

The media center application offers the same functionality as the other applications and set-top boxes in this context, e.g., users are able to watch and record television broadcasts, listen music, and view photographs. Currently, it provides full control over digital television content, including an advanced electronic program guide (EPG). As illustrated in Figure 1, the system has two graphical user interfaces. A large television screen acts as the main GUI, while the mobile phone GUI provides additional information.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 836–837, 2009.

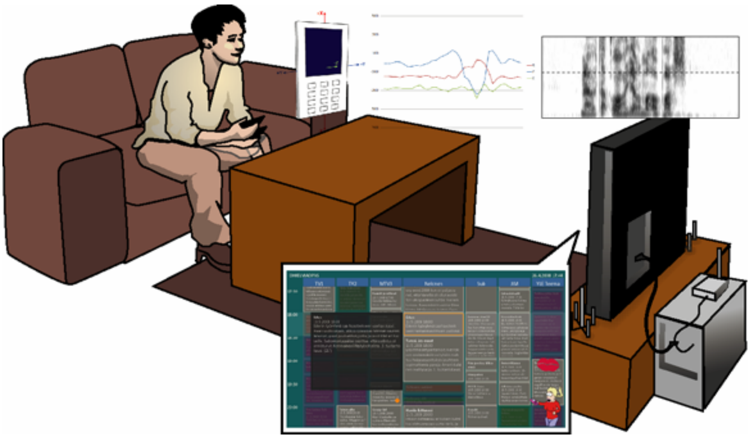


Fig. 1. Media center setup and the EPG user interface

Users have full control over the media center via its speech user interface. For example, it is possible to record multiple episodes with a single utterance (*"Record all the Tom the Tractor shows this week*"), and to highlight programs based on their genre ("*Show me all the children programs tomorrow morning*"). The mobile phone keypad and gestures can be used for navigation and selections either independently or combined with each other. Gestures are made by moving the mobile phone in specific patterns. For visually impaired users, the system provides a tight integration between the Focus + Context GUI and synthesized speech output. Both the telephone and the mobile phone include speech output capabilities. In addition, haptic feedback is given using the vibration component of the mobile phone using a set of nine haptic patterns (hapticons). For physically impaired users, we have implemented a wheelchair mounted wireless microphone solution with voice activity and blow detection.

The first version of the media center application was demonstrated in a long-term public pilot study in a local museum (Rupriikki, Tampere) between June 2008 and March 2009. In addition, laboratory experiments have been carried out [2]. Here, we demonstrate a fully functional live system with additional functionality implemented for later pilot studies with physically and visually impaired users.

This work is supported by the Technology Development Agency of Finland (TEKES) under the Ubicom-programme in the "Ambient Intelligence Based on Sound, Speech and Multisensor Interaction"-project (TÄPLÄ, grant 40223/07).

References

1. Turunen, M., Hakulinen, J., Hella, J., Rajaniemi, J.-P., Melto, A., Mäkinen, E., Rantala, J., Heimonen, T., Laivo, T., Soronen, H., Hansen, M., Valkama, P., Miettinen, T., Raisamo, R.: Multimodal Media Center Interface based on Speech, Gestures and Haptic Feedback. In: Proceedings of Interact (2009)
2. Turunen, M., Melto, A., Hella, J., Heimonen, T., Hakulinen, K., Mäkinen, E., Laivo, T., Soronen, H.: User Expectations and User Experience with Different Modalities in a Mobile Phone Controlled Home Entertainment System. In: Proceedings of MobileHCI 2009 (2009)

Social Circles: A 3D User Interface for Facebook

Diego Rodrigues and Ian Oakley

Lab:USE, Department of Mathematics and Engineering, University of Madeira
Campus Universitário da Penteada, Funchal, 9000-390, Portugal
diegoajrodrigues@gmail.com, ian@uma.pt

Abstract. Online social network services are increasingly popular web applications which display large amounts of rich multimedia content: contacts, status updates, photos and event information. Arguing that this quantity of information overwhelms conventional user interfaces, this paper presents Social Circles, a rich interactive visualization designed to support real world users of social network services in everyday tasks such as keeping up with friends and organizing their network. It achieves this by using 3D UIs, fluid animations and a spatial metaphor to enable direct manipulation of a social network.

Keywords: Social network, visualization, direct manipulation, 3D UI, animation.

1 Introduction

Web 2.0 has ushered in an age of media rich, connected and highly interactive web applications. Among the most popular of these are online social networking services, such as Facebook (www.facebook.com) which boasts millions of users and hits per day. The appeal of online social networking services is clear: they enable users to easily form persistent networks of friends with whom they can interact and share content. However, online social network services have rapidly evolved into highly complex systems which contain a large amount of personally salient information derived from large networks of friends [1]. This paper argues that the amount of information presented in these systems exceeds the abilities of users to view and absorb it using traditional desktop user interfaces: it is slow, inefficient and frustrating. To address this issue, it describes the design of Social Circles, a rich visual interface focused on effectively presenting the key information that users require from social network services. This approach differs from previous work on the visualization of social networks which has been largely focused on analytical tasks and inspired by the diagramming conventions of social network theory [e.g. 2].

2 Social Circles

The design of Social Circles was fundamentally informed by the desire to place all the relevant content relating to a user's friends on a single screen. The overall goal is to enable an at-a-glance overview of activity in the entire network and rapid, seamless navigation to particular content items of interest. In order to achieve this, the system uses a spatial metaphor and the principles of direct manipulation. It argues that a spatial

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 838–839, 2009.

Fig. 1. a) User categorization UI (left), **b)** Rows UI (mid) and **c)** Circles UI (right)

metaphor can communicate important structural aspects of a user's network (such as the "social distance" of particular individuals, or the presence of particular salient groups and clusters) in a way that conventional UI techniques, involving the largely textual construction of lists or groups, cannot. Furthermore, an animated spatial interface supports direct interactions in the form of not only selection, but also hovering over items to highlight and dragging content to alter properties and change configurations.

Reflecting these motivations, the Social Circles UI represents a user's friends as visual icons resembling name cards (or Polaroid photos) arranged in 3D space. It allows users to classify friends according to three levels of social distance (best friends, social friends and acquaintances) and illustrates this distinction spatially. Two arrangements are shown in Figure 1, one based on differently sized cards arranged in three horizontal rows and the other on cards in three concentric circles. In the former UI, card size represents social distance while in the later it is the radius of the circle. In both interfaces, users can promote and demote friends simply by selecting and dragging them to desired new level. The cards also display extra information about the friends they depict. A border to the cards (shown yellow in the figures) and conspicuous pictorial icons highlight friends with recently updated status text or photo albums. Hovering over a card causes it to expand in place and reveal a friend's name; clicking causes it to move to the center of the screen and show the most recent text status. The card can then be clicked again to display a full screen profile page of the friend, or the close widget (in the corner of a card) can be selected to collapse the card back down to its default presentation. The interfaces differ in their presentation of the dense social network content: the rows UI uses a carousal metaphor which scrolls content in response to user input, whereas the circular UI clusters friends into alphabetic groups which expand when the cursor hovers over them. We are currently exploring the usability and scalability of these different design choices.

Social Circles is implemented as a cross-platform Facebook application using Adobe Flash, the Papervision3D engine and the GreenSock tweening platform.

In conclusion, this paper introduced Social Circles, a rich user interface for online social network services using a spatial metaphor, 3D rendering and animation. Future development of this work will include adding an interface for photo exploration and subjective and objective user study. We will also release a public beta of Social Circles to assess the real world applicability of this interface, design and approach.

References

1. Joinson, A.N.: Looking at, looking up or keeping up with people?: motives and use of facebook. In: Procs. CHI 2008 (2008)
2. Perer, A., Shneiderman, B.: Integrating Statistics and Visualization: Case Studies of Gaining Clarity during Exploratory Data Analysis. In: Procs of CHI 2008 (2008)

Socio-Technical Evaluation Matrix (STEM): A Collaborative Tool to Support and Facilitate Discussion on Socio-Technical Issues of a Design Process

Souleymane Boundaouda Camara and José Abdelnour-Nocera

Thames Valley Univeristy, St Mary's Road, Ealing, London, W5 2RF, UK
souleymane.camara@tvu.ac.uk

Abstract. STEM is an interactive web application designed to support and facilitate socio-technical discussions in a collaborative environment. This interactive demo shows how STEM aims to encompass issues with collaborative tools such as organisation of discussion threads by relevancy, interdependency and iteration of previous discussions.

Keywords: Collaborative work tool, socio-technical evaluation, dependability.

1 STEM: Overview

STEM[1] is an interactive web application designed to support and facilitate collaboration in IT system design where, social and technical factors need to be matched to identify critical issues towards system acceptance, use and fitness for purpose. Emerging from an ongoing research in the Village eScience for Life (VeSeL[2]) project, this tool tries to offer an improved environment for group discussion based on pre-defined framework, methods or models.

The demonstration of the tool will highlight the rational for STEM and an interactive online session of its main features based on its usage in the VeSeL project as described in the following sections.

2 Why STEM: Features and Rational

STEM builds on a rationalised approach and philosophy of understanding the use of HCI moderated collaborative work which offers:

- Distinction of Critical Social and Technical Factors of System Design/Evaluation. Socio-technical evaluation highlights the importance and dependability of a system based the social impact of actors (users, designers) and the technological impact. STEM allows users to comment on the implication of design decisions for both users and the technology. A decision that

[1] STEM overview available at: http://itcentre.tvu.ac.uk/~VeSeL/matrixsample/.
[2] Visit http://www.lkl.ac.uk/projects/vesel/ for more on the VeSeL project.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 840–841, 2009.

is expressed for one is therefore evaluated in its implication for the other in a form based interface.

- Adaptability to Pre-Defined Design Frameworks and Methodologies. Most collaborative tools, open source[3] or proprietary (Lotus Notes[4], Freeman[5], Huddle[6]) are often too generic with no underpinning HCI framework or methodologies. They aim to manage workflows, documents and communications within a group. With STEM, users propose dimensions and relevant attributes of an approach to adopt and evaluate their processes against them. The demonstration will show how these dimensions and attributes can be defined.
- Better Organisation of Discussion Thread: Many collaborative tools tend to display discussion entries in a chronological manner or simply alter previous entries as the discussions feed in new or better ideas (Google docs, Wikipedia). Following a topic thread on these tools is therefore very difficult. STEM interface allows participants to decide if their contribution is in support or conflicts with an existing main entry. It then displays the comments by relevance and in relation to each other saving time and motivating participants to easily follow/contribute to a discussion. Furthermore, a moderator of a discussion can summarise the discussion by attributes if there are too many or sufficient entries.
- Iterative Capability for Consistent and Continuous Design Processes: Often in IT system design, processes follow an iterative route. Requirements are gathered, refined and turned into a design prototype. Then, prototypes are discussed, summarised for evaluation and so on. STEM goes a step further by integrating the summary of previous stages into the discussion of the actual stage. The reason for this is that some dimensions may not or can not be addressed in previous stages. By bringing them forward, STEM maximises the chance of them being acted upon at the earliest possibility.

These features are currently being evaluated to refine the next version of the tool.

3 Conclusion

STEM is not a platform for technicality elucidation, document exchange or a heavy communication-driven collaborative tool. Rather, it is an environment for discussing cross-party issues on a predefined approach. It is an attempt to address the shortcomings of current collaborative tools in HCI based design.

[3] Open source tools: http://php.opensourcecms.com/scripts/show.php?catid=4&cat=Groupware
[4] Lotus Notes by IBM: http://www-01.ibm.com/software/lotus/
[5] Freeman Online: http://www.freemanco.com/showorganizers/collaborativeTools.jsp
[6] Huddle: http://www.huddle.net/what-is-huddle

Take Three Snapshots - A Tool for Fast Freehand Acquisition of 3D Objects

Gabriele Peters and Klaus Häming

University of Applied Sciences and Arts, Computer Science, Visual Computing, Emil-Figge-Str. 42, D-44221 Dortmund, Germany
{gabriele.peters,klaus.haeming}@fh-dortmund.de

Abstract. We introduce a tool which allows an untrained user to take three images of an object freehand with a simple consumer camera. From these images a 3d model of the visible parts of the object is reconstructed within seconds and visualized realistically. From a research point of view we propose solutions for three weaknesses of the state-of-the-art reconstruction pipeline. These contributions allow for a more robust and a considerably faster reconstruction process than before, which can be used, e.g., to create new types of interfaces or to assist in creating virtual environments.

Keywords: Object Acquisition, Object Reconstruction, User Interfaces, Interaction Techniques, Visualisation Techniques, Usability.

1 Introduction

There is a growing demand for realistic 3d models of real-world objects in a large number of fields of applications. One way to obtain such models is the cumbersome manual creation of 3d point clouds and meshes, which is still the prevalent method. Automatic methods require complex technical equipment or highly skilled experts or both. Existing methods for object acquisition such as [1] are far from being user-friendly or accessible to the broad public.

2 A Solution with an Intuitive Interface

We present a prototype of an intuitive acquisition tool which allows an untrained user to take a few images of an object freehand from slightly different but random viewpoints with an uncalibrated consumer camera. From these images a 3d model of the visible parts of the object is reconstructed. The model is visualized as a realistic, freely rotatable model. The scientific background of our demonstration is visualized in Fig. 1. Our contribution to the image processing step consists in a new combination of feature detection and description based on [2] providing a higher robustness. For consistency filtration we have developed a two-stage process based on [3], which facilitates a faster selection of relevant features. Our third contribution consists in a new method to texture-map the point cloud, resulting in an arbitrarily rotatable 3d surface.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 842–843, 2009.

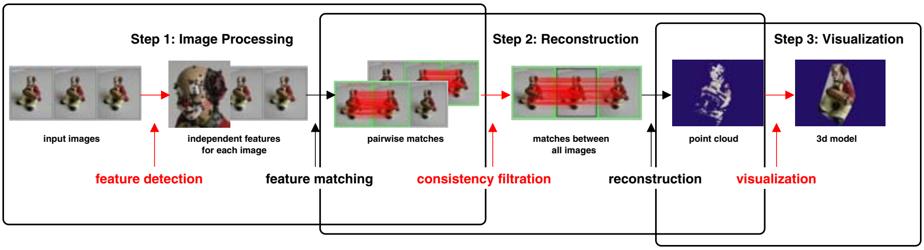


Fig. 1. 1. Image processing with feature detection and matching between image pairs. 2. Filtration of consistent feature matches and reconstruction of 3d points. 3. Visualization with texture mapped on the point cloud. Those parts of this pipeline we contribute to are highlighted red.

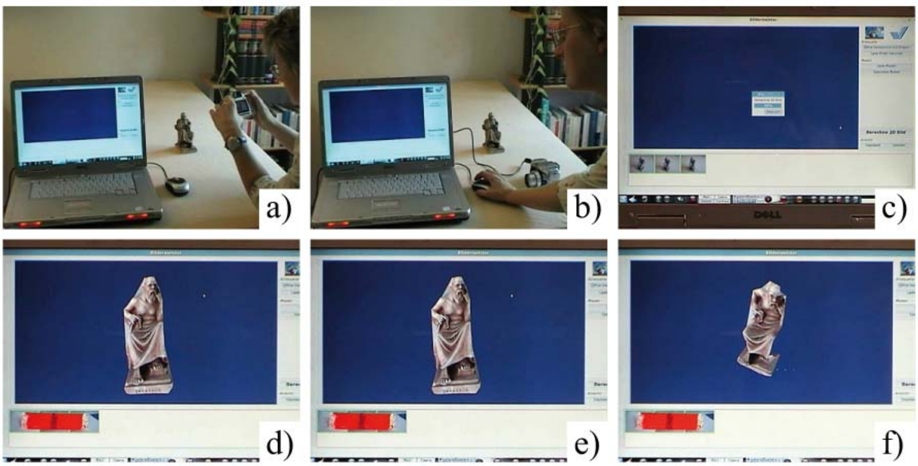


Fig. 2. a) Take snapshots. b) Connect camera to computer, initiate image loading and model creation. c) Progress bar while the model is generated. d) Generated 3d model is displayed. e),f) Visualization of model from different viewpoints.

Fig. 2 shows the operation of our system in practice. We hope this tool is not only useful for professional applications but also attractive to the broad public.

References

1. Saxena, A., Sun, M., Ng, A.Y.: Make3d: Depth perception from a single still image. In: Fox, D., Gomes, C.P. (eds.) AAAI, pp. 1571–1576 (2008)
2. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60(2), 91–110 (2004)
3. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24(6), 381–395 (1981)

Blended Interaction Design: A Spatial Workspace Supporting HCI and Design Practice*

Florian Geyer

Human-Computer Interaction Group, University of Konstanz
Universitätsstraße 10, 78464 Konstanz, Germany
florian.geyer@uni-konstanz.de

Abstract. This research[1] investigates novel methods and techniques along with tool support that result from a conceptual blend of human-computer interaction with design practice. Using blending theory with material anchors as a theoretical framework, we frame both input spaces and explore emerging structures within technical, cognitive, and social aspects. Based on our results, we will describe a framework of the emerging structures and will design and evaluate tool support within a spatial, studio-like workspace to support collaborative creativity in interaction design.

Keywords: interaction design, design research, conceptual blending, blending theory, design tools, creativity support.

1 Motivation

Interaction design is broadly perceived as a recent trend in human-computer interaction (HCI) that emphasizes on designing user experiences with technology [1,2]. Its interpretations differ between the two converging perspectives: "Interaction design as a design discipline" and "Interaction design as an extension of HCI" [3]. In recent years, HCI research developed methods, techniques, and tools that aim on supporting interaction designers in practice. However, due to the complexity of the developed methods and inappropriate integration, application in practice remains sparse [4]. One of the reasons for this is that interaction design research "has not been grounded in or guided by a sufficient understanding and acceptance of the nature of design practice" [5, p.56]. Consequently, there is a growing interest in the role and nature of design in HCI research [6,7]. These contributions have already increased awareness in interaction design practice [1,8] and will influence research over time. However, due to the tension between interaction design and HCI [3] as well as HCI and software engineering (SE) [9], attempts specifically addressed at supporting design practice in HCI have not yet been successful [5]. One possible solution to this discussion is the proposal to merge, or rather blend, these disciplines into a new one [10, p.198]. Tensions

* Advisor: Harald Reiterer.

[1] This dissertation work is supervised by Prof. Dr. Harald Reiterer, Human-Computer Interaction Group, University of Konstanz, Germany, harald.reiterer@uni-konstanz.de

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 844–847, 2009.

between design thinking and engineering philosophies may then result in a distinctive, creative domain that has emergent structures of its own.

2 Research Approach

In order to effectively blend design practice and HCI it is necessary to understand and describe design rationality [5], a notion that has been labeled with concepts such as the *reflective practitioner* [11] or the *thoughtful designer* [7]. More detailed descriptions of disciplined behavior of designers can be found in [1,12] and [13]. These authors show that following activities, among others, are fundamental for design rationality: *Sketching* is at the core of design and is used to explore relationsships between design ideas, form and function as well as big picture and details. Exploring *many different alternatives* in a *conversational* way while constantly *shifting focus* between *ill-defined* problems and solutions is the prevalent work style. *Spatial* structures, like *design studios* that make judgements visible and application of *design critique* techniques are the foundation of design argumentation. Eventually, *process awareness* is essential to deal with the complexity of design tasks.

However, HCI research struggles with integrating these *designerly ways* [12] into interaction design practice. One reason for this is that, due to a focus on engineering, most methods actually employed in practice impose a *linear progression* through design tasks [14]. Another reason is the fact that user interface design tools merely focus on the task of *constructing a single built* instead of *supporting* the messy nature of designer's explorative and spatial thinking modes [5]. A possible solution therefore has to consist of two blends, namely a blend of methods and techniques, as well as a blend of work styles and tool support.

2.1 Theoretical Foundations

We ground our research within blending theory, which is a recent framework developed by Fauconnier and Turner [15] to deal with online meaning construction. Except from Imaz and Benyon [10] and Markussen and Krogh [2], blending theory has received only little consideration in HCI research. To account for a deep understanding of how people construct meaning as they think and act, Fauconnier and Turner introduced the notion of "conceptual blending" [15]. They describe the basic principles as follows: "A conceptual blend operates in two input mental spaces to yield a third space, the blend. Partial structure from the input spaces is projected into the blended space, which has emergent structure of its own." [15, p.149]. Notable here is that the blend develops a new structure that did not exist in either of the two input spaces. This is realized by modifying existing structures through the cognitive operations of composition, completion, and elaboration. Hutchins [16] added the notion of "material anchors" to the original framework. In doing so, he claims awareness to the relevance of physical form within input spaces and the bodily interaction with the physical world. By employing both concepts as a framework for describing and mapping key features of the input spaces, we are confident to create a distinctive blend that has its own structure. Nevertheless, we also seek to provide a reference in application of blending theory as a foundation to interaction design.

2.2 Method and Proposed Solution

Based on the described foundations, we are blending the two mental and physical input spaces HCI and design practice into a novel emerging structure that we call "Blended Interaction Design". In previous research, we gained valuable expertise with the integration of agile SE methods into HCI that we will bring into this input space [17]. We expect that the tension between the proposed input spaces leads to a novel, more creative perspective.

As a first step, we explored the structure of our input spaces and examined related work. We found that an important prerequisite for a successful blend is the reduction of complexity that results from the evolution of both spaces. Therefore, we decided to frame our input spaces to include only those opposing structures that we found most significant within the technical, cognitive, and social aspects.

Simulation. Visualizing design ideas and concepts is the integral activity within both input spaces that drives the design progress. While HCI has a focus on modeling and prototyping, sketching as the central design activity is more opportunistic [1]. Overall, these different activities lead to two separate perspectives on process awareness, shifting between problem space and solution as well as the generation of alternatives.

Creativity. Creative work styles in design differ fundamentally from work styles with engineering tools commonly used in HCI [14]. Most design tools do not explicitely support divergent thinking, generation of alternatives, and decision-making, which often leads to fixation effects [14] and undesired results. In contrast, design techniques emphasize on the spatial nature of synthesis in design, where juxtaposed artifacts allow reflection-in-action [11].

Collaboration. Both input spaces are interdisciplinary and inherently social domains. Therefore, a blend has to accommodate common understanding in representations, expressions, and evaluation methods. Engineering methods in HCI often bury design artifacts and decisions in file systems. Studio settings and spatial workspaces may provide a more practical environment for collaborative design.

Each of these frames has structures that contribute to the blend. We are currently applying the principles of blending theory to construct an image of a proposed solution. Based on the results, we will then develop and describe a *framework of methods* that emerges from the mental structures of the blend. Concurrently, we will design and implement visualization and interaction techniques that integrate the physical structures within a *spatial workspace* based on a computation-augmented design studio setting with interactive whiteboards, large high-resolution displays, and tabletops. Therefore, we seek to combine physical interaction concepts that built upon the principles of natural interaction (e.g. digital pen and paper, tangible design artifacts) with our object-oriented multiscale interface paradigm ZOIL [18]. Eventually, we will assess our efforts with evaluation measures and case studies. For evaluation of developed methods, we will draw upon feedback from the research community and our industrial partners while we seek to evaluate the physical workspace within lecture projects and class assignments.

3 Expected Contributions

This research will contribute to interaction design and HCI research within several areas. On the one hand, we will provide reference and case studies for application of blending theory as a foundation to interaction design. On the other hand, we will contribute to the appreciation and support of design practice in HCI. Eventually, we will - based on elaborated needs for simulation, creativity support and collaboration - design, and evaluate novel visualization and interaction techniques for tool support within computation-augmented design studio settings.

References

1. Buxton, B.: Sketching User Experiences: Getting the Design Right and the Right Design. Morgan Kaufmann, San Francisco (2007)
2. Markussen, T., Krogh, P.G.: Mapping Cultural Frame Shifting in Interaction Design with Blending Theory. International Journal of Design 2(2), 5–17 (2008)
3. Löwgren, J.: Interaction Design. from Interaction-Design.org (2008), `http://www.interaction-design.org/encyclopedia/interaction_design.html`
4. Rogers, Y.: New Theoretical Approaches for HCI. Annual Review of Information, Science and Technology 38, 1–43 (2004)
5. Stolterman, E.: The Nature of Design Practice and Implications for Interaction Design Research. International Journal of Design 2(1), 55–65 (2008)
6. Winograd, T., Bennett, J., De Young, L., Hartfield, B.: Bringing Design to Software. Addison Wesley Professional, Reading (1996)
7. Löwgren, J., Stolterman, E.: Thoughtful Interaction Design. MIT Press, Cambridge (2004)
8. Moggridge, B.: Designing Interactions. MIT Press, Cambridge (2006)
9. Seffah, A., Gulliksen, J., Desmarais, M.C.: Human-Centered Software Engineering - Integrating Usability in the Software Development Lifecycle. Springer, Heidelberg (2005)
10. Imaz, M., Benyon, D.: Designing with Blends: Conceptual Foundations of Human Computer Interaction and Software Engineering. MIT Press, Cambridge (2007)
11. Schön, D.A.: The Reflective Practitioner: How Professionals Think in Action. Basic Books (1983)
12. Cross, N.: Designerly Ways of Knowing. Board of International Research in Design. Birkhäuser, Basel (2007)
13. Lawson, B.: How Designers Think, The Design Process Demystified. Architectural Press (2005)
14. Terry, M., Mynatt, E.D.: Recognizing Creative Needs in User Interface Design. In: Proceedings of the 4th Conference on Creativity & Cognition, pp. 38–44 (2002)
15. Fauconnier, G., Turner, M.: The Way We Think: Conceptual Blending and the Mind's Hidden Complexities. Basic Books (2003)
16. Hutchins, E.: Material Anchors for Conceptual Blends. Journal of Pragmatics 37(10), 1555–1577 (2005)
17. Memmel, T., Geyer, F., Rinn, J., Reiterer, H.: Tool-Support for Interdisciplinary and Collaborative User Interface Specification. In: Proceedings of the IADIS IHCI (2008)
18. Jetter, H., König, W.A., Gerken, J., Reiterer, H.: ZOIL - A Cross-Platform User Interface Paradigm for Personal Information Management. In: CHI 2008 Workshop - The Disappearing Desktop: Personal Information Management (2008)

Designing an Artificial Robotic Interaction Language*

Omar Mubin

Department of Industrial Design, Eindhoven University of Technology, The Netherlands
{o.mubin,c.bartneck,l.m.g.feijs}@tue.nl

Abstract. The project described hereunder focuses on the design and implementation of a "Artificial Robotic Interaction Language", where the research goal is to find a balance between the effort necessary from the user to learn a new language and the resulting benefit of optimized automatic speech recognition for a robot or a machine.

Keywords: Artificial Languages, Speech Interaction, Automatic Speech Recognition.

1 Introduction

Speech is a natural means of information exchange for humans. Therefore, improving speech interaction technology in Human Computer Interaction could lead to a more pleasant interaction. Speech Interaction is confronted by various issues, such as: ambiguity in natural dialogue, un-robust speech recognition and un-synchronization between software and hardware [1]. Recent attempts to improve the quality of the technology of automatic speech recognition for machines have not advanced enough. The limitations prevailing in current speech recognition technology for natural language is a major obstacle behind the unanimous acceptance of Speech Interfaces. Existing speech recognition is just not good enough for it to be deployed in natural environments, where the ambience also influences its performance.

Generally in speech interfaces the focus is on using natural language and given their unsuitability, it is time to find a different balance in the form of a new language. Recent research in speech interaction is already moving in this direction, as stated in [2], constraining language is a plausible method of improving recognition accuracy. In [3] the user experience of an artificially constrained language ("Speech Graffiti") was evaluated and it was concluded that 74% of the users found it more satisfactory than natural language and also more efficient in terms of time. The field of handwriting recognition has followed a similar road map. The first recognition systems for handheld devices, such as Apple's Newton were nearly unusable. Palm solved the problem by inventing a simplified alphabet called Graffiti which was easy to learn for users and easy to recognize for the device. Therefore we aim to construct an "Artificial Robotic Interaction Language" where an artificial language is a *language deliberately invented or constructed, especially as a means of communication in computing* [4]. In linguistics, there are numerous artificial languages which address a user

* Advisors: Christoph Bartneck and Loe Feijs.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 848–851, 2009.

perspective by making communication between humans easier and/or universal; however there has been little or no attempt to optimize a spoken artificial language for automatic speech recognition. Therefore the main goal of our research is constructed on the basis of two sub goals. Firstly the language should be learnable by the user and secondly, the language should be optimized for efficient automatic speech recognition by a machine or a robot.

2 Overview of Artificial Languages

As a first step of our research an overview of artificial languages has been carried out [5], to ascertain what we could learn from existing Artificial Languages, especially in reference to what could be easier to learn for humans. The overview was carried out across two aspects, namely morphology or grammar and phonology. Various encyclopedias [6] define the major properties of a language of which morphology and phonology are two key aspects.

2.1 Morphological Overview

In summary, there are two major approaches of morphological design amongst artificial languages: The first is to have very few grammatical markings, leaving it to the interpretation of the speakers and the context, and secondly, some languages have inflections but their grammatical rules are consistent across all words. The question that emerges is that which grammar type would be easier to learn and which would be less ambiguous. Shown below (see Table 1) are the two main approaches of grammar design amongst artificial languages. Grammar-I has very few grammatical markings (as indicated by the 'X') and therefore interpretation is determined by the speakers, the context or the word order. Grammar-II has inflections but the grammatical rules are consistent across all words within each category.

Table 1. Two main grammar types in artificial languages

Grammatical Category	Grammar I	Grammar II
Aspect	X	X
Case	X	Basic Levels: for e.g. Possessive, Nominative
Gender	X	Male, Female
Mood/Modality	X	X
Numbering	X	Singular, Plural
Person References	First, Second, Third	First, Second, Third
Tense	X	Past, Present, Future
Polarity	Positive, Negative	Positive, Negative
Voice	X	Active, Passive
Definiteness (Articles)	X	Definite (the), Indefinite (a)
Word Order	SVO	SVO

2.2 Phonological Overview

Another important metric upon which artificial languages could be overviewed was determined to be the domain of Phonetics. Deciding the manner in which to use phonemes as part of our phonological classification of artificial language was an important decision. Extending from our research goal of designing an interaction language that is easy to learn for humans, we extracted a set of the most common phonemes/segments present in the major languages of the world, based on the number of speakers of a language as indicated in the Ethnologue [7]. The overview utilized the UCLA Phonological Segment Inventory Database (UPSID), see [8] and [9]. The database provides a large inventory of all the existing phonemes of 451 different languages of the world. The number of phonemes documented in the database amount to 919. What we were seeking for was a list of phonemes found in only the major languages of the world. This resulted in a net total of 23 phonemes [5]. Interesting trends were observed; certain dental consonants were not found in any of the artificial languages that we overviewed. One reason why this might have occurred is that most artificial languages stem from Germanic or Western languages, whereas the dental consonants are found in Indic or Asian languages only. Trends that have been observed in natural languages with regards to the most common consonants (for e.g. ‘m’, ‘p’) were to some extent replicated for the case of artificial languages. The mirroring effect between natural and artificial languages extended to vowels as well. The vowels ‘a’, ‘o’ and ‘u’ were found in all the artificial languages that we classified and the vowels ‘i’ and ‘e’ were absent in only one artificial language.

3 The Design of the Artificial Interaction Language

We have presented a morphological overview of artificial languages where, two primary grammar types were discussed. In the future, we aim to evaluate which of the mentioned grammar types will be easier to learn for our intended artificial language and which will be less ambiguous, using methods as advocated in [10]. Moreover, our phonological overview has revealed a set of phonemes that might be desirable to include in the artificial language to render it conducive for human learnability. However for both aspects of morphology and phonology what also needs to be determined is how both could contribute to improve speech recognition. For example unique phonemes that have less confusion amongst them would be easier to recognize [11]. Similarly, selecting a particular grammar type could also influence the quality of speech recognition, and we aim to determine this in the future. Another interesting variable is word length, as shorter words tend to be confused with each other in automatic speech recognition. Therefore we also aim to evaluate the role the length of a word could play in improving the quality of speech recognition.

As a first step in the design process we aim to inherit the vocabulary set or word concepts of the simple artificial language Toki Pona [12]. It has 118 word concepts and sufficiently caters for the needs of a simple language. We aim to adapt the pronunciation of the words of Toki Pona based on the requirements of word length and phonetic information. For example, given that Toki Pona is a simple language it has some words which are very short; of course to be easier to learn for humans. However to assist speech recognition, some of its words will need to be elongated based on a specific methodology, which will attempt to improve the phonetic discernability of words and also be scalable to allow the generation of new words. Additionally, we aim to start the design from Grammar Type II and gradually remove grammatical markings and rules to move towards Grammar Type I, an evolutionary trend that can also be noticed in natural languages [13].

In the future, we shall also investigate speech recognition engines to ascertain "*the exact criteria that makes speech easy for machines to recognize*". As a combination of our initial endeavors we will then move towards designing the "Artificial Robotic Interaction Language". Our intention is to carry out future research in the form of several cycles as a spiral model. Each cycle typically would have four phases: requirements, design, implementation and evaluation.

Contribution to HCI. We intend to deploy our interaction language within the domain of robotics, however our proposed interaction language does not necessarily have to be restricted to robots only, but it could be applied to any behavioral products that employ speech interaction. Moreover, HCI is moving towards the domain of Ambient Intelligence, where technology is invisible in the background but there are still objects that mediate/interact, for e.g. robots. Therefore the design of an Artificial Robotic Interaction Language would lie in the heart of next generation HCI.

References

1. Kulyukin, V.A.: On natural language dialogue with assistive robots. In: ACM conference on Human-robot interaction, pp. 164–171 (2006)
2. Rosenfeld, R., Olsen, D., Rudnicky, A.: Universal speech interfaces. Interactions 8, 34–44 (2001)
3. Tomko, S., Rosenfeld, R.: Speech Graffiti vs. Natural Language: Assessing the User Experience. In: Proceedings of HLT/NAACL (2004)
4. A Dictionary of Psychology. In: Colman, A.M. (ed.) Oxford Reference Online. Oxford University Press, Oxford (2006)
5. Mubin, O., Bartneck, C., Feijs, L.: A Morphological and Phonological Overview of Artificial Languages. In: Interspeech 2009 (2009) (submitted)
6. David, C.: The Cambridge encyclopedia of language. Cambridge University Press, Cambridge (1997)
7. Gordon, R., Grimes, B.: Ethnologue: Languages of the World. Summer Institute of Linguistics, SIL International (2005)
8. UPSID Info., `http://web.phonetik.uni-frankfurt.de/upsid_info.html`
9. Maddieson, I.: Patterns of Sounds. Cambridge University Press, Cambridge (1984)
10. Mubin, O., Shahid, S., Bartneck, C., Krahmer, E., Swerts, M., Feijs, L.: Using Language Tests and Emotional Expressions to Determine the Learnability of Artificial Languages. In: CHI 2009 Conference on Human Factors in Computing Systems, pp. 4075–4080 (2009)
11. Hinde, S., Belrose, G.: Computer Pidgin Language: A new language to talk to your computer? Hewlett-Packard Laboratories
12. Toki Pona - the language of good, `http://www.tokipona.org/`
13. Grudin, J., Norman, D.: Language evolution and human-computer interaction. In: Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society, pp. 611–616 (1991)

Designing Mobile Service Experiences, the Role of Emotional Dimension*

Teresa Sarmento

Faculdade de Engenharia da Universidade do Porto
Rua Dr. Roberto Frias, s/n 4200-465 Porto, Portugal
teresasarmento@esad.pt

Abstract. This research project aims to improve the methods for translating emotional factors experienced by users into characteristics of the mobile service interfaces.

Keywords: service design, user experience, mobile interfaces, new services, service experience prototyping.

1 Motivation

The mobile phone is undoubtedly a powerful tool to connectivity and many other functions; it's like a Swiss army knife [1, 2], and goes far ahead of voice communication The Portuguese mobile phone network reaches 12,2 millions of subscribers with a 115,7% penetration rate[3]. This widespread use of mobile phones has created a huge opportunity for service providers [4], and many companies, such as banks and retailers, now offer their services through mobile networks. But to develop successful mobile services, it is important to understand customer experience requirements and to carefully design the service experience desired by customers. As the new world of opportunities for mobile services has emerged we need to classify elements connected with the interest in mobile phones and their services supporting the co-creation of value through the dedicated design of new particular and interactive applications, interfaces and its specific mobility[5]. In addition to improving core technologies, they must also focus on providing a rich and satisfying mobile-communication experience to end users [6]. The intentions of using mobile services go beyond the technology acceptance model and Herbjon Nysveen et al [7, 8] conclude, from their cross comparison study, that it depends on the service characteristics and the value of its perceived expressiveness.

In this context, it is of the greatest importance to understand user's emotional and experiential behavior[9]. In addition to improving core technologies, developers must focus on providing end users a rich and satisfying mobile-communication experience.

* Advisors: Lia Patricio and José Bartolo.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 852–855, 2009.

1.1 Literature Review and New Challenges Raised by Mobile Service Design

User experiences are very important for mobile service success, as the context of use and conditions are constantly and rapidly changing. However, service experiences and especially mobile service experiences need further research as referred by Manfred Thüring [2] factors such as aesthetic qualities and emotional experiences play an important role in addition to instrumental aspects and as he concludes further research topics concern the temporal characteristics of human–technology interaction and further dimensions of non instrumental qualities offer demanding opportunities to reach beyond ‘classical’ usability approaches.

Other previous studies [10] confirmed that emotional value and customer satisfaction affect mobile service experiences. Accordingly methods for translating user experience factors into service design require additional research. It is also recognized that service design requires interdisciplinary competences. The inter-disciplinary research can access research proposals more consistently and close partnerships with industry can help academics to develop relevant research agenda leading to interdisciplinary tools, models and frameworks [11] requiring the joint contribution of new service development, Interaction design, engineering

At this point, sketching and prototyping the service experience deserves particular attention[12]. Due to the distinctive characteristics of services (particularly intangibility and co-creation of value), prototyping services is quite different from prototyping products (the service tools and physical environment, but also the service process and the people, both customers and service employees). Several research frameworks can be useful, such as IDEO experience prototyping [13]or the service theatre framework [14]. However, a systematic approach to service experience prototyping is still lacking.

2 Research Questions

a. What are the most important experience factors that contribute to satisfaction and usage of mobile services?
b. How can these experience factors be translated into mobile service design? In particular, how can these experience factors be translated into the different components of a service experience prototyping?
c. What should be the components of a service experience prototype and how do they contribute to a better mobile service design and its testing?

3 Method

Empirical ground – the research project will involve joint work with a company which is developing a new mobile service, and comprises two main activities: the study of mobile service experiences and the development of new mobile service experience prototypes, through which new methods will be developed and tested.

Study of mobile service experiences: Through a framework of an iterative evaluation of the mobile service design, it is expected to access experiential factors and to refine them along its development and test its sequential releases. It will start with an exploratory study with a first set of interviews and observation of mobile service users to iterative walkthroughs and a survey. The research will make use in addition from a developer's point of view of several different techniques from sketching to paper prototyping and simulations in order to evaluate a longitudinal progression of the method.

Development of new methods for translating experience factors into mobile service design, through service experience prototyping. Along the new mobile service project, several releases will be iteratively developed, and several service experience prototypes will be created to translate experience factors. This iterative process will also allow for a systematic analysis of how the different service prototype components can contribute ~~for~~ to better user testing, better service design, and better results of the overall project.

4 Expected Research Contributions

Considering it as important to prototype the interface as it is to prototype the context of interaction, this research project will pursue a relationship between the service experience and service design through:

a. Better understanding of user experience factors in mobile services, identifying the factors and understanding how they impact customer satisfaction and usage of these services. Moreover, it will be interesting to understand how experience factors of mobile services differ from experience factors ~~for~~ of regular services.
b. Improvement and development of new methods for translating experience factors into service design, in particular for improving service experience prototyping.

References

1. Ballard, B.: Designing the Mobile User Experience. John Wiley & Sons, West Sussex (2007)
2. Thüring, M., Mahlke, S.: Usability, aesthetics and emotions in human-technology. International Journal of Psychology 42(4), 253–264 (2007)
3. ANACOM, A.N.d.C.: `http://www.anaco.pt` (2008)
4. Jones, M., Marsden, G.: Mobile Interaction Design. John Wiley & Sons, Ltd., West Sussex (2006)
5. Moggridge, B.: Designing Interactions. The MIT Press, Massachusetts (2007)
6. Subramanya, R.S., Yi, B.K.: Enhancing the user experience in mobile phones. Computer, 114–117 (2007)
7. Herbjon Nysveen, P.E.P., Thorbjornsen, H.: Intentions to use Mobile services: Antecedents and cross-service comparisons. Journal of the Academy of Marketing Science 33, 330–346 (2005)
8. Davis, F.D.: Perceived Usefulness, Perceived ease of use, and user acceptance of Information Technology. Mis Quarterly 13, 319–340 (1989)

9. Norman, D.: Emotional Design - Why we love or hate everyday things. Basic Books, New York (2004)
10. Lim, H., Widdows, R., Park, J.: M-loyalty: winning strategies for mobile carriers. Journal of Consumer Marketing 23/4, 208–218 (2006)
11. Cambridge University, I.: Succeeding through Service Innovation, Developing a Service Perspective on Economic Growth and Prosperity. In: Cambridge Service Science, Management and Engineering Symposium, Cambridge UK (2007)
12. Buxton, B.: Sketching User Experiences. Morgan Kaufmann, San Francisco (2007)
13. Buchenau, M., Suri, J.F.: Experience Prototyping. In: DIS 2000. ACM, Brooklyn (2000)
14. Fisk, R.P., Grove, S.J., John, J.: Interactive Services Marketing. Houghton Mifflin Company, Boston (2008)

Development of a Method for Evaluating the Usability of In-Vehicle Information Systems (IVISs)

Catherine Harvey

Transportation Research Group, School of Civil Engineering and Environment,
University of Southampton, Highfield Campus, Southampton, Hampshire, SO17 1BJ,
United Kingdom
c.harvey@soton.ac.uk

Abstract. This project involves the development of a model-based method for the evaluation of design concepts for In-Vehicle Information Systems (IVISs). The evaluation method is aimed specifically at assessing the usability of the Human Machine Interface (HMI) associated with these systems. It is aimed at the very early stages of design and is intended to provide designers with a quick and easy method of evaluating HMI concepts. The evaluation method will be validated and refined using the results from user trials conducted in a driving simulator and on-road, in an instrumented vehicle.

Keywords: Driving, Usability, Ergonomics, Interface, Evaluation, Design.

1 Project Overview

Supervision: University of Southampton: Prof. Mike McDonald(m.mcdonald @soton.ac.uk), Prof. Neville Stanton (n.stanton@soton.ac.uk), Dr. Pengjun Zheng (p.zheng@soton.ac.uk); Jaguar Cars Ltd.: Dr. Carl Pickering (cpicker6@ jaguar.com).
Research Area: Evaluation of the usability of in-vehicle interfaces.
Research topic description: This project involves the development of a method for evaluating the usability of In-Vehicle Information System (IVIS) design concepts.

2 Introduction

Secondary driving tasks are those which are not directly related to the primary task of driving, and typically include any functions relating to communication, comfort, navigation and 'infotainment'. Traditionally, secondary functions have been operated via a series of hard switches mounted on the vehicle's dashboard. Today, in the premium sector, and increasingly with volume brands, these functions are integrated into a single menu-based interactive system, with only the most high-frequency and high-importance controls left as hard switches. These 'In-Vehicle Information Systems' (IVISs) make use of a screen-based interface, which reduces

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 856–859, 2009.

the cluttered appearance of the dashboard and is an aesthetically superior solution to the traditional layout.

The ease with which a driver can interact with an IVIS is determined by the Human Machine Interface (HMI). The HMI influences a driver's ability to input information to the IVIS, to receive and understand information outputs, and to monitor the state of the system. A HMI which optimises the interaction between the driver and IVIS will enhance the 'usability' of the system and consequently improve the driver's experience of using it. Usability is comprised of many factors such as effectiveness and efficiency of use, learnability, user attitude and acceptance, flexibility, and safety [1], [2]. All of these factors, plus many more, are important in the design of a HMI, although the relative importance of each of these factors will depend on the specific application. When considering the usability of an IVIS specifically, factors such as effectiveness and efficiency of use, learnability, and safety are likely to be most important because secondary tasks must be easy to use and to learn, and, perhaps most crucially, must not present a risk to the user in terms of safety.

3 The Current State of The Art

A review of the literature has identified a necessity for an evaluation method for IVIS interfaces which is targeted specifically at the concept stage of the design process. There has been a tendency in the past to leave consideration of human-machine interaction until late in the design process when much of the design has already been finalised. This has meant that even if usability issues have been identified it is usually too late or too costly to make the necessary changes to design.

The review has also shown that existing methods for in-vehicle interface evaluations are largely qualitative, relying heavily on subjective opinions of usability, which are susceptible to bias. Examples include an evaluation method for automotive cockpits in which ergonomics experts rated attributes of the cockpit design on a rating scale where the levels of satisfaction were denoted by sad or smiley faces [3], and a checklist of statements which evaluators had to assess as being 'true' or 'false' [4]. It is evident that these evaluation techniques are open to bias from evaluators and are a relatively crude method of assessment. A more quantitative method, based on measurable attributes of usability, such as interaction times, performance error rates or user eye movements, is desirable because it would allow different interfaces to be compared directly and would also remove experimenter bias from evaluation. However, aspects of usability such as aesthetic appeal and user satisfaction are virtually impossible to assess objectively and it is therefore likely that the completed evaluation technique will comprise a mix of objective and subjective measures, which will combine to give an overall estimate of IVIS usability.

4 Objective

The objective of this project is to develop a method for evaluating the usability of IVIS interfaces. The evaluation method will be developed for use by automotive manufacturers in the concept design phase of IVIS development. This method will have an advantage over other, mainly subjective, IVIS evaluation techniques because it is based on quantitative metrics and therefore allows different interfaces to be directly compared. Evaluation at an early stage in the design process will also ensure that only those concepts which offer a real advantage over existing IVISs will be taken forward for further development, resulting in a time and cost saving to automotive manufacturers. It is hoped that the evaluation method will also facilitate the development of new HMI concepts for use in vehicles.

5 Methods

The evaluation method will be developed as part of an iterative process, in which the findings from various stages of the research will feed back into the project to aid the development of the method.

5.1 Modelling Primary and Secondary Driving Tasks

The primary driving task and a selection of sample secondary driving tasks will be modelled individually, and then combined into a single model of driving. The main element of the model will be the Critical Path Analysis technique, which will be used to predict task interaction times. This will be used specifically to model efficiency of use, which is one of the main facets of IVIS usability. Efficiency is useful as an indicator of the safety of the interaction because it signifies how much time a driver spends interacting with secondary tasks and how this could potentially conflict with primary task performance. The model will be incorporated into an evaluation method which will also include measures of wider aspects of usability, including learnability, flexibility and user satisfaction.

5.2 Validation of Evaluation Method

The evaluation method will be applied to various existing and potential IVIS interfaces to predict levels of usability for these systems. A series of user trials will also be conducted using a driving simulator and in on-road tests, using an instrumented vehicle. The trials will be used for the empirical evaluation of the usability of the same selection of IVIS interfaces. These results will be used to validate the predictions from the evaluation method. The method, including the task interaction time model, will then be refined according to these results.

6 Conclusions

The research element of this project will contribute to further understanding of what constitutes usability for In-Vehicle Information Systems and how these aspects of usability can be evaluated using a combination of objective and subjective techniques. This work will address the deficiencies of existing evaluation techniques by taking a more quantitative approach and also by focussing on very early-stage concept design.

Acknowledgements

This research is sponsored by Jaguar Cars Ltd. and the Engineering and Physical Sciences Research Council (EPSRC).

References

1. Shackel, B.: Ergonomics in design for usability. In: Harrison, M.D., Monk, A.F. (eds.) People and computers: designing for usability. 2nd Conference of the British Computer Society Human Computer Special Interest Group, pp. 44–64. University Press, Cambridge (1986)
2. Nielsen, J.: The usability engineering lifecycle. IEEE computer 25, 12–22 (1992)
3. Bhise, V., Dowd, J., Smid, E.: A comprehensive HMI evaluation process for automotive cockpit design. In: SAE World Congress. SAE International, Warrendale, Pennsylvania (2003)
4. Stevens, A., Board, A., Allen, P., Quimby, A.: A safety checklist for the assessment of in-vehicle information systems: a user's manual. Technical report, Transport Research Laboratory (1999)

Evaluating Human Computer Interaction through Self-rated Emotion*

Danielle Lottridge

Interactive Media Lab, Mechanical & Industrial Engineering, University of Toronto,
5 King's College Rd, Toronto ON M5S3G8, Canada
danielle.lottridge@utoronto.ca

Research Area. Affective self-report, Emotion, Evaluation methods.

Related Submissions. "Emotional Bandwidth: Information Theory Analysis of Affective Response Ratings Using a Continuous Slider" [12].

Research Topic. This thesis investigates the tools, psychometric characteristics, physiological relationships, and overall benefits and limits of continuous self-rated emotion as a measure of conscious affective experience.

1 Research Problem

Emotions are an important part of the user experience in human machine interaction [1]. User interface and product design requires satisfactory methods for evaluating and comparing experiences, and generally obtaining user feedback, including emotional feedback. In order to be widely used by designers, researchers, usability engineers and ergonomists, measures of users' emotional responses need to be standardized, and need to be relatively lightweight and easily usable.

Interviews and questionnaires are standard ways to assess emotion after system interaction. Some, such as the self-assessment manikin (SAM) [2], are used to rate websites. Some tools break reactions down into emotional categories for product feedback (e.g., [3]). However, summative emotions occurring after interaction has occurred are not suitable for evaluating changing interface aspects and interactions since they do not provide moment-to-moment ratings. Further, judgments may correspond to mode, average or end feelings [4]. In order to provide continuous feedback, some researchers have explored physiological measurements (e.g., [5]). These measures represent a fundamental component of emotion, and are often analyzed using self-report data, since physiological data are complex to interpret. Others [6] have explored qualitative methods to capture emotional responses in time. These investigations aim to deepen understanding of the human emotional experience, rather than provide a pragmatic way to evaluate emotional interactions.

Recent work has proposed sliders as a useful way to continuously measure self-reported emotion [7,8,9]. My thesis extends this work, to examine: what tools (sliders

* Advisor: Mark Chignell, chignell@mie.utoronto.ca

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 860–863, 2009.

and variations) can we design to facilitate affective self-report? What are the behavioural characteristics (psychometric properties) of how people use these tools? What are physiological correlates to the ratings? And what are relationships between self-reports to tasks and interaction performance? Throughout this thesis, we examine the benefits and limits of continuous affective ratings, which correspond to participants' impressions or perceptions of their emotional status.

1.1 Hypothesis/Claim

I believe that measures of conscious affective experience can give insight into mild, near neutral and generally "difficult to label" states, which may or may not correlate with physiological manifestations. I believe that different tools will facilitate different aspects of emotional self-reporting, and will hold unique psychometric properties. I posit that participants will display reliable individual differences in how they self-report their emotions. This work aims to produce rigorously validated emotion self-report tools with known psychometric properties for affective evaluation.

2 Methods

A series of experiments are proposed to progressively test different aspects of self-rated emotion. An initial step was to determine the discriminatory power of affective self-report. I collaborated in a study that used a single slider, and found a significant difference between two videoconference experiences [10]. We reanalyzed the data to uncover psychometric properties that differentiated four types of response patterns (submitted to Interact).

My first experiment, conducted in 2008, compared the three devices in a within-subjects study with 12 participants, to investigate the validity of the tools and the tradeoffs between the devices [11]. We presented 12 short videos. The ratings generally reflected the anticipated emotional content of the videos, with a certain amount of individual variation. Participants stated a preference for the touchscreen device. However, the 1 slider and 2 slider devices were found to be superior in terms of reliability measures. The results led to development of revised emotion rating tools and selection of more immersive videos to elicit stronger emotions in future studies.

The second experiment will be conducted in the spring 2009 with the goals to validate the iterated versions of the tools, discover the learning curves for emotional-self-reporting, individual differences and psychometric properties. The design is similar to the first experiment, except with longer videos and individual difference measures such as personality and emotional intelligence. The third experiment will be conducted in summer 2009. This third experiment will repeat the second experiment, but in Japan and with Japanese participants. The stimuli will be identical except that the instructions will be in Japanese and the videos will be narrated in Japanese. This experiment will give insight into the cultural differences that may exist in affective self-report.

The fourth and last experiment in this thesis will involve a health care-related task, such as browsing for chronic disease information online. The goal is to relate affective self-report to outcome measures such as browsing behaviour and retention of information. Participants will report their affective state throughout the experience. This experiment is planned for fall 2009. It is hoped that feedback from the Interact Doctoral Consortium will facilitate insights into tradeoffs for the design of this study.

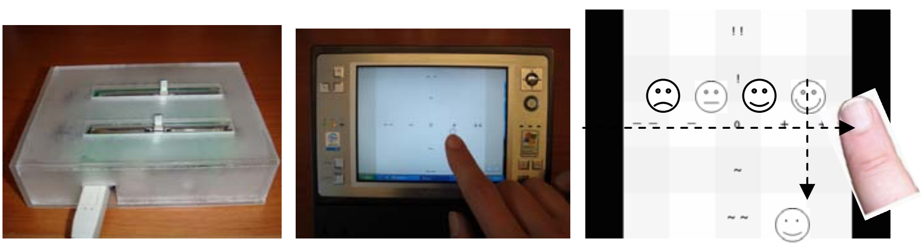
3 Proposed Solution

We designed three prototypes: a tangible one-slider model, representing valence; a two-slider model, representing valence and arousal; a two-dimensional touch-screen model, with an animated face that follows the participants' current location (Figure 1). In this case the expression on the face changed to match the position of the cursor within the two-dimensional space of valence and arousal; the curvature of the mouth changes with horizontal motion and the size (circular area) of the eyes change with vertical motion.

One outcome of this research is the discovery of psychometric properties to characterize use of sliders as scales for affective self-reports. Individuals can display varying affective capacities, which may change over time in response to the stimuli, environment and context. We used the Shannon-Weaver measure of informational entropy to quantify the rating usage bandwidth, which we refer to as Emotional Bandwidth [12].

As mentioned in the methods section, we aim to investigate the time taken to habituate to continuously self-report one's emotional state, and how individual and cultural differences affect ratings. Lastly, we plan to relate self-report to outcome measures such as behaviour, understanding and judgments of experience.

Contributions. This research is developing self-report tools for measuring emotional responses to user interfaces and interactions. The proposed thesis is original because it is the first to design and validate two dimensional self-report tools, discover their psychometric properties and test individual and cultural differences. The final experiment in this thesis will extend this research to a task scenario to investigate how ratings interact with behaviour and other outcomes. We aim for a rigorous, continuous method of measuring self-reported emotion, which is a valuable addition to traditional HCI measures such as response time, accuracy, usability, satisfaction, and workload.



(a) Two-slider Emotrace (b) Touchscreen Emotrace (c) Dynamic Facial Feedback

Fig. 1. The mouth changes with horizontal motion; the eyes change size with vertical motion

References

1. Norman, D.: Emotional Design: Why we Love (or Hate) Everyday Things. Basic Books, New York (2004)
2. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psych. 39, 1161–1178 (1980)
3. Desmet, P.: Measuring emotion: Development and application of an instrument to measure emotional responses to products. In: Blythe, M.A., Monk, A.F., Overbeeke, K., Wright, P.C. (eds.) Funology: From Usability to Enjoyment, pp. 111–123. Kluwer Academic Press, The Netherlands (2003)
4. Kahneman, D.: Objective happiness. In: Kahneman, D., Diener, E., Schwarz, N. (eds.) Well-being: The foundations of hedonic psychology, pp. 3–25. Russell Sage Foundation, New York (1999)
5. Mandryk, R.L., Inkpen, K.M.: Physiological Indicators for the Evaluation of Co-located Collaborative Play. In: Proceedings of Computer Supported Collaborative Work CSCW 2004, pp. 102–111. ACM Press, New York (2004)
6. Isbister, K., Höök, K., Sharp, M., Laaksolahti, J.: The sensual evaluation instrument: developing an affective evaluation tool. In: The Twenty-Fourth Annual SIGCHI Conference on Human Factors in Computing Systems CHI 2006, pp. 1163–1172. ACM Press, New York (2006)
7. Laurens, G. Desmet, P.: Designing a research tool: A case-study in how design and research can cross-fertilize each other, In: DRS 2008 Design Research Society Biennial Conference, pp. 362(1)–362(10). DRS digital archive (2008)
8. Baumgartner, H., Sujan, M., Padgett, D.: Patterns of Affective Reactions to Advertisements: The Integration of Moment-to-Moment Responses into Overall Judgments. J. of Market'g Research 34(2), 219–232 (1997)
9. Mauss, I.B., Levenson, R.W., McCarter, L., Wilhelm, F.H., Gross, J.J.: The tie that binds? Coherence among emotional experience, behavior, and autonomic physiology. Emotion 5, 175–190 (2005)
10. Ranjan, A., Birnholtz, J., Balakrishnan, R.: Improving Meeting Capture by Applying Television Production Principles with Audio and Motion Detection. In: The Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems CHI 2008, pp. 227–236. ACM Press, New York (2008)
11. Lottridge, D., Chignell, M.: Emotrace: Tracing Emotions through Human-Product Interaction. In: Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting HFES 2009. CD-ROM (2009)
12. Lottridge, D., Chignell, M.: Emotional Bandwidth: Information Theory Analysis of Affective Response Ratings Using a Continuous Slider. In: Proceedings of INTERACT 2009 (2009)

Human-Computer Interaction Techniques in Firefighting*

Sebastian Denef[1,2]

[1] Fraunhofer FIT, Germany
Schloss Birlinghoven, 53754 Sankt Augustin, Germany
sebastian.denef@fit.fraunhofer.de
[2] CIT Cork Institute of Technology, Ireland

Abstract. This thesis investigates the design of human computer interaction techniques for ubiquitous computing solutions in firefighting.

Research Area: Human-Computer Interaction Design for Safety Critical Ubiquitous Computing.

1 Research Problem

For HCI, research on interaction techniques is a core concern. A human-computer interaction technique provides users "a way to carry out an interactive task" [1]; or, as [2] put it, an interaction technique "represents an abstraction of some common class of interactive task"; it can be classified as a primitive element of an interactive system. Lacking the same means of input and output, ubiquitous computing systems do not support desktop computing interaction techniques; they demand for novel interaction techniques tailored for the specific use [3].

Tailored interaction techniques are crucial for ubiquitous computing systems supporting people who work in hazardous, safety-critical environments. Interaction techniques need to closely match on-site situation as the mixture of cognitive and physical load, protective clothing and special equipment leaves very little space for human-computer interaction.

For the domain of firefighting, such restrictions are imminent. Fighting fire is a multi-faceted task in a hazardous environment that changes dynamically during an intervention. The work on the first line of intervention is complex and requires a set of very specific skills and tools [4]. It is often conducted in buildings full of smoke, which can only be accessed by crawling on the floor due to heat distribution. To cope with an ever-changing environment, firefighters rely on intensive training, high levels of improvisation and teamwork [5].

Current HCI research targeting the firefighting domain, proposes systems to support navigation, health monitoring or communication [6-10]. While some researchers avoid the constraints of the first line of intervention by focusing on the work that takes place in relatively safe and less restrictive environments [11], research projects

* Supervisors: Prof. Dr. Arnd Steinmetz[2], Prof. Dr. Reinhard Oppermann[1].

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 864–867, 2009.

targeting the first line of intervention often use firefighting scenarios as a motivation or explanation for the need of a certain technological development [12, 13]. Here, the approach is to adapt interaction techniques designed for everyday mobile technology [12] or to rely on relatively novel means of interaction such as head-mounted displays [13] without evaluating their fit with real-world conditions. The work that specifically focuses on interaction techniques for firefighting [14] focuses on certain aspects of firefighting practice but does not consider the overall firefighting practice such as the collaborative nature of the job.

Firefighting practice is, however, very sensitive to technological change and likely to be disturbed by systems that do not take into account the full complexity of the task [5]. As of today, firefighting on the first line of intervention remains an activity with almost no IT support [4], lacking human computer interaction techniques that are able to cope with the challenging setting.

Recent HCI research shows that a situated perspective provides means to study social, situated actions in dynamic use contexts and that interaction design needs to respect the underlying role of embodiment in interaction [15]. While [15] provides a research paradigm covering the relevant aspects, the existing methods have been mainly applied in far less restricted environments [16] leaving unanswered the question if and how such open methods can be applied for the design of interaction techniques to be used in highly dynamic, safety critical environments.

2 Working Hypothesis

Design methods of the situated perspectives HCI research paradigm provide meaningful contributions to the challenge of designing interaction techniques for ubiquitous computing systems that support people working in highly dynamic, safety critical conditions.

3 Methodology

This research applies an action research method [17] and an phenomenologically situated research paradigm [15].

The *initial diagnosis* motivating this research was conducted in form of a study that introduced a triggering artifact [18] in firefighting practice by means of a wizard-of-oz prototype of an indoor navigation system. A full account of the findings can be found in [5].

The *actions* to be taken were planned using a broad set of methods following the situated perspectives research paradigm. They comprise empathic studies [19], workplace studies, design probing [20], experience prototyping [21] and qualitative in-field evaluations of prototypes:

- To build empathy, studies to allow forming an embodied understanding of this work are set up, as firefighting is essentially different from daily experiences. The work-studies take place in professional firefighting training apartment buildings that simulate typical interventions and in training facilities that allow experiencing real fire conditions. First results of this empathy building approach can be found in [5] and [22].

- Workplace studies of the equipment used by firefighters analyze tool support in firefighting and extract the qualities that make the existing tools fit the challenging environment. Empirical material for this study has been conducted and initially analyzed as presented in [4].
- Design probes [20] of consumer interactive systems featuring novel interaction techniques used by firefighters in their private life help to join the expertise of firefighters with their usage experience of novel interaction techniques. For this activity, jackets with interactive textiles have already been handed out to firefighters.
- Experience prototypes [21] of novel interaction techniques allow to gain results early in the design process. An initial experience prototype for visual feedback inside the SCBA mask has been constructed and evaluated with firefighters.
- Qualitative in-field evaluations of prototypes will analyze critical incidents and feedback by firefighters to learn about the fit of interaction techniques within firefighting work practices.

To *specify learning* the results from individual studies will be reflected in terms of their contribution and value for the design of interaction techniques.

4 Proposed Solution

A broad set of methods, inspired by ethnographic and design research of the situated perspectives HCI research paradigm, provides the means to design interaction techniques for ubiquitous computing solutions to be used in hazardous environments. This methodological foundation supports the design of interaction techniques for ubiquitous computing solutions in firefighting scenarios.

5 Expected Contributions

Scientific contributions to the HCI research community are expected to be twofold:

- On a theoretical level, the research will provide an understanding if and how design research methods following the situated perspective research paradigm can be applied and for the design of interaction techniques for ubiquitous computing solutions to be used in highly dynamic, safety-critical environments.
- On a practical level, a set of interaction techniques will be developed and evaluated to support firefighters working inside burning buildings.

References

1. Raisamo, R.: Multimodal Human-Computer Interaction: a constructive and empirical study University of Tampere (1999)
2. Jacob, R.: Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces. In: Advances in Human-Computer Interaction, pp. 151–190. Ablex Publishing Co. (1993)
3. Jacob, R., Girouard, A., Hirshfield, L.M., Horn, M., Shaer, O., Solovey, E.T., Zigelbaum, J.: CHI2006: what is the next generation of human-computer interaction? Interactions 14, 53–58 (2007)

4. Denef, S., Ramirez, L., Dyrks, T.: Letting Tools Talk: Interactive Technology for Firefighting. In: Extended Abstracts CHI 2009, pp. 444–447. ACM, Boston (2009)
5. Denef, S., Ramirez, L., Dyrks, T., Stevens, G.: Handy navigation in ever-changings Spaces: an ethnographic study of firefighting practices. In: DIS 2008, pp. 184–192. ACM, New York (2008)
6. WearIT@Work Project (2008), http://www.wearitatwork.com
7. Ubiquitous IP Centric Government & Enterprise Next Generation Networks (2008), http://www.u-2010.eu (cited, 2008 12/12)
8. landmarke project, http://landmarke-projekt.de (cited, 2008 12/12)
9. Tri-Sentinel (2008), http://www.trisentinel.com/solutions.htm (cited 2008 12/12)
10. FeuerWhere (2008), http://www.feuerwhere.de/ (cited 2008 12/12)
11. Landgren, J., Nulden, U.: A study of emergency response work: patterns of mobile phone interaction. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 1323–1332. ACM, San Jose (2007)
12. Jiang, X., Chen, N.Y., Hong, J.I., Wang, K., Takayama, L., Landay, J.A.: Siren: Context-aware computing for firefighting. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 87–105. Springer, Heidelberg (2004)
13. Wilson, J., Bhargava, V., Redfern, A., Wright, P.: A Wireless Sensor Network and Incident Command Interface for Urban Firefighting. In: 4th Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services, 2007. MobiQuitous 2007, pp. 1–7 (2007)
14. Naghsh, A.M., Roast, C.R.: Designing user interaction with robots swarms in emergency settings. In: Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges, pp. 519–522. ACM, Lund (2008)
15. Harrison, S., Tatar, D., Sengers, P.: The Three Paradigms of HCI. alt.chi (2007)
16. Gaver, W., Boucher, A., Law, A., Pennington, S., Bowers, J., Beaver, J., Humble, J., Kerridge, T., Villar, N., Wilkie, A.: Threshold devices: looking out from the home. In: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems. ACM, Florence (2008)
17. Baskerville, R.: Investigating information systems with action research. Commun. AIS 2 (1999)
18. Mogensen, P., Robinson, M.: Triggering Artefacts. AI & Society 9, 373–388 (1995)
19. Leonard, D., Rayport, F.: Spark Innovation Through Empathic Design. Harvard Business Review, 102–113 (1997)
20. Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B.B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., Eiderböck, B.: Technology probes: inspiring design for and with families. In: CHI 2003, pp. 17–24. ACM, Ft. Lauderdale (2003)
21. Buchenau, M., Suri, J.F.: Experience prototyping. In: DIS 2000, pp. 424–433. ACM, New York City (2000)
22. Dyrks, T., Ramirez, L., Denef, S.: Designing for Firefighters—Building Empathy through Live Action Role-Playing. In: Abstracts of the International Conference on Information Systems for Crisis Response and Management., Gothenburg, Sweden, May 10–13 (2009)

Retrieval of User Interface Templates Based on Tasks*

Jordan Janeiro[1]

[1] Technische Universität Dresden, Department of Computer Science, Institute for Systems Architecture, Nöthnitzerstr. 46, 01187 Dresden, Germany
{jordan.janeiro,thomas.springer,alexander.schill}@tu-dresden.de
[2] PUC-Rio, Computer Science Department, Rua Marquês de São Vicente, 225, Gávea, Rio de Janeiro, RJ, Brazil
simone@inf.puc-rio.br

Abstract. Using template structures is an interesting approach to develop user interfaces. A designer can predefine such structures, following best user interface practices, to be constantly reused. However, it may be a problem to navigate through a repository of templates to find a suitable one for a certain application. Therefore, we propose a retrieval mechanism for templates based on its supported tasks.

Keywords: user interface; template; information retrieval; natural language; task terminology; search engine.

1 Introduction

The user interfaces subject is still a complex open problem in computer science. According statistics, 50% of the source code and development time of a system are due to the user interface design [1]. Therefore, there is naturally the requirement to automate such process. Indeed, there are approaches which aim to automatically generate user interfaces, analyzing the logics of a system. However, such automatically approaches generally present usability problems [2].

A suitable approach, as an alternative for full manual or automatic user interface creation, is the use of *templates* [3]. The idea of using templates is that a designer can previously define a generic user interface for certain types of system functionalities and publish them, in the sense that designers or even non-designers can reuse the templates to easily create end user interfaces.

Generally, in software companies, such repositories may contain a large number of templates and thus, it becomes difficult for developers to find a template for a specific purpose. Therefore, the goal of this work is to support developers on finding templates based on aspects as, the user tasks the template supports and its structural description. Such aspects are specified by a natural language query and have to be matched with a formal description attached to the templates. Our approach should identify the most relevant templates which cope with the semantics of such query.

* Advisors: Thomas Springer[1], Simone DJ Barbosa[2], Alexander Schill[1].

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 868–871, 2009.

The details of our approach are illustrated by a running example; a personal web page scenario, which manages a photo album. For this scenario, we defined a template, as presented by Figure 1 (b).

2 Template Search Mechanism

Our approach proposes an engine which searches for templates, based on its semantic information. Figure 1 (a) presents an overview of the steps involved in the search mechanism. First, the user describes the template he searches, by specifying its semantic information in a query; the engine receives the query and analyses it to find the matching templates and, finally, the engine presents the templates, ranking them by relevance. The following subsections provide more details about the search mechanism just described.

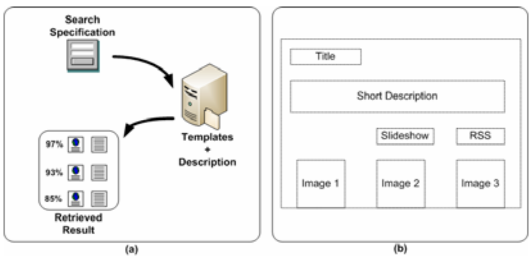


Fig. 1. (a) Methodology for searching templates based on its description. (b) template for a personal photography web page.

2.1 Template Description Model

This work proposes a standard definition for a model which describes user interface templates, called here *template description model*. By template description we mean semantic information such as: the possible tasks a template supports, from the end user point of view; the modality in which a task is executed; the elements which compose a certain template and the device which the template targets. Such information is then attached to the templates to allow the search engine to perform semantic searches.

Our model is based on previous approaches [4][5] which aim to describe taxonomies for user interface tasks, which can be performed by end users. Based on the model we propose here, we identify that the template of our use case might perform the following tasks using the notation proposed here [5], for example: *UseInformation(slideshowHyperlink)*, *UseInformation(viewThumbnail)* and *ProvideInformation(subscribeRSSFeed)*. With such tasks, we can describe, for example, that our use case template, allows the user to follow the slideshow hyperlink, to view the slideshow of the photos which are currently in the web page; or simply to view thumbnail pictures; or to allow the user to subscribe to news (RSS feeds) associated to the web page.

2.2 Query Formulation

Describing the semantics of a template in any kind of query language may imply the addition of unnecessary complexity for the template adopter, such as learning a new language and obtaining enough experience with it to describe precisely the goal of a query. Therefore, we chose to use a natural language, in this case English, as the query language for the search mechanism. For example, for our use case example we can formulate a query like, "*a user interface which presents a photo album as thumbnails. The user can: click on the thumbnails and view full images, the user can view the photo album as a slideshow and the user can subscribe to the RSS feeds service*". We assume that the template adopters intuitively use the task terms in the query which can be further extracted and mapped to the tasks compiled by our template description model.

2.3 Search Matching

If in one hand the use of English as a natural language provides flexibility on specifying a query for templates, in the other hand it introduces the complexity of identifying its goals. As we allow the user to formulate a query freely using the English language, our major challenge is to analyze such a query to identify which terms match the terminology, specified by our model. For this purpose, our approach uses the unstructured information management (UIM) framework[1], a system which performs the recognition of named terms (structured data) from a text (unstructured data) by information extraction techniques. The framework achieves its goal by integrating a structured data source, which contains a terminology, and then identifying these terms in an unstructured data (text). In our case, the structured data refers to the template description model terminology and the unstructured data refers to the query that the user formulates in natural language.

As the user formulates a query using natural language, we must consider that the query may not specify exactly the terminology foreseen in our model, due to the possibility to formulate it with synonyms of our terminology, or even formulate the purpose of a query in a different way. To support the solution of this problem, we intend to use existing terminology databases, such as *WordNet*[2] or *SAPTerm*[3], which aim to describe concepts and its relationship. We can use such database in combination with the UIM framework to identify the goal of a query formulated with other terms, which may be synonyms.

3 Related Work

Many approaches in literature present proposals to define task taxonomies, based on the interaction between end users and systems [4][5][6]. Although these studies describe a set of tasks a certain user interface supports, they do not propose any practical kind of integration of taxonomies that formally describe user interfaces.

[1] Unstructured Information Management Architecture, http://uima-framework.sourceforge.net
[2] WordNet, http://wordnet.princeton.edu/
[3] SAPTerm, http://help.sap.com/saphelp_nw04/helpdata/en

With similar concepts presented in this work, the Web Semantic concept [7] proposes the attachment of additional information (semantic) to web services to allow its discovery, in a further step. There are some existing semantic service matchmakers which perform such discovery processes based on the attached semantic, by non-logic, logic or hybrid matching techniques [8]. In our approach, the semantic we attach to the templates represent tasks and our search engine represents the semantic service matchmakers, which searches the templates based on the attached tasks.

4 Conclusion

This paper presents an approach which retrieves templates, annotated with the template description model, through natural language queries. For achieving this goal, we define a model, mainly based on task terminologies, which aims to describe the tasks a template supports, by the end user point of view. We annotate the templates using such a model to perform searches of templates based on tasks. We use natural language as a query language, to allow the user to easily formulate queries to find the templates. To support such an approach, we adopt the UIM framework for processing the queries and identify the tasks defined within it. In combination with the UIM framework, we use a terminology database to recognize the synonyms of the tasks.

References

1. Myers, B., Rosson, M.: Survey on User Interface Programming. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 195–202. ACM Press, New York (1992)
2. Spillner, J., Feldmann, M., Braun, I., Springer, T., Schill, A.: Ad-Hoc Usage of Web Services with Dynvoker. In: Mähönen, P., Pohl, K., Priol, T. (eds.) ServiceWave 2008. LNCS, vol. 5377, pp. 208–219. Springer, Heidelberg (2008)
3. Nielsen, J.: User Interface Directions for the Web. In: Communications of the ACM, pp. 65–72. ACM Press, New York (1999)
4. Zhou, M., Feiner, S.: Visual Task Characterization for Automated Visual Discourse Synthesis. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 392–399. ACM Press, New York (1998)
5. Byrne, M., John, B., Wehrle, N., CrowThe, D.: Tangled Web We Wove: a Taskonomy of WWW Use. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 544–551. ACM Press, New York (1999)
6. Shneiderman, B.: The Eyes Have It: a Task by Data Type Taxonomy for Information Visualizations. In: IEEE Symposium on Visual Languages, pp. 336–343. IEEE Press, New York (1996)
7. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web, Scientific American, vol. 284, pp. 34–43 (2001)
8. Klusch, M.: Semantic Web Service Coordination. In: CASCOM: Intelligent Service Coordination in the Semantic Web, pp. 59–104. Springer, Heidelberg (2008)

Supporting Aphasics for Capturing, Organizing and Sharing Personal Experiences*

Abdullah Al Mahmud

Department of Industrial Design, Eindhoven University of Technology
P.O.Box 513, 5600 MB Eindhoven, The Netherlands
{a.al-mahmud,j.b.o.s.martens}@tue.nl

Abstract. When a person, due to brain injury or another disease, suffers in his or her ability to speak, it becomes inherently cumbersome to share needs, emotions, and experiences through personal stories and social interaction. This paper describes the aim and progress of the author's dissertation, which focuses on designing a support system to share daily experiences for people suffering from expressive aphasia.

Keywords: Aphasia, Sharing experiences, Contextual interview.

1 Introduction

Much of our social life consists of sharing daily stories with other people. However, sharing personal stories can be extremely difficult for people with limited verbal ability, such as those suffering from expressive aphasia. As a result, aphasia often leads to increased social isolation and possibly to depression. Enabling aphasics to share their daily experiences will help them to become more socially active and to re-engage with their preferred life style. Aphasia is an acquired communication disorder that is caused by brain injury or trauma. Aphasia affects language comprehension and generation [4], such that people's ability to express themselves verbally suffers. Augmentative and Alternative Communication (AAC) devices such as TouchSpeak [7] are widely used in aphasia therapy as well as during the post-therapy period but there are various limitations of AAC. They contain isolated symbols whose meanings have to be learned and aphasics need to retrieve phrases or words to construct meaning in order to support very simple stories. Therefore, the way these activities are designed end up posing a cognitive challenge for the users. Moreover, AAC devices support needs-based interaction for functional communication. They lack the support for sharing experiences based on real life events, which are crucial aspects to social interaction.

Photographs have several functions such as preserving memories about the past, narrating and organizing everyday experiences, and communicating and building social relations [3]. Photography has been proposed within 'Aphasia Talks' [6] as a way of facilitating self-expression in aphasics for reintegration, improving socialization and allowing recreation. The usefulness of digital photographs within

* Advisor: Jean-Bernard Martens.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 872–875, 2009.

conventional storytelling has been reported (see e.g iTell [5]). The challenge or issue of how people with special needs can be empowered to use photos to share experience, does not yet seem to have been explored in depth. A study has been reported to help aphasics in building story using photos [2]. However, there were certain limitations of the prototype such as the editing of the story had to be done at the time of capture and there was no support for organizing the captured photos. Moreover, the story-building task was limited just by adding a simple emoticon and sound and lacking the order and comprehensibility of the captured materials. Therefore, our objective is to design a supporting system for capturing and sharing personal experiences via photos considering the limitations described above.

The main research question of this PhD thesis is to investigate whether and how capturing everyday activities that assist in storytelling can help to reduce the social isolation and passive position of aphasics. We specifically focus on aphasics who have been discharged from the rehabilitation center and who need to continue their life with a serious disability. The main reason for this is that in earlier phases of diagnosis and treatment, therapists and patients tend to concentrate on training whatever speech abilities that remain.

2 Methodology and Conducted Research

Since aphasics cannot express their thoughts verbally, it is not easy to involve them in the design process. Existing research focused on different methods such as participatory design by involving domain experts in a user centered manner etc [1]. We involved speech therapists in the design process since they are able to judge the capabilities of aphasics. To know more about how aphasics and how their partners are involved in the communication process we visited one aphasic and his partner. Moreover, we also explored whether and how photos can be used to tell past experience. The main findings were a. photos can be used for aphasics to recall past experience b. Personally relevant photos are needed to narrate an experience c. providing personally relevant cues are helpful such as 'who', 'what', to make the process of retelling easier.

We realized that it was not easy to engage aphasics and elicit design requirements. Therefore, we have adopted a different methodology. We conducted an experiment with two non-aphasic persons and they were asked to explain how they would capture and share their personal experiences in an unfamiliar language. This would mirror the task that aphasics would need to carry out if they would narrate their story using photos. We argue that involving non-aphasics in the early design phase will give us better insight to understand the problem of expressive aphasics. Participant was told to take photos of events and express that photo event(s) in an unfamiliar language. The main objective was to find out how people construct and share narratives from photos to others given that they did know a particular language, similar to the condition of aphasics. The main findings are a. not many pictures are needed for communicating key everyday experience b. people miss joining words, phrases to establish a connection between photos while telling in other language c. people prefer to narrate their experience by making clusters based on activity and time. The findings are used to enrich the first conceptual prototype, which is described below.

3 Proposed Solution

The design we propose will help aphasics to capture everyday experiences and later to share it with significant others in a face-to-face setting. There are two aspects of the proposed design. First is the capturing of everyday experiences. It became apparent that a traditional camera is not suitable for aphasics to capture daily experiences [6]. Therefore, we propose to build a suitable capturing device such as using an off the shelf sunglass camera to capture picture, sound and video (see Fig 1). The captured device will be operated by a remote controller and that can easily be handled by only one hand. Later, these captured materials can easily be transferred to a laptop or PC for reviewing with the help of our appropriate sharing software application. A suitable docking station will be designed for transferring the captured materials easily.

Computing technology can provide support to construct a particular experience from the captured materials. The support can be in the form of help in organizing and editing the captured materials and guiding user to recall in constructing experiences. As a first step in designing the first prototype, we have created an application that could help aphasics to organize and share personal experiences (see Fig 1). The application facilitates aphasics to organize their collected photos according to time and date. Moreover, photos can be organized according to different clusters such as 'who', 'where' 'what' and a story can be defined in terms of these principles. To help in building a narrative to share experience a template is provided. The template has the option to fill appropriate pictures according to the attributes of who, where and what. There is an option for adding personal media such as audio, text, inbuilt icons in the picture to augment its meaning. Being able to respond to questions is obviously, what

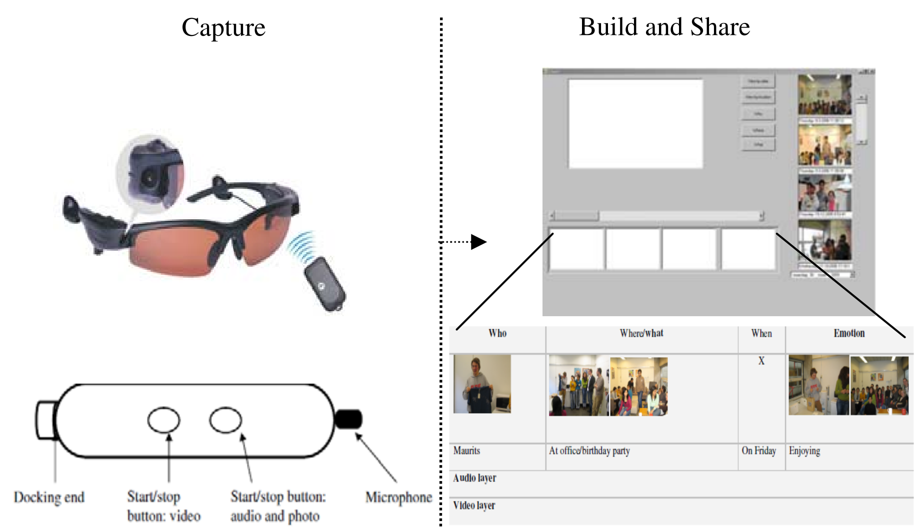


Fig. 1. The capturing device with external controller (left), the application for building and expressing narrative of experiences (right)

makes the difference between simply telling a story and using storytelling as a means for social interaction. This aspect therefore needs to be carefully developed and tested in interaction with the end user.

4 Research Plan

First, we will evaluate the suitability of the capturing device. The tool will be evaluated in a longitudinal study to investigate how aphasics use the device in their daily life. Moreover, we want to compare how aphasics and non-aphasics use the device and can communicate by using it effectively. We will incorporate the findings while evaluating it iteratively in to the final design. The next step is to evaluate the application for constructing and sharing daily experiences. It is equally important to evaluate how the narrative template that we proposed helps aphasics to reconstruct their experiences. Moreover, several design cycles with proxies are planned, also because we want to explore how the use of an experience sharing application evolves over time as users become more proficient with it.

Expected Contributions. Prior research has not explored how capturing and sharing could improve social interaction and confirming social affiliations for aphasics. We expect to show experimentally how the passive position of aphasics can be alleviated by our proposed solution. The method that we deploy will also contribute for designing assistive technologies for aphasics. The research and design outlined in this paper, although aimed at people with (expressive) aphasia, may be relevant for other user groups, such as people suffering from dementia or Alzheimer disease.

Acknowledgement. I would like to thank my supervisor Prof. Jean-Bernard Martens for his continuous support. Thanks also go to speech therapists, the person with aphasia and his partner for their cooperation.

References

1. Allen, M., McGrenere, J., Purves, B.: The Field Evaluation of a Mobile Digital Image Communication Application Designed for People with Aphasia. ACM Trans. Access. 1(1), 1–26 (2008)
2. Daemen, E., Dadlani, P., Du, J., Li, Y., Erik-Paker, P., Martens, J.B., de Ruyter, B.: Designing a free style, indirect, and interactive storytelling application for people with aphasia. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4662, pp. 221–234. Springer, Heidelberg (2007)
3. Harrison, B.: Photographic visions and narrative inquiry. J. Narrative Inquiry 12(1), 87–111 (2002)
4. Hillis, A.E.: Aphasia: Progress in the last quarter of a century. Neurology 69, 200–213 (2007)
5. Landry, B.M., Guzdial, M.: iTell: supporting retrospective storytelling with digital photos. In: Proc. DIS 2006, pp. 160–168 (2006)
6. Levin, T., Scott, B., Borders, B., Hart, K., Lee, J., Decanini, A.: Aphasia Talks: photography as a means of communication, self-expression, and empowerment in persons with aphasia. J. Topics in stroke rehabilitation 14(1), 72–84 (2007)
7. TouchSpeak, http://www.touchspeak.co.uk/

The Role of Personal and Shared Displays in Scripted Collaborative Learning

Sara Streng

Media Informatics, University of Munich, University of Munich,
Amalienstr. 17, 80333 Munich, Germany
sara.streng@ifi.lmu.de

Abstract. Applying technologies in collaborative learning is an important but difficult task. In this work I will investigate ways of supporting "scripted" collaborative learning with technologies. "Scripted" means that the learning phases, activities and tasks are predefined by educationalists to provide some form of guidance. In particular, I focus on the role of personal and shared displays in different learning phases. Furthermore, I examine how user interfaces can be designed in a way that guides the learner's activities. The proposed contribution is a set of guidelines as well as a software framework that support the creation of scripted collaborative learning applications.

Keywords: Collaborative Learning, CSCL Scripts, Shareable User Interfaces.

1 Introduction

In many learning scenarios, educators expect a benefit from using computer applications for specific purposes. For example, computers are more efficient in processing or visualizing large chunks of information, they facilitate efficient search and history functions as well as simultaneously editable copies, and much more. There are even learning applications that would not be possible without the support of computers, e.g. ArgueGraph [3].

Despite the advantages of computers, their application in learning environments remains a difficult task to solve. Often usability problems are so profound, that many people still prefer pen and paper. In collaborative learning additional problems arise because traditional desktop computers and laptops are not practicable to be used by multiple people at the same time. Although a number of interesting technologies such as interactive surfaces are available nowadays, they are rarely applied in educational fields.

This research aims at analyzing different technological setups and their applicability in various situations in the area of collaborative learning. This includes the choice of technological devices, the seating arrangement, the design of the user interface as well as the way learners are guided. In the long run, the goal is developing computer applications that are as easy to use as the according pen-and-paper scenario, while simultaneously providing the benefits mentioned above.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 876–879, 2009.

2 Problem Statement and Research Questions

Educationalists use collaboration scripts to specify how learners should be guided. This guidance is considered necessary because students rarely engage in effective learning activities spontaneously [5]. A script prescribes a sequence of activities to be carried out by the learners. For example, after reading a text paragraph, one learner summarizes its content. The other person listens and subsequently gives feedback. In the end both learners elaborate on the results and switch roles for the next paragraph [2]. The first research question related to this is: Which guidance functions can be provided by technology support in collaborative learning? My fundamental hypothesis is that well-designed computer programs can be used to guide certain aspects of the learners' activities. For example, in argumentative learning students can be prompted to build argumentation sequences or justify their arguments by designing the user interface in a way that suggests the according activities.

A second important question regards the type of implementation to choose. Many collaboration scripts contain alternating individual and collaborative working phases. Accordingly, technology support for scripted collaboration can be designed in a way such that the interfaces, in particular the displays, correspond to activity types: Individual displays (e.g. laptops, PDAs, tabletPCs) may be used for individual phases, shared displays (interactive surfaces, wall displays) for collaborative phases. To what extent does the usage of display types affect variables connected to learning, such as motivation, the subjective effort for the learners, and of course the learning outcome? In more detail, the following sub-questions will be investigated:

- Which devices and user interfaces are suitable for specific phases of collaborative learning scenarios? Is there a positive effect of the phase-specific choice of technology on the learning outcome? Is there a negative effect caused by the integration of several heterogeneous devices on usability, and how can this effect be counter-balanced?
- How can the integration of personal and shared workspaces be optimized technologically, e.g. by smooth handover or by defining personal workspaces as subspaces of a large shared space (e.g. an interactive table)?
- How does the seating arrangement affect technology-supported collaborative learning? E.g. under which conditions is a loose gathering around a large wall display preferable over a tabletop seating arrangement?

3 Related Work

Kobbe et al. [5] proposed a specification of collaboration scripts comprising the educational as well as the computer science perspective. It is the most elaborated approach in this direction and the foundation for any standardization or framework for collaboration scripts.

Specific scripts, which have either been adopted by or at least heavily influenced current projects are ArgueGraph [3], MURDER script [2], Social Script [8], Structured Academic Controversy [4] or Peer Reviewing [7].

4 Approach and Methodology

Coming from an HCI background it is important for me to get familiar with the educational perspective on collaborative learning. As a foundation I chose the field of collaboration scripts for several reasons. First, as the research field is several decades old, there is a collection of scripts that have proven to be effective. These scripts provide a detailed specification of collaborative learning scenarios including important educational aspects that need to be kept in mind when thinking about suitable user interfaces. Second, the specification proposed by Kobbe et al. [5] structures the domain, which allows for an analysis on an abstract level as well.

The chosen methodology is two-fold. On the one hand, hands-on experience will be gathered. By experimenting with personal and shared displays as well as different guidance strategies in various scenarios, the research questions will be explored. The scenarios are derived from effective collaboration scripts that could be improved by using technologies. The ideas on how technology could be applied are sketched and discussed with educational psychologists. During this process 2x2 factorial experiment designs are built that cover research questions from both HCI as well as educational psychology. The concepts are then implemented and evaluated in small interdisciplinary projects. In the end, guidelines will be created from the gained experience.

In parallel a software framework will be developed, which supports the implementation of scripted collaborative learning applications. The framework will incorporate structural and technical aspects. The structural aspects include components and mechanisms as proposed by Kobbe et al. [5], as well as strategies for learner guidance. The technical aspects concerns the interaction between heterogeneous devices, e.g. application handovers from personal to shared displays or vice versa. Regarding this issue there is already existing work [1] that can be built on. The framework and the mini-projects are built in a coherent, iterative process. The framework supports the development of the applications, which are built in the mini-projects. In reverse, insights from the practical experience will influence the development of the framework.

5 Research Conducted and Preliminary Results

Two projects have already produced preliminary results. The first one is a program that supports two learners in training text comprehension based on the *MURDER* script [2]. Both learners are provided with laptops. In this case a shareable user interface was not chosen because the two learners have oppositional roles (summarizer and listener). Two user studies have been conducted in a seminar to evaluate the usability as well as effects on variables connected to learning. The results show that subjects who used the program were equally motivated and showed similar effort as the control group who used pen and paper. In a post questionnaire about 75% of the participants stated that they would prefer the computer application over the pen and paper version.

In the second project a tabletop application for argumentative learning was developed. The *ArgueTable* guides the learners in substantiating their arguments and building argumentation sequences. A preliminary case study suggests that the quality of argumentation can be improved by means of this guidance [6].

For the realization of these projects, a first version of a framework has already been built. However, there are still many complex issues that need to be considered in a more elaborated version.

6 Summary and Future Steps

Collaboration scripts offer a great foundation to study the effects of technology in group learning sessions. Taking advantage of these possibilities, I am empirically exploring the impact of using different technologies in group learning sessions. In particular, two research questions are examined: The role of personal and shared displays as well as learner guidance through careful user interface design.

According to my approach there are two goals I want to pursue. First, practical experience will be gained by experimenting with different technological setups in different scenarios, which will in the end be transformed into guidelines. The second goal is developing a framework that supports the design and implementation of scripted collaborative learning applications.

References

1. Boring, S.: Interacting in multi-display environments. In: Proceedings of the Doctoral Colloqium at the Ubicomp 2007 (2007)
2. Dansereau, D.F., Collins, K., McDonald, B., Holley, C.D., Garland, J., Diekhoff, G.: Development and evaluation of a learning strategy training program. Journal of Educational Psychology 71(1), 64–73 (1979)
3. Jermann, P., Dillenbourg, P.: Elaborating new arguments through a CSCL script. In: Andriessen, J., Baker, M., Suthers, D. (eds.) Arguing to Learn: Confronting Cognitions in Computer-Supported Collaborative Learning environments, pp. 205–226. Kluwer Academic Publishers, Amsterdam (2003)
4. Johnson, D.W., Johnson, R.T.: Constructive conflict in the schools. Journal of Social Issues 50(1), 117–137 (1994)
5. Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hmlinen, R., Hkkinen, P., Fischer, F.: Specifying computer-supported collaboration scripts. International Journal of Computer-Supported Collaborative Learning 2(2-3), 211–224 (2007)
6. Stegmann, K., Streng, S., Halbinger, M., Koch, J., Fischer, F., Hussmann, H.: eXtremely Simple Scripting (XSS): A framework to speed up the development of computer-supported collaboration scripts. In: CSCL2009 Community Events Proceedings. ISLS (2009)
7. Trahasch, S.: From peer assessment towards collaborative learning. In: Frontiers of Education, FIE 2004, 34th Annual, October 2004, vol. (2) (2004)
8. Weinberger, A., Fischer, F., Stegmann, K.: Computer-supported collaborative learning in higher education: scripts for argumentative knowledge construction in distributed groups. In: Proceedings of the CSCL 2005, pp. 717–726. International Society of the Learning Sciences (2005)

Towards a Flexible User Simulation for Evaluating Spoken Dialogue Systems*

Dmitry Butenkov

Quality & Usability Lab, Deutsche Telekom Laboratories, TU Berlin,
Ernst-Reuter-Platz 7, 10587 Berlin, Germany
{dmitry.butenkov,sebastian.moeller}@telekom.de

Abstract. The main aim of research is to introduce a new data-driven user simulation approach for the quality and usability evaluation for spoken dialogue systems.

Keywords: Human-Computer Interaction, Human Factors, Statistical User Simulation, Usability Evaluation, Spoken Dialogue Systems, Experimental Design.

1 Problem Definition

Measuring the quality and usability of a spoken dialogue system (SDSs) is a complex continuous process requiring a series of expensive experiments to be established. In controlled laboratory experiments, participants very often tend to demonstrate unnatural communication behavior (e.g. no hang offs, barge-ins, or too much focused on the final goal). Also, a lot of test cases and subject groups have to be covered to provide meaningful evaluation results [1]. Thus, the costs of thorough usability tests in the design phase of spoken dialogue systems are often unreasonably high.

A potential solution is the evaluation of systems in their early development phase. The authors of [2] and [3] have shown that even at initial stages of SDS development usability problems can be discovered. The earlier problems can be identified, the cheaper their correction is. The authors of [2] and [4] have recently reported that semi-automated evaluation and testing can significantly decrease the expenses for quality and usability experiments. Following this idea, a simulation of user-system interactions may help to identify usability problems early and at low costs, and it might even be possible to base a tentative evaluation of the entire systems on simulated (instead of real user) interactions. In some cases time costs can be also be reduced significantly in comparison to a classical subjective evaluation approach. Of course, this solution is a deal between complexity and unification, but it must be able to recover complex enough usability problems.

Some attempts were made so far: the authors of [5] provided interesting, but very limited personality generation service, while authors of [3] and [4] reported first attempts to develop flexible user simulation systems with a focus on usability

* Advisor: Sebastian Möller.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 880–883, 2009.

evaluation. Despite of these prospective results, user simulation is still used not very frequently, except as an assessment tool to enable various machine learning algorithms by enlarging the original learning corpus [6]. Moreover, the authors of [2] and [7] have shown that all modern user simulators in the field of SDSs are either user- or system-centric. The idea of the present research is to develop a balanced concept for a data-driven simulation of user behaviour towards SDSs, covering several typical user patterns and scalable up to several common systems.

2 Research Hypotheses

A major part of commercial German SDSs has been evaluated in the Voice Awards competition being held since 2003 [8]. Although, there is no detailed information on every particular system available, an analysis indicates that the respective systems are relatively simple in comparison to current research systems, for instance from the smart home control field. The logical consequence is to try to reduce the simulation problem to standard slot-filling mechanisms. Thus, the first research hypothesis is that typical commercial German SDS can be reduced to a standard slot-filling problem.

A second research hypothesis follows from the previous one: since the problem scope is significantly reduced, the generalization and unification may be more straight-forward. This actually means that the same simulation approach can be applied to several systems or several configurations of one system. Of course this is not a panacea and in practice it turns into very different performances on every particular system under test. However, up to now the generalization problem in SDS field was considered to be very complex, so the authors of [2], [6], and [7] note the appropriate solutions are still of high demand.

A third hypothesis recalls typical SDS user patterns. The authors of [9] did a trustworthy marketing research and usability expert workshop. As a result, they have shown that "for a complex field, such as interaction behavior and usability requirements, a one-dimensional description with types might be an oversimplification". Therefore, we propose to make use of a three-dimensional descriptive model of a particular user instance covering memory, behavioral, and motivation (MBM) aspects. It is important to note that a prospective fourth dimension is the emotional state of the user, which might be considered in an eMBM model. However, due to the complex, latent nature of this phenomenon, and because of some technical issues (e.g. reliable emotion recognition techniques) it still stands a very complex task even for a separate scientific research. Thus, the final hypothesis claims the MBM model captures the major aspects of users' behavior for the target SDSs described in the two previous hypotheses.

3 Methods Proposed

The Mental Models (MeMo) project described in [4] was a recent pioneer attempt to develop a probabilistic rule-based user model for usability evaluation. A big work has been done so far, resulting in positive interest from the scientific community. Therefore, the present research proposal will take up some interesting ideas introduced in MeMo. However, the practical experience has shown that a rule-based approach has some significant drawbacks: the derived rules are often not precise enough to encode all behavioral aspects of human users. On the other hand, due to inherent corpora limitations, a purely data-driven method can not be safely applied to derive the complete user model from the observed interactions. Thus, a hybrid solution of both data-driven and rule-based descriptive models is considered for implementation. In order to predict user behavior, interaction parameters are learned from data, and a top logic of decision-making is affected by rules. Coupled by adaptive motivation system able to react on dynamic environment changes, this paradigm should enable enough flexibility, and coverage of patterns unseen in learning data.

Behavioral aspects are divided into two following classes (both represented by rich set of features): personal user characteristics and observable performance parameters. At first, a cluster analysis is performed on observable performance parameters to detect user stereotypes from the learning data. The factor analysis is then done to match these stereotypes into the user taxonomy and define the appropriate set of personal user characteristics. Of course, this task might fail due to missing or inconsistent data. Gathering the domain-specific knowledge is also a challenge. Therefore, some semantic web search techniques will be used for this propose, supported by an ontology-based reasoning mechanism.

Besides that, learning the optimal task completion path is also an interesting problem for simulating realistic task-driven interactions. Therefore, an appropriate machine learning algorithm will be used to compute this path. Recently, various reinforcement learning (RL) techniques became widely popular in SDS development and related research fields. However, several important constraints make the pure RL not as attractive in practice as it looks in theory. For instance, high computational costs coupled by learning data requirements make it infeasible in practice to find the optimal solution. Moreover, once configured, many classic RL algorithms do not provide enough flexibility and scalability. Last but not least, there is an oversensitivity problem with a meaningful reward function that is very often hand-crafted empirically. Thus, we propose to implement advanced RL technique dealing with the named limitations efficiently. A prospective approach might be the fuzzy reinforcement learning (FRL) paradigm [10].

4 Sketch of Proposed Solution

The proposed solution is planned as a part of the SpeechEval project carried out in collaboration with the German Research Center for Artificial Intelligence (DFKI). The issuer simulation will be implemented on the Ontology-based Dialog Platform (ODP) supporting several standard voice interfaces. The solution itself is being implemented as a standard Java plug-in supported by ODP. The frontend ASR and speech synthesis services are high-quality ready-to-use business solutions by Nuance. A MySQL database engine is a backend transactions support.

The learning phase is based on the Voice Awards contest corpus [8] provided by DFKI. The development phase grounds on a research SDS, the Bochum Restaurant Information System (BoRIS) and its support corpora gathered in several experimental series [1]. Besides that, Let's Go! Lab [11] and ICT [9] projects' results are used to study usability-related user characteristics.

5 Conclusions

User simulation for evaluation is a young growing field of research. It is still open for significant contributions through applied and fundamental research activities that surely are both of scientific interest and commercial relevance. Thus, the major goal of this research project is to introduce and verify an innovative hybrid approach to user simulation, joining the best results from the SDS field with special focus on usability. The authors hope the research will change the classic role of user simulation in the development and evaluation of SDSs.

A minor goal of this project is to introduce the special FRL technique, adapted for SDS and optimized for usability criteria. Moreover, the structure of this research encapsulates the opportunity to incorporate and validate the results from previous projects in the field, done at T-Labs in cooperation with various partners. Last, but not least is the generalization problem of user simulations across systems and user groups, that stays unsolved up to now. We would like to demonstrate that it is possible to generalize our solution at least within different configurations of the same system.

References

1. Möller, S.: Quality of Telephone-based Spoken Dialog Systems. Springer Science + Business Media Inc., New York (2005)
2. Ai, H., Weng, F.: User Simulation as Testing for Spoken Dialog Systems. In: Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue, Columbus, Ohio (2008)
3. Chung, G.: Developing a Flexible Spoken Dialog System Using Simulation. In: Proceedings of ACL 2004, Barcelona, Spain (2004)
4. Möller, S., Englert, R., Engelbrecht, K.-P., Hafner, V., Jameson, A., Oulasvirta, A., Raake, A., Reithinger, N.: MeMo: Towards Automatic Usability Evaluation of Spoken Dialogue Services by User Error Simulations. In: Proceedings of InterSpeech 2006, Pittsburgh (2006)
5. Mairesse, F., Walker, M.: PERSONAGE: Personality Generation for Dialogue. In: Proceedings of the 45th Annual Meeting of the ACL, Prague (2007)
6. Pietquin, O.: User Simulation/User Modeling: State of the Art and Open Questions. CLASSiC Project Consortium Meeting, Issy Les Moulineaux (2008)
7. Schatzmann, J., Georgila, K., Young, S.: Quantitative Evaluation of User Simulation Techniques for Spoken Dialogue Systems. In: 6th SIGdial, Lisbon, Portugal (2005)
8. Steimel, B., Jameson, A., Jacobs, O., Pulke, S.: Voice Awards 2007: Die Besten deutschsprachigen Sprachapplikationen. Project Deliverables (2007)
9. Hermann, F., Niedermann, I., Peissner, M.: Development of Usability-Oriented User Taxonomy. Report on Procedure, Methods, and the Resulting User Taxonomy (2007)
10. Berenji, H.: Tutorial: Fuzzy Reinforcement Learning. In: IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2007. Imperial College, London (2007)
11. Eskenazi, M., Black, A., Raux, A., Langner, B.: Let's Go Lab: a platform for evaluation of spoken dialog systems with real world users. In: InterSpeech 2008 proceedings (2008)

User Aware Technology: From Inter-human Awareness to Technology-User Awareness*

Ditte Hvas Mortensen

Bang & Olufsen A/S
Peter Bangsvej 15, Struer, Denmark
Department of Psychology
Aarhus University, Aarhus, Denmark
dittehm@psy.au.dk

Abstract. The present project explores how the development of user-aware technology may be supported by applying knowledge from research in psychology on nonverbal aspects of human activity in context.

Research Area: Human-Technology Interaction, User-Aware Technology, Multimodal Interaction.

1 User-Awareness

As programmed technology is embedded into a growing array of objects demanding the user's attention, the importance of creating seamless and intuitive interaction that puts the least possible strain on the user's attention span becomes apparent. To achieve this there has been an increased interest in creating user-aware technology in recent years, especially within the field of pervasive computing. User-awareness can be understood as part of the broader concept of context-aware computing [1]. By being aware of different aspects of the user and her behavior, user-aware technology tries to adapt to the user without receiving explicit commands. A common approach for achieving this has been based on trying to predict the user's activity and wishes, by collecting extensive behavioral information about the user in question [2]. This is often a substantial task, as relevant information about a specific user is a constantly changing and in principle infinite field. Other common attempts at this are more technology driven and define user awareness more narrowly, based on what is relevant to a specific product [e.g. 3].

The present project seeks to contribute to former work by being situated between these two approaches. It focuses on examining more general principles for how technology can be aware of users, by exploring what the most basic objective indicators about the user's intentions may be, and how technology can be made aware of these. This is done through focusing on the user's nonverbal behavior, such as what the user's gaze-direction or movements indicate about her intentions.

The theoretical foundation of the project is based on research in cognitive and developmental psychology that study basic patterns of multimodal nonverbal human

* Advisors: Klaus B. Bærentsen, Søren Bech and Marianne Graves Petersen.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 884–887, 2009.

behavior. At the most basic level the aspects are e.g. body movements, facial expression, gaze direction, gesture, and prosody such as tone, tempo, and pauses in speech [4-7]. This research suggests that we have large cognitive resources for expressing ourselves and perceiving the intentions of others nonverbally. This behavior is often outside of our conscious field of attention, and takes place as a constant adjustment to the activity we are involved in, to other persons and to objects around us. If a user is trying a new technology for the first time, for instance, her movements during the interaction will be slow and hesitative, but when she becomes more accustomed to the interaction they will change to being faster and goal oriented. This change in movements takes place outside of the user's conscious focus, and can often be easily interpreted by others watching her. It could be said that we act more or less instinctively on this level, rather than using a substantial amount of knowledge and model building [8, 9].

Because most nonverbal behavior take place outside of our conscious attention, it is important that the user-aware technology works outside of the user's conscious focus as well. The user should be able to focus on what she wishes to achieve from the technology, rather than on the operational details of the interaction. One way to help accomplishing this is by using the information given by the user's nonverbal behavior, to interpret what her intentions are, in order to narrow down the offered interaction possibilities.

1.1 Hypothesis

The hypothesis of the project is that technology which supports the user's interaction by being aware of, and reacting to, objective indicators of the user's intentions in nonverbal behavior is easier and more pleasurable to use.

2 Method

The project is divided into three parts:

1. An analysis of basic nonverbal human behavior, through qualitative pilot studies and a thorough review of psychological research in the area, to find the most relevant user information for technology
2. An analysis of the technological possibilities and limitations in creating user-aware technology.
3. An experimental testing of the hypothesis by implementing the results from part 1 and 2 in concrete functions.

2.1 Progress So Far

In part 1 and 2 of the project, relevant technologies and aspects of nonverbal behavior are being analyzed and pilot tested individually. Nonverbal behavior can be understood as a multimodal activity. Thus, technology that reacts to the objective indicators of the user's intentions, through perception of nonverbal behavior, should function multi-modally [10]. It is however necessary to begin by analyzing the aspects of nonverbal behavior individually, through a review on the research done on each to identify which aspects are most basic to expressing intentions, what their precise function are as well as the technologies available to monitor them.

Through the review of the literature done so far, it has been decided to focus the further studies on two aspects of nonverbal behavior, namely gaze-direction and movements/ locomotion. These aspects are generally acknowledged as some of the most fundamental ways of expressing intentions, studies have shown that intentions expressed through gaze direction and movement can be shown and perceived very early in human development, before we are able to talk. Gaze-direction and movements, can also be implemented in a wide variety of technologies, most interfaces already require visual and haptic interaction, whereas for instance prosody would require verbal interaction through speech recognition, which is more rare [4, 11].

Besides the individual aspects, it has also been studied how it is possible to set up some general rules for an awareness of the user's intentions, when it is recognized that activity cannot be separated from the context that it is a part of [12]. As a part of this a qualitative pilot study of the activity of solving a task, focusing on gaze related behavior, has been carried out. This explores if an analysis of the specific action being performed can be generalized to other settings where similar actions are being performed.

2.2 Planned Studies

In part 3 of the project, the hypothesis will be tested experimentally. Two A/B experiments will compare a technology that uses gaze direction and movements to perceive and react to the user's intentions, with an otherwise similar technology that does not use this. In each experiment it is planned to compare the two versions of the technology on the dependent variables precision in interaction, time, and user preference. If the assumption that it is easier to use technology that reacts to nonverbal aspects of human behavior is correct, it is to be expected that study participants are faster and more accurate with the user-aware technology, and that they prefer this form of interaction.

To ensure the validity of the experiments and in order to avoid incorrectly rejecting the hypothesis, it is very important that the stimuli are designed correctly, so that they actually represent relevant nonverbal behavior connected to the expression of intentions. This will be ensured by combining qualitative studies with an extensive review of research, and by focusing on the most basic principles of communication.

3 Expected Contributions

Expanding the scope of human technology interaction from explicit commands through pressing buttons, touch screens and the like, to a wider range of human behavior potentially has benefits for all types of applications. Technology that is able to perceive and react to the basic aspects of nonverbal user behavior would make use of the cognitive resources we have for perceiving and expressing these, without much increase in the cognitive load involved in the interaction, because they exist outside of the user's conscious field of attention.

References

1. Schmidt, A., Beigl, M., Gellersen, H.-W.: There is more to context than location. Computers and Graphic 23(6), 893–901 (1999)
2. Dourish, P.: What we talk about when we talk about context. Personal and Ubiquitous Computing 8(1), 19–30 (2004)
3. Qvarfordt, P., Zhai, S.: Conversing with the User Based on Eye-Gaze Patterns. In: CHI 2005, pp. 221–230. ACM Press, Oregon (2005)
4. Turati, C., Simion, F., Milani, I., Umilta, C.: Newborns' preference for faces: What is crucial? Developmental Psychology 38(6), 875–882 (2002)
5. Brooks, R., Meltzoff, A.N.: The Importance of Eyes. Developmental Psychology 38(6), 958–966 (2002)
6. Krauss, R.M., Morrel-Samuels, P., Colasante, C.: Do conversational hand gestures communicate? Journal of Personality and Social Psychology 61(5), 743–754 (1991)
7. Henderson, J.M.: Human gaze control during real-world scene perception. Trends in Cognitive Sciences 7(11), 498–504 (2003)
8. Suchman, L.: Human-Machine Reconfigurations: Plans and Situated Actions. Cambridge University Press, New York (2006)
9. Bateson, G.: Redundancy and Coding. In: Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution and epistemology, pp. 417–431. Jason Aronson, Northvale (1972)
10. Gibson, J.J.: The Senses Considered as Perceptual Systems. Houghton Mifflin, USA (1966)
11. Tapia, E.M., Intille, S.S., Larson, K.: Activity Recognition in the Home Using Simple and Ubiquitous Sensors. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 158–175. Springer, Heidelberg (2004)
12. Barker, R.G.: Ecological Psychology. In: Concepts and Methods for Studying the Environment of Human Behavior. Stanford University Press, USA (1968)

User eXperience: Tools for Developers[*,**]

Anssi Jääskeläinen

Lappeenranta University of Technology, Department of Information Technology,
Skinnarilankatu. 34, 53850 Lappeenranta, Finland
{jaaskela,Jari.Porras,Kari.Heikkinen}@lut.fi

Abstract. This dissertation aims to narrow the user experience gap between the design and test phases in software development by offering UX tools for the developers.

Keywords: User experience, UX, software development, query tool, database.

Research area: Tools for Design, Modeling, Evaluation.

1 Introduction

User experience (UX) has been an important topic in e.g. travel business for a long time [1] and currently UX is gaining interest in the field of software development as well [2], [3], [4]. UX related requirements are raised up in the design phases and in the test and marketing phases, but there is still a huge gap between those. Programmers might see the UX requirements created by the designers but since the programming language or used tool does not support the designed features they are overlooked and passed by or replaced by similar functionality created by the programmer itself. The problem is that programmer seldom possesses enough awareness about UX to be able to make acceptable compromises [5]. This means that a lot of money might be wasted on programming that needs to be redone due to programmers' wrong decisions that has not been verified with the designers.

A simple solution would be to approve every change and replacement from the designers, but often e.g. in many projects the deadlines and deliverables have been pre-set and there simply is no time to approve every single change. My personal experience in multiple projects at university has revealed this same fact; deadlines are usually set to real tight due to financial and funding aspects.

Intention of my dissertation is to shrink the gap between the design and test phases by offering UX tools for developers. A resulting tool set can help developers to solve simple UX vs. reality problems without consulting UX designers and to reduce fatal mistakes in UX caused by bad decisions. By all means, this tool is not and will not be a replacement for the designers! It only tries to increase the likelihood that the whole development process is completed in time and with better results, thus every decisions

[*] Supervisor: Jari Porras.

[**] Advisor: Kari Heikkinen.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 888–891, 2009.

does not have to go through the designers. My claim is based on this chapter and is. "*By offering UX tools for developers the likelihood of successful software development process is increased.*"

1.1 Related Work

Multiple studies about the effect of usability in UX have been conducted. E.g. According to [6] usability is a key issue when designing DVD menus that improve the overall UX. On the other hand, in [2] authors have studied the possibility to use existing usability reports as tools to increase the influence of UX on the strategically important levels in the company. Although, these and many other studies, have proven that usability forms a great amount of UX it is certainly not the only affecting matter. Scholars have studied that, in addition to usability, the following factors have impact for UX.

- Emotional reactions: heart rate, galvanic skin response and heart rate variability[7]
- Fundamental needs, wishes, necessities of the user / target [8].

2 Current Status

Foundations to my dissertation were experiences and the work carried out in two projects called Ampers[1] and Virtahepo[2]. This dissertation continues the path built by those projects by combining four different sections presented in Figure 1. The first section is experience economy, where the research aim is to raise personal understanding of UX in economy point of view and to understand why it is considered so important in business. This section is completed. Two papers were written and published [9, 10] and several reference applications, were built. These papers and applications contributed to guidelines for multiple aspects of actual UX among different population groups and offered a path to integrate ISO 9126 software quality standard for measuring UX.

The second section is software metrics and the idea is to test how well the existing software metrics suits for measuring the UX level. This section is completed. Two different multiplayer games for youths were implemented by using MUPE (Multi-User Publishing Environment) platform[3]; Greenhouse and Detector. Greenhouse worked as a primary reference when the first set of high level UX metrics were rationalized from the ISO 9126 software quality standard while utilizing the learning from experience economy section [11]. Detector was used in a publication which introduces context-aware service development under MUPE, and analyzes differences between professional and amateur developers [12].

[1] http://www.tbrc.fi/eng/projects/?PCID=32&PID=33
[2] http://virtual.vtt.fi/virtual/virtahepo/
[3] http://www.mupe.net/

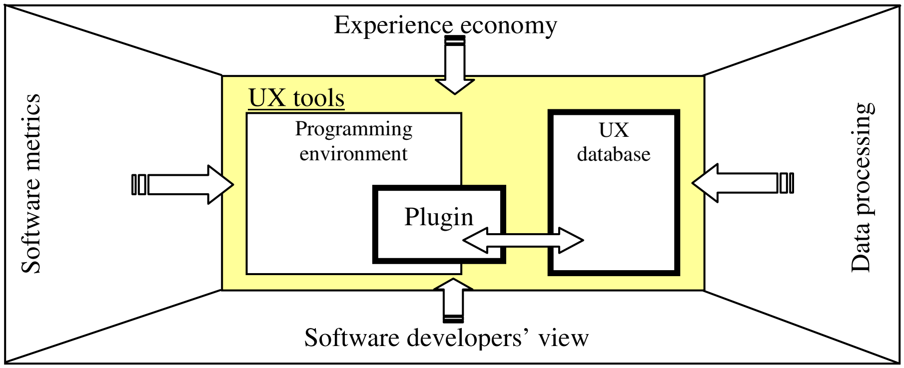


Fig. 1. Sections of UX tools

Currently sections 3 (software developers' view) and 4 (data processing) are under way. Combined goal from these sections is to widen my personal aspects with the experience gained from other software developers and to gain sufficient amount of data to achieve statistical reliability. So far the work in these sections has been conducted by creating UX surveys for programmers in different programming related courses. Surveys intents to find out what causes the feeling of experience while developing and using some software. Secondary goal is to study the effects of UX in learning and programming.

A second iteration of surveys is currently being done in equivalent programming courses to collect enough data to achieve statistical reliability. This collected data is then being analyzed and transferred into a database. Preliminary results and evaluations indicate some promising points that should be further studied, e.g. a story behind a programming assignment could make it more pleasant and therefore more experiencing. Results from these surveys will be published in suitable journals or conferences.

2.1 Future Work

Currently the stub of database contains only one table with about hundred data rows, but after adequate amount of reliable data has been collected, the database structure will be enhanced. Of course, before this enhancing process can begin, the collected data must be validated. A great helping hand is given by one M.Sc. student, whose topic is testing and validating collected data. I really look forward to see the results, thus they can really help me in final sections of my dissertation. Target of this last phase is to enhance the database so that it can match the specific needs of the tool for the developers. Intended deadline for final database structure is the end of this year, but of course the data collecting will continue further on to create larger and more thorough database.

After the database structure is near to finish it will probably be released to test use in our university to collect real usage statistics from the programmers. This valuable usage data can help to further develop the structure and efficiency of the database. Later on a plug-in for programming environment (e.g. Eclipse or Visual Studio) is developed with the ability to make highly specified queries to UX database with a simple mouse clicks. With this possibility, it is easy for the programmer to check the current project or functionality under development against the information held by the database. In addition to easy queries, this plug-in will also contain the formulas to count and predict the UX rate of the defined user type. Plug-in will present the data with the format specified by the developer. Possible formats are going to be plain numbers, different types of charts and statistical views.

2.2 Expected Contributions

- Shrink the “UX free” gab between the design and test phases
- Functional UX database
- UX tools for developers (plugin connected to UX database).

References

1. Lapland Centre of Expertice for the Experience Industry (LEO), Articles on Experiences, http://tinyurl.com/dm49dp
2. Rönkkö, K., Winter, J., Hellman, M.: Reporting user experience through usability within the telecommunications industry. In: Proc. International workshop on Cooperative and human aspects of software engineering 2008, pp. 89–92 (2008)
3. Ashley, J., Desmond, K.: Success with user-centered design management. Interactions 12(3), 27–32 (2005)
4. Najafi, M., Toyoshiba, L.: Two Case Studies of User Experience Design and Agile Development. In: Proc. AGILE 2008, pp. 531–536 (2008)
5. Buxton, B.: Sketching User Experiences: Getting the Design Right and the Right Design. Morgan Kaufmann Publishers, San Francisco (2007)
6. Költringer, T., Tomitsch, M., Kappel, K., Kalbeck, D., Grechenig, T.: Implications for designing the user experience of DVD menus. In: Ext. abstracts on Human factors in computing systems CHI 2005, pp. 1565–1568 (2005)
7. Lottridge, D.: Emotional Response as a Measure of Human Performance. In: Ext. abstracts on Human factors in computing systems CHI 2008, pp. 2617–2620 (2008)
8. Keinonen, T.: User-centered design and fundamental need. In: Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges, pp. 211–219 (2008)
9. Jääskeläinen, A., Heikkinen, K.: Designing and implementing www-based multi-user virtual environment: Case youth work. In: IADIS e-Society 2006, pp. 261–265 (2006)
10. Jääskeläinen, A., Heikkinen, K.: Designing, Implementing and Testing Experiential Multi-User Virtual Community with MS-Patients. IADIS Int. Jour. WWW/Internet, 147–164 (2007)
11. Jääskeläinen, A., Heikkinen, K.: Utilisation of GQM method to map ISO 9126 for experience measurements. In: SOFTCOM 2008 (2008)
12. Jääskeläinen, A., Lautamäki, J.: Analysing Context-Aware Service Development under MUPE Platform. In: ASWN 2008 (2008)

A Dynamic Environment for Video Surveillance

Paolo Bottoni[1], Maria De Marsico[1], Stefano Levialdi[1,2], Giovanni Ottieri[1], Mario Pierro[1], and Daniela Quaresima[1]

[1] Department of Computer Science "Sapienza" University of Rome, Via Salaria. 113, 00198 Roma, Italy
[2] DEI, Carlos III University, Madrid
{bottoni,demarsico,levialdi}@di.uniroma1.it, mat101@gmail.com, {jowans,daniela.quaresima}@libero.it

Abstract. Video surveillance systems must support multiple streaming and prompt alert notification. We propose a two-tiered environment: a supervisor defines presentation layouts and model interface reactions to alerts; a surveillant watches synchronized videos, adapts layouts, and is notified with alerts.

Keywords: Video surveillance, Interface reconfiguration, Synchronization.

1 Introduction

Video surveillance is evolving from analog to digital, decoupling video sources from presentation devices and enabling random access to scenes, automatic processing of anomalies and advanced interfaces [1]. However, surveillants, still have to react to alerts and automatic recognition of individuals or suspect behaviours is complex [2,3]: small variations in algorithms parameters cause false negatives / positives, wasting attention resources or diminishing users' trust in the system. An effective interface should support event interpretation via multiple cameras, and ensure proper information on alerts for timely redirection of attention. Hence, we organize subsets of videos into *views*, with suitable stream layouts, typically related to camera arrangement. The interface always presents the whole stream set at reduced resolution; and a scheduler ensures adaptable frame rates for video refresh. Flow synchronization provides optimal resource usage and presentations suited to the task. Stream rearrangement and resynchronization may comply with surveillants' dynamic needs or with supervisor's policies – e.g. to present streams from an alert area. The overall interface structure is kept constant, while content is dynamically adapted adapted to ensure full control whilst reducing attention stress.

2 Related Work and Requirement Analysis

Research on usability of interactive digital surveillance systems is recent. Research prototypes often assume well-controlled and synchronized settings and exclude humans from the processing loop. However, [4] advocates user involvement through

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 892–895, 2009.

usable interfaces. Focusing on interaction, we consider the impact of usability flaws encountered by surveillants [5], especially related to performance degradation over time and divided attention among cameras. Commercial systems map camera banks onto window banks (*views*): a surveillant selects cameras to trace activities through views. With many cameras, small images make this difficult; moreover, it is hard to anticipate where a traced subject will reappear. Help comes from camera maps, and from experimental layout algorithms arranging peripheral streams according to their relative positions: a "geographic" context helps tracing a subject [6]. Our scheduling derives from Ptolemy's HDF [7]. Each iteration is a state in which flows stay constant and the number of required activation cycles is determined. After each iteration, admissible modifications can result in revision of the scheduling.

Literature and interviews with surveillants provided the main requirements: continuous area overview and rapid focus switch on relevant information in case of alarms. Users also need to modify stream arrangements and frame rates and to relate each stream with the different views. To this aim, our system provides a map and uses identifiers and color codes to indicate camera positions in a window grid in which streams are played. Continuous information is kept for auditing: by logging all modifications of the interface organization, either due to user interaction or to generated alerts, and keeping track of users' accesses. Interviews with experts also highlighted the need for two classes of users, *supervisors* and *surveillants.*

3 System Use and Architecture

A network of services constitutes the server side. One service feeds the interface and the streams, the others perform recognition activities. Fat clients allow access as *Supervisor* or *Surveillant*. Supervisors manage stream identity and layout in a view, and define a set of predefined views to address frequent or critical needs. They also define or add servers for automatic alert triggering, and include or exclude cameras from views, marking them on the map. Supervisors can select sources of different types (files, web-accessible or closed-circuit cameras) and specify their paths and access rights. Both supervisors and surveillants can further define parameters, e.g. source frame rate, position and size. Modifications of frame rates or of the composition of the active view require rescheduling. The surveillant environment implements the *overview + detail* paradigm [8]. A panel constantly provides the whole streaming set in a reduced size and a map of the cameras with positions and observation cones. Such elements are kept constant as top (*Stream*) and bottom-right panels (*Map*). The bottom-left panel (*Focus*) presents the currently selected view: windows presenting streams at the chosen occupy grid cells to allow a closer inspection of specific sources. Fig. 1 shows an interface screen with a chosen *view*. A surveillant can adapt a presentation by adding or removing streams to and from the *Focus*. Modifications cannot be made permanent, except for the dedicated *surveillant view*. Alerts can activate a predefined view, or push specific information in the normally empty *alert cell*. Fig. 1 shows how subject recognition from a database of suspects triggers the visualization of the source stream in the alert cell, while the database record is presented in the *Information Panel* (above the *Map* Panel).

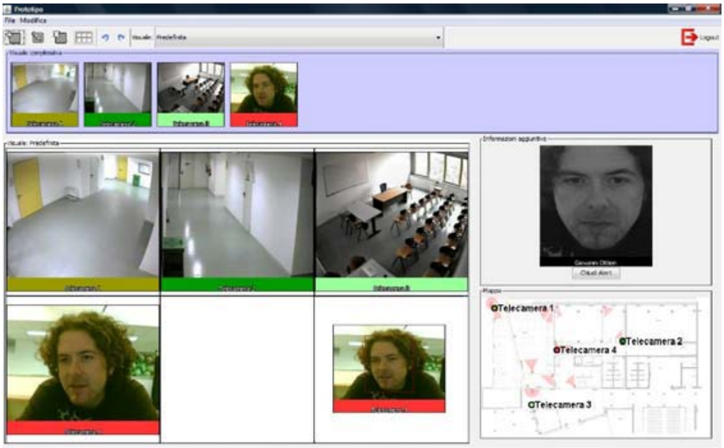


Fig. 1. In an alert situation the *Information* panel presents info on the recognized person, while the live stream in which the person was recognized is shown in a cell of the *Focus* panel

An alert can cause replication of a stream on several windows: potentially dangerous situations are rapidly assessed, by attracting attention through interface changes, possibly complemented by sound signals. Frame rates of streams can be singly adapted, e.g. streams from less visited zones may be kept at a low frame rate, automatically increased if something occurs there. Typically, frame rates in the *Stream* panel are slower than in the *Focus* one. A `Scheduler` manages the single rates. Events are classified according to possible impact on the video presentation rate, and identified by interpreting stimuli from the user or from servers. Each interpretation defines a request, associated with a command sequence, possibly leading to rescheduling and to redefining the *Focus* panel. Each modification request is routed to a specific `Command Manager` for each panel.

4 System Usability

The usability analysis was introduced in the development process through a cognitive walkthrough, and an analysis based on Nielsen's heuristics. We then classified the identified design flaws according to Cognitive Dimensions [9], only considering those relevant to the designed tasks. We especially analyzed view adaptation in the supervisor and surveillant interfaces. As for the notions of *notation*, *environment* and *media* upon which the analysis is based, we respectively identified: a) the collection of windows within the *Focus* panel; b) the user commands to add, remove, or change visualization parameters of selected videos; and c) the *Focus* panel, where the user can manipulate symbols. Due to lack of space we only highlight some examples of the evaluation and redesign process. For <u>Abstraction: types and availability of abstraction mechanisms</u>, we observed that abstraction creation and management are straightforward. However, it was difficult to understand at first sight that the view can contain any view from a menu. A simple workaround was to expand the label of the view panel. A problem with Premature commitment: constraints on the order of doing things was related to the expected position of a newly added element. We decided to use a fixed grid for positions.

5 Conclusions

We supports surveillants in rapidly focusing attention on relevant sources, while maintaining an area overview. A first tier defines a supervisor environment, while in a second, the surveillant can change predefined settings according to current needs. The system can present relevant streams when predefined events occur. This avoids missing relevant events due to difficult orientation within a stream bank, and reference to physical locations. Users can personalize frame rates.

Acknowledgments. Partially supported by MIUR – PRIN 2006.

References

1. Valera, M., Valestin, S.A.: Intelligent distributed surveillance system: a review. IEEE Proceedings of Image Signal Process 152, 192–204 (2005)
2. Micheloni, C., Snidaro, L., Piciarelli, C., Foresti, G.L.: Exploiting Temporal Statistics for Events Analysis and Understanding. In: Proc. ICIAP 2007, pp. 530–535. IEEE CS, Los Alamitos (2007)
3. Bremond, F., Thonnat, M.: Issues of Representing Context Illustrated by Video-Surveillance Applications. Int. J. of Hum.-Comp. St. 48, 375–391 (1998)
4. Yamasaki, T., Nishioka, Y., Aizawa, K.: Interactive Retrieval for Multi-Camera Surveillance Systems Featuring Spatio-Temporal Summarization. In: MM 2008, pp. 797–800 (2008)
5. Keval, H.U., Sasse, M.A.: Man or a Gorilla? Performance Issues with CCTV Technology in Security Control Rooms. In: 16th World Congress on Ergonomics Conference (2006), http://hornbeam.cs.ucl.ac.uk/hcs/publications/IEA2006_Keval.pdf
6. Girgensohn, A., Shipman, F., Turner, T., Wilcox, L.: Effects of Presenting Geographic Context on Tracking Activity between Cameras. In: Proc. CHI 2007, pp. 1167–1176. ACM Press, New York (2007)
7. Brooks, C., Lee, E.A., Liu, X., Neuendorffer, S., Zhao, Y., Zheng, H.: Heterogeneous Concurrent Model and Design in Java. Ptolemy II Domains, vol. 3, UCB/EECS-2008-37 (2008)
8. Cockburn, A., Karlson, A., Bederson, B.B.: A review of overview+detail, zooming, and focus+context interfaces. ACM Comp. Surv. 41(1), 1–31 (2008)
9. Green, T.R.G., Petre, M.: Usability Analysis of Visual Programming Environments: A 'Cognitive Dimensions' Framework. Journal of Visual Languages and Computing 7, 131–174 (1996)

An Integrated Approach for Creating Service-Based Interactive Applications

Marius Feldmann[1], Jordan Janeiro[1], Tobias Nestler[2], Gerald Hübsch[1], Uwe Jugel[2], André Preussner[2], and Alexander Schill[1]

[1] Technische Universität Dresden, Department of Computer Science, Institute for Systems Architecture, Computer Networks Group
{marius.feldmann,gerald.huebsch,jordan.janeiro, alexander.schill}@tu-dresden.de
[2] SAP Research CEC Dresden, 01187 Dresden, Germany
{uwe.jugel,tobias.nestler,andre.preussner}@sap.com

Abstract. While the implementation of business logic and business processes based on service-oriented architectures is well-understood and covered by existing development approaches, integrated concepts that empower users to exploit the Internet of Services to create complex interactive applications are missing. In this paper, we present an integrated approach that fills this gap. Our approach builds upon service annotations that add meta-information related to user interface generation, service dependencies, and service composition to existing service descriptions. Services can be composed visually to complex interactive applications based on these annotations without the need to write any code. The application code is generated completely from the service composition description. Our approach is able to support heterogeneous target environments ranging from client/server architectures to mobile platforms.

Keywords: services, composition, annotations, interactive applications, MDA.

1 Introduction

Service-oriented architecture is a design concept that implements system functions as a set of highly reusable and distributed services. These services are composed to implement business applications and processes. Existing work in this area mainly focuses on these aspects and does not sufficiently address the composition of services to create service-based interactive applications. In these applications, the user directly interacts with services via a rich user interface (UI). In this context we assume that reuse in service-orientation is not limited to the implementation of system functions. Reuse can be extended towards the aspects user interaction and service composition through *services annotations* that attach meta-information to services, thereby creating *annotated services*. The UI of a service-based interactive application can be composed out of reusable *UI fragments*, each of which is attached to a data type, parameter, operation or the service itself as an annotation. For example, a service operation that provides user authentication would be associated with one UI fragment

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 896–899, 2009.

for entering login information, and two other UI fragments for displaying the authentication result ('successful', 'not successful'). With regard to service composition, service annotations specify dependencies between services. For example, a service for editing user profiles will require an authenticated user. It depends on the presence of an authentication service in the composition. This dependency can be expressed in its annotation, including a reference to the required authentication service. The main benefit of service annotations is that they limit the effort for the development of service-based interactive application to a purely model-driven, visual composition of annotated services. In section 2, we discuss our approach in more detail. We summarize and conclude our findings in section 3.

2 Development Approach

In this section, we present our integrated development approach for service-based interactive applications. It builds upon the basic concept of service annotations which are explained in more detail below. The annotated services build the foundation for the composition of interactive applications. For this purpose, we define a *composite application model* (CAM). Annotated services can be composed in a visual way and transformed to CAM artifacts. The service composition includes arranging UI fragments into pages, defining the page flow, and the data flow between annotated services. A CAM describes the aspects of an interactive application and can be transformed into executable code, which implements the interactive application through *model-to-code transformations* in a process called runtime generation. Due to the technology independence of the underlying annotations and the CAM, multiple target platforms can be supported. We also define a methodology that ties all of these parts together in an integrated development process.

2.1 Methodology

Our development methodology defines three steps: *service annotation*, *service composition* and a fully automatic *runtime generation*. These steps cover the evolution from a set of services to the complete service-based interactive application and are supported by corresponding tools.

The service annotation step supports technical service descriptions, like WSDL[1]- or WADL[1]-files and produces service annotations. The developer, who annotates a service, performs the role of *service annotator* and uses the *service annotation tool*. Such role requires experience in UI development and technical background knowledge about the services which are to be annotated.

The service composition step builds upon annotated services and produces a CAM. The role of *service composer* defined for this step does not require any programming skills and can be performed by an end-user [1] with a basic understanding of the composition concepts. The service composer identifies and composes the annotated services that are required to implement the interactive application. To produce sophisticated applications, the service composer can add all necessary UI layout definitions, page-, and data-flows using the *composition tool*. During the composition process a

[1] http://www.w3.org/TR/wsdl; http://wadl.dev.java.net

CAM instance is continuously updated, to generate executable code for the platform chosen by the service composer.

2.2 Service Annotations

Service annotations are reusable information fragments which are associated *i)* with service definitions, *ii*) with service operations, *iii*) with input parameters, *iv*) with return values, *v*) with failure messages, or *vi*) with structured data types to facilitate the creation of service-based interactive applications.

We categorized the annotations in three groups: *visual*, *behavioral*, and *relational*. Visual annotations concern static UI aspects, behavioral annotations concern the aspect of the behavior of UI elements, and relational annotations concern dependencies between services. Examples of visual annotations are: labels, tooltips, MIME-type information, design templates, and grouping or ordering of UI elements. Behavioral annotations are, for example: rules for client-side input validation, the specification of a data source for an automatic form completion and the specification of a time interval in which a service performs the call of a certain operation (e.g. information retrieval for news tickers). Relational annotations describe service dependencies that express relations between services from a functional and a data perspective, e.g. that a service requires authentication information and a reference to the service that can provide this information.

More than 20 types of annotations have been identified and formally defined in a meta-model. An instance of this model holds all information provided during the annotation task and a reference to the associated service description. The technology-independence of the information within the model assures a maximum level of compatibility with existing and future target platforms.

2.3 Service Composition

The purpose of service composition is the development of interactive applications based on annotated services. The usage of the service annotations enables a simplified design and creation of an application in a visual manner following the concept of *integration at the presentation layer* [2]. The goal of this concept is to hide the complexity of the actual programming task by representing a service entirely by its corresponding UI. The service composer only works with visual representations to build the application.

Service composition integrates annotated services through the definition of pages, and the specification of page- and data-flows. Each page represents a screen in the final application. Pages consist of UI elements that are arranged according to a layout. They enable the interaction between the user and services. The service annotations are used to automatically infer UI elements, page layout, and additional behavior (e.g., suggestion capability). The service composer may modify the generated pages manually by, for example, relocating UI elements within a page or between pages. To build multipage applications, pages can be linked by specifying a page flow. Data flows specify the flow of information between services. For example, the return value of a service operation may be an input parameter of another service operation. These data flows can be partially derived from service dependencies defined in the annotations.

The CAM that results from the service composition contains all necessary information for generating an executable application for the platform that has been selected by the service composer.

2.4 Runtime Generation

After finishing the design-process, the resulting instance can be transformed into an executable application via a model-to-code transformation process. Due to their high relevance the focus has been directed to fat and rich clients (e.g. AJAX-based applications). For every platform supported by the composition tool such a code generation mechanism is provided. For the prototypical implementation two platforms are supported. Web applications are supported by mapping the CAM to AJAX toolkits embedded into the Spring framework. Code generation is also provided for Google Android to validate the approach for a mobile fat client platform. After the final application has been generated, it can be deployed automatically using different deployment strategies for the supported target platforms.

3 Conclusion and Outlook

This paper presented an approach for the visual creation of interactive applications based on a technologically heterogeneous service infrastructure. The central idea of this approach is the introduction of service annotations. These reusable information fragments are attached to services and can be used within the service composition to semi-automatically instantiate a composite application model for describing interactive applications. Instances of the composite application model serve as input for a model-to-code transformation that generates executable applications.

Future work will concentrate on improving the tool environment for service annotation and service composition. The service composition tool will be evaluated continuously in end-user studies to achieve the intended goal of enabling end-users to create applications without programming skills. These applications will be more sophisticated than those that can be created using current mashup platforms [3]. Furthermore, the mentioned meta-models are improved in an iterative manner based on real-world examples. We also consider making the service annotation model available as an open standard.

References

1. Nestler, T.: Towards a mashup-driven end-user programming of SOA-based applications. In: Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services (2008)
2. Daniel, F., et al.: Understanding UI Integration: A survey of problems, technologies, and opportunities. In: IEEE Internet Computing (2007)
3. Hoyer, V., Fischer, M.: Market Overview of Enterprise Mashup Tools
4. Proceedings of the 6th Int. Conf. on Service Oriented Computing (2008)

Implicit Interaction: A Modality for Ambient Exercise Monitoring

J. Wan[1], M.J. O'Grady[2], and G.M.P. O'Hare[2]

[1] School of Computer Science & Informatics, University College Dublin (UCD), Belfield, Dublin 4, Ireland
[2] CLARITY: Centre for Sensor WWW Technologies, University College Dublin (UCD), Belfield, Dublin 4, Ireland
jessie.wan@yahoo.co.uk,
{Michael.J.OGrady,Gregory.OHare}@ucd.ie

Abstract. Ambient Exercise refers to the implicit exercise that people undertake in the course of their everyday duties - a simple example being climbing stairs. Increasing awareness of the potential health benefits of such activities may well contribute to an increase in a person's well-being. Initially, it is necessary to monitor and quantify such exercise so that personalized fitness plans may be constructed. In this paper, the implicit interaction modality is harnessed to enable the capturing of ambient exercise activity thereby facilitating its subsequent quantification and interpretation. The novelty of the solution proposed lies in its ubiquity and transparency.

Keywords: Ambient exercise, Implicit interaction, Pervasive health.

1 Introduction

A major challenge for heath professionals is how to motivate people to incorporate more physical activity into their daily routine. Increased incidences of obesity and cardio-vascular disease are just two common health problems that could be reduced if people engaged in more physical activity. Though multiple factors contribute to the prevalence of each, the sedentary lifestyle that is common in many societies at present is suspected to be a key contributing factor. Addressing this can be difficult as it is frequently not a life style choice but a side-effect of one's profession. Thus there is a need to reconcile normal daily activities with health and fitness requirements.

Ambient Exercise refers to the implicit exercise undertaken in the course of everyday activities. Examples include, walking to a local convenience store, climbing stairs and so on. Though not formal exercise per se, nevertheless, the cumulative effect of such physical activity could be significant. Increasing awareness of such exercise, and motivating individuals to incorporate more of it into their everyday lives, is a desirable objective. A prerequisite to the successful fulfilment of this objective is the capture, analysis and interpretation of ambient exercise activities such that they can be quantified and incorporated into fitness regimes. However, fundamental to the integration of such exercises into an individual's daily routine is the autonomous and

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 900–903, 2009.

transparent monitoring of such activities. The implicit interaction modality offers one potential model by which this can be achieved.

1.1 The Implicit Interaction Modality

Implicit interaction [1] occurs when users interact with an arbitrary system in a determinable but subconscious manner. It has been harnessed in a number of domains, for example, in electronic tourist guides. Such interaction may be captured as discrete events or as a result of continuously monitoring a stream of data – a pertinent case in this discussion being the continuous monitoring of physiological signals. Harnessing such an interaction modality minimizes the need for explicitly requiring user attention, thereby ensuring that users focus uninterrupted on their tasks.

1.2 Related Research

A number of applications have been described in the broad pervasive health domain that enables monitoring of various activities. (PmEB) [2], hosted on a mobile device, monitors caloric consumption and expenditure; however, details must be manually entered to enable the calculation of an updated caloric balance. SHAKRA [3] detects patterns in GSM signal strength fluctuation to infer the physical activities of the carrier. HealthGear [4] harnesses a suite of body sensor technologies to detect incidences of Sleep Apnea. Similarly, MOPET [5] hosts a suite of wearable sensors to monitor physical fitness activities in an outdoor scenario. Ermes et al [6] explore activity identification in supervised and unsupervised settings. In the next section a prototype is introduced that is distinct in its focus on implicit interaction and harnessing of the intelligent agent paradigm.

2 Engineering an Ambient Exercise Monitor (AEM)

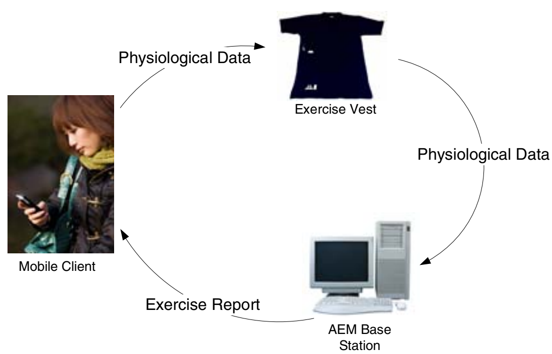
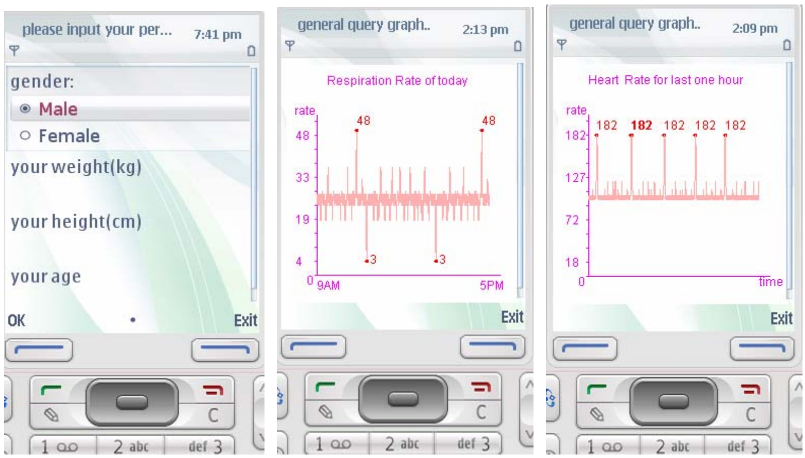


Fig. 1. Architecture of AEM

The architecture of AEM is outlined in Figure 1. Key components include:

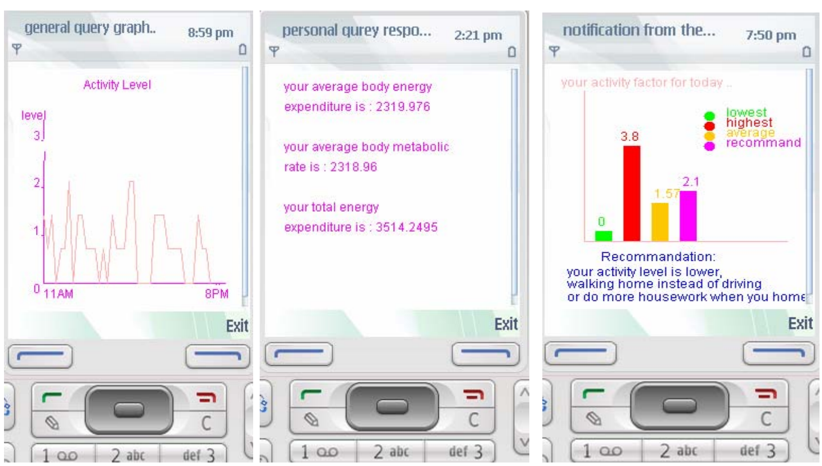
– Exercise vest: A vest incorporating a number of integrated physiological sensors, including heart rate, respiration rate, activity level and skin temperature, is worn by the user. Data from these sensors is continuously collected in a transparent manner and stored on the vest thus enabling activity monitoring in both indoor and outdoor environments. It should be noted that location-sensitive data is not collected, thus significantly reducing privacy concerns.

– AEM Base Station: When in the vicinity of the base station, sensor data is seamlessly uploaded and stored in the data base. A Web interfaces defines the protocols required for database access.



(a) (b) (c)

Fig. 2. User profile construction (a); daily respiration history (b); hourly heart rate record (c)



(a) (b) (c)

Fig. 3. Acquiring the user profile (a); energy expenditure (b); and activity recommendation (c)

- Mobile Client: The mobile client incorporates two embedded agents to deliver key AEM functionality. A User Agent manipulates the interface, communicates with the user, and manages their queries. The Analysis Agent interprets queries, and interfaces with the AEM database.

The AEM client has been implemented in J2ME and deployed on a Nokia N95. The server has been implemented in Java and MySQL. A sports vest, augmented with appropriate sensors has been harnessed to deliver the necessary physiological data.

An initial profile must be constructed to seed the AEM functionality (Fig. 2a). Heart beat and respiratory rate records can be constructed for arbitrary time intervals (Fig. 2b and Fig. 2c). A summary of the activity level is illustrated in Fig 3a while Fig 3b demonstrates the average daily energy expenditure. This can be compared with various norms, resulting in appropriate actions being recommended (Fig. 3c).

3 Future Work

A number of avenues for augmenting the AEM are being explored. The model currently used for calculating energy expenditure is generic so an adaptable model that can be personalised for individual users is being considered. Secondly, it is intended to realise an Intelligent User Interface (IUI) that will provide feedback to users so as to encourage them to partake in further exercise activities, as circumstances dictate. Agents have proved successful when developing IUIs for mobile devices [7] so incorporating this technology should not prove problematic. Finally, it is intended to validate the approach using an initial series of user evolutions that will, it is hoped, lead to further refinement of the AEM.

Acknowledgments. This material is based upon works supported by the Science Foundation Ireland (SFI) under Grant No. (Grant No. 07/CE/I1147)

References

1. O'Grady, M.J., O'Hare, G.M.P., Keegan, S.: Interaction Modalities in Mobile Contexts. In: Virvou, M., Jain, L. (eds.) Studies in Computational Intelligence (SCI), vol. 104, pp. 89–106 (2008)
2. Tsai, C.C., Lee, G., Raab, F., Norman, G.J., Sohn, T., Griswold, W.G., Patrick, K.: Usability and feasibility of PmEB: a mobile phone application for monitoring real time caloric balance. Mob. Netw. Appl. 12(2-3), 173–184 (2007)
3. Anderson, I., Maitland, J., Sherwood, S., Barkhuus, L., Chalmers, M., Hall, M., Brown, B., Muller, H.: Shakra: Tracking and Sharing Daily Activity Levels with Unaugmented Mobile Phones. Springer Science + Business Media (2007)
4. Oliver, N., Flores-Mangas, F.: HealthGear: A Real-time Wearable System for Monitoring and Analyzing Physiological Signals. In: International Workshop on Wearable and Implantable Body Sensor Networks, pp. 61–64 (2006)
5. Buttussi, A., Chittaro, L.: MOPET: A context-aware and user-adaptive wearable system for fitness training. Artificial Intelligence in Medicine 42(2), 153–163 (2008)
6. Ermes, M., Parkka, J., Mantyjarvi, J., Korhonen, I.: Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. IEEE Transactions on Information Technology in Biomedicine 12(1), 20–26 (2006)
7. O'Hare, G.M.P., O'Grady, M.: Addressing mobile HCI needs through agents. In: Paternó, F. (ed.) Mobile HCI 2002. LNCS, vol. 2411, pp. 311–314. Springer, Heidelberg (2002)

Interacting with Casework Documents Using Timelines

Morten Bohøj and Niels Olof Bouvin

Department of Computer Science, Aarhus University
Åbogade 34, 8200 Århus N, Denmark
bohoej@cs.au.dk, bouvin@cs.au.dk

Abstract. We present a way of creating an overview of administrative procedures using timelines. Our design also provides for manipulation of the information given on the timeline through interacting directly with the timeline. As an example application, a prototype supporting administrative procedures surrounding parental leave in Denmark, has been developed and evaluated.

Keywords: Timeline, visualization, manipulation, casework.

1 Introduction

We present a method of using timelines to visualize administrative procedures in the public sector and a way of interacting with the timeline. The visualization of administrative procedures may be interesting in itself, but the main contribution of his work is the interaction with and manipulation of the timeline. The application is a web based application constructed using the Google Web Toolkit.

As a demonstrator for our approach, we have chosen to focus on an application for parental leave, as this is a complex administrative procedure under Danish law. In Denmark you are entitled to leave subsidized by the municipality if you are working. This leave consists of a variety of leave types e.g. four weeks of maternal pregnancy leave and two weeks of paternal postnatal leave. A total of 52 subsidized weeks to be split between the parents are available. These 52 weeks can be extended with 8 or 14 weeks if taken as direct extension to the 52 weeks, but with the same amount of money spread over a greater time span. The 52 weeks can be spread over the first nine years of the child's life, allowing for options such as two hours every other Friday. The systems available today at the municipality cannot handle this flexibility very well and case workers try to adapt by counting days in paper calendars and starting/stopping cases manually. There is no system for the parents to keep track of their leave, only if they themselves keep track.

Our design gives easier access for parents to information and this way takes some of the workload from the caseworkers. Our approach is not limited to handling parental leave applications and can be generalized to other casework applications.

The work presented herein was done under the auspices of the eGov+, a project looking into ways of using web technologies to improve communication between the citizens and the Danish public sector.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 904–907, 2009.

2 Related Work

Timelines are well suited to represent chronologically occurring events, and several systems and UI widget sets have in recent years been developed to exploit this on the web. The focus of most systems are to facilitate browsing with some support for authoring timelines for others to read. A typical example of this class of timeline sites is TimeRime [1], which supports authoring of rich media timelines (rich text, images, and links) and sharing of the same. Other systems emphasis the sharing of timelines, as seen on TimelineIndex [2], which support embedding timelines on other websites. The technologies vary - many systems have previously used Flash (as seen in e.g., TimeGlider [3]), yet modern UI widget sets such as SIMILE Timeline [4] demonstrates what can be accomplished in a modern web browser.

While these systems excel in visualizing chronological events, they are not intended as tools for other activities rather than the presentation of the timeline itself. This is one of the defining differences to the work reported herein, as we actively explore the timeline as the locus of collaborative activity and planning.

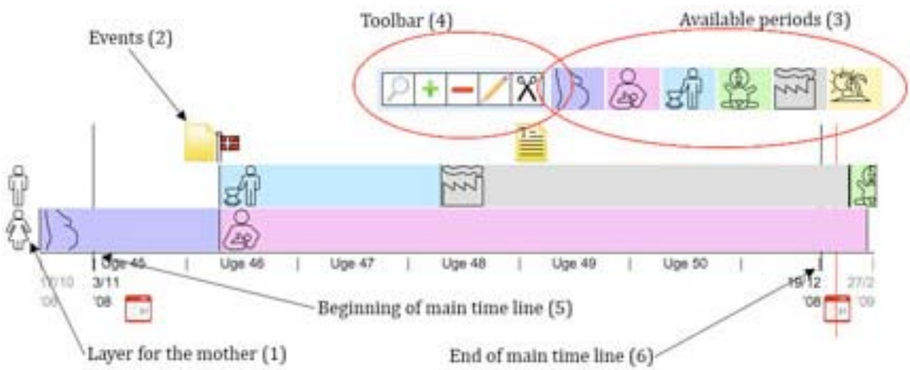


Fig. 1. Screenshot of our prototype

3 Visualization and Manipulation with Administrative Procedures

Parental leave is a serious business - jobs are put on hold to care for the newborn, and the family's economy must be taken into consideration. Parental leave is under Danish law supported by the state (and very often additional leave is supported by the employer depending on trade union agreements). The leave is divided into a number of different periods, such as pre- and post-natal, each period governed by different rules. Thus, the combination of leave periods (which excluding pre- and postnatal leave can shared equally between parents) is quite crucial, when a family is planning the time (usually at least a year) after birth. To help parents plan their leave periods in an accessible manner, we have devised a collaborative web-based tool.

During our initial designs, we investigated several approaches to the presentation and overview of administrative procedures. We present in this poster our work on using timelines as an overarching structural mechanism. Maternity leave consists in

Denmark of different periods of time, where a parent has paid leave of absence from work. Certain parts of this leave can only occur at specific times, and the timeline is well suited to highlight when such periods are applicable. The overall concept is that the timeline should give the applicant a better view of when everything is happening. Our timeline, shown in Figure 1, consists of four layers, the timeline itself, a layer showing periods related to the mother, labeled (1), a layer showing the periods related to the father, and a top layer showing different events, labeled (2), such as birth, expected birth, as well as the documents, which form a critical part of the communication between the parents and the municipality.

The timeline consists of three horizontal sections. The middle section, between (5) and (6), shows a six week time period centered on the current date, and at either end compact sections showing the events which occur outside the main section. One can change the span of the middle section and thereby zoom in and out on the timeline.

The timeline does not only provide visualization and overview of the casework, it is also a tool to interact with and manipulate the information. Interaction with the timeline is done with a combination of direct manipulation and the use of tools selected from a toolbar, labeled (4) in Figure 1.

We have chosen to combine both direct manipulation and tool manipulation because this gives a greater flexibility in the applications supported. Relying only on direct manipulation, where the action executed is chosen by the location or context of the mouse, may restrict the actions available, because only one action can be done at one place, or mouse actions will have to be combined with key presses, giving the user more combinations to remember. We have chosen to include direct manipulation as it gives a nice immediate interaction.

Adding periods of leave is done by dragging periods from the period panel, labeled (3) in Figure 1, and onto the timeline in either the layer for the mother, or the layer for the father. A two-week period will then be added to the respective timeline, starting at the point where the period was dropped. The length of the period can be changed by dragging the ends, making it smaller or larger. Moving the periods is done by clicking a period and moving it to another location. Moving a period off the timeline deletes it. Periods can be cut in two, using the cut tool. This allows the user to move or delete only part of a period.

Events in the timeline, such as birth or application documents, can be either static or allow manipulation. An event such as birth should be static once created and not allow manipulation, whereas events such as application documents should allow manipulation. Events cannot be moved by dragging and the placement on the timeline is decided by internal state of the event.

4 Prototyping

We have conducted user evaluations of our design using prototypes throughout our design process. These prototypes have been shown to both citizens and caseworkers from one of the collaborating municipalities. In the early stages of the design process, we employed paper mock-ups to facilitate feedback on our timeline concept, as well as the general interaction and manipulation. A working prototype implementing some of the features of the design has also been developed and evaluated through workshops. Both citizens and caseworkers have been given the opportunity to interact with the timeline and associated tools. This process led us to our current prototype which supports dragging periods to the timeline and interacting using tools.

5 Generalisability and Future Work

The visualization of casework is not restricted to parental leave and should not only be considered relevant for the outsiders, the citizens, but also for the insiders, the caseworkers. Our timeline design allows for all types of periods and events to be implemented and applied to the timeline. Public work is, at least in the Danish administrative tradition, divided across different departments, such as social, traffic, planning, environment, etc. A case, perceived by the citizen as just one case, may well span several departments and multiple caseworkers. Here, the timeline could be employed to show the current state of the application, letting different periods/layers signify departments or case workers, on whose desk the case (figuratively if not literally) currently resides. This would not only provide the citizens with an insight into the state of their application, but also allow caseworkers to follow cases as they moved to other departments. An example of such a use for the timeline can be found in the ordinary processing of applications (e.g. building permits), which can involve a number of case workers and several departments.

We are currently working on extending the validation of time periods input by the user, so that illegal combinations are caught as early as possible, and also enabling users to see the (financial) consequences of their choices. A simple example of such a needed validation is that you can only qualify for municipal paid leave of absence if you have been employed for 13 weeks in the past year. Some parts of the law apparatus can readily be formalized into a machine decidable system, while other parts will require the experience of a caseworker - we aim with our design to let the caseworkers concentrate on the areas where their expertise is actually needed, rather than being bogged down by trivial, but time consuming tasks that can be handled by our system. We also intend to extend the prototype by implementing more direct manipulation such as allowing dragging the ends of the periods to manipulate length, to allow evaluation of interaction.

References

1. TimeRime, `http://www.timerime.com/`
2. TimeLineIndex, `http://www.timelineindex.com/`
3. TimeGlider, `http://www.timeglider.com/`
4. SIMILE Timeline Widget, `http://www.simile-widgets.org/timeline/`

Measuring Emotional Wellbeing with a Non-intrusive Bed Sensor

Gert van der Vloed and Jelle Berentsen

Industrial Engineering & Innovation Sciences
Eindhoven University of Technology, Den Dolech 2. 5612 AZ, Eindhoven
{G.v.d.Vloed,J.Berentsen}@tue.nl

Abstract. We examine the possibility of using non-intrusive bed sensor measures to ascertain the emotional wellness of an individual. To this end we did a convergent validation study to determine whether heart rate and respiration measures provided by the bed sensor correlate with self reports through questionnaires. The results show that negative affect and worry tension positively correlate with nocturnal respiration.

Keywords: Emotions; Emotional Wellbeing; Physiological Measures.

1 Introduction

Emotions are a primary determinant of wellbeing. In the evolution of our species, reflexive and emotional responses to our environment played a vital role. On being confronted with a predator a human displays the emotion of fear and consequently the sympathetic nervous system triggers the rapid release of energy resources to initiate the fight/flight response. Without this sympathetic response, the possibility of adaptively coping with the such an emotional situation would radically decrease. Although these emotional reactions were evolutionarily selected for in a time very different from modern times they still function as a quick means of the human organism for evaluating its current condition and for selecting between possible behavioral responses [4].

When one gets positively or negatively excited, the sympathetic branch of the autonomic nervous system (ANS) gets activated. This sympathetic activation increases the readiness of the organism to come into action; it raises heart rate, increases respiration, raises blood pressure and decreases heart rate variability. Negative emotions such as anger and fear are associated with stronger increases in sympathetic activation than positive emotions such as happiness and contentment [2]. In addition to these stronger activations, the physiological consequences of negative emotions are longer lasting compared to positive emotions [2, 7]. Furthermore, it has been shown [3] that daily stress, worry and rumination lead to an increased heart rate (HR) and a decreased heart rate variability (HRV). Detecting changes in physiological measures (such as HR, HRV, respiration and blood pressure) can thus be indications of changes in that person's emotional wellbeing (in addition to being indicators of other problems

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 908–911, 2009.

such as coronary heart disease). This raises the possibility of identifying changes in emotional states through measuring differences in ANS activity. These physiological changes should then be more indicative of negative emotions than positive emotions due to their greater intensity and temporal extent. There has been some interesting work in HCI research on the use of non-intrusive emotion detection through measuring physiology (e.g.[1, 6, 10]). Our research differs from these studies in that our focus lies on detecting long term changes in emotional wellbeing instead of detecting short term instantaneous effects of emotional states.

A predominance of negative emotional states leads to a decrement in emotional wellbeing and even physical wellbeing. There is a strong link between mental health and physical health. Numerous studies have shown that psychological stress can lead to decreased immune response (e.g. [5]). Stress also lowers the number of cytotoxic lymphocytes called “natural killer cells” that protect the body not only against infectious diseases but also play an important role in the rejection of tumors. Therefore, monitoring a person's wellbeing and responding to threats to this wellbeing (such as sustained negative emotions) should not only be effective in enhancing that person's emotional wellness but also enhance that person's physical wellbeing, care dependency and ultimately life expectancy.

The standard method for measuring changes in a person’s emotional wellbeing is having them fill out questionnaires. In order to avoid this we investigate automated alternatives that measure physiological and/or behavioral changes that can be indicative of emotional change. The goal of our research is to provide non-intrusive automated measures of emotional wellbeing.

2 Measuring Nocturnal Physiology Using a Bed Sensor

Taking nocturnal measures of heart rate, respiration and activity provides a unique opportunity to gather data in a non-intrusive way over a long stretch of time. There are two additional advantages associated with the use of these physiological signals as a measure for emotional wellbeing. First, one would expect the general emotional state to be least contaminated by specific situational cues during sleep. The individual is essentially unresponsive to sensory stimulation (except perhaps the comfort of the bed, temperature and excessive noise). From the perspective of an individual's general emotional state this would thus provide an ideal measurement, because of the relative absence of sensory stimulation in the temporal proximity. The ANS activity would thus reflect a more general context independent measure of the individual's emotional state at that time. Relatedly, taking nocturnal measures increases the chance of distinguishing negative emotional states from positive emotional states because of their greater temporal extent.

As a means for taking physiological measures in a non-intrusive way we chose to investigate the use of a bed sensor. The bed sensor used in our research is the Emfit DVM (discrete vitals monitoring) sensor (as depicted in Figure 1) which is a thin film dynamic sensor, installed under the mattress. It responds to small pressure changes caused by a person’s ballistocardiographic (BCG), respiration movements and other bodily movement during sleep. The output of the sensor is a single voltage signal corresponding with all pressure changes caused by the a person lying on the mattress.

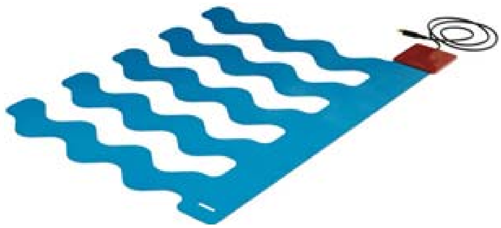


Fig. 1. Dicrete vitals monitoring sensor

The digital data acquisition and signal processing unit uses specialized algorithms to calculate heart and respiration rates and movement activity from the sensor signal. The sampling frequency of the sensor is 200 Hz.

Because the sensor is thin and capable of taking recordings of the above mentioned physiological traits through the mattress it is minimally intrusive. The sensor causes no noticeable difference for a person lying in bed (apart perhaps from being conscious that the sensor is there independent of differences in sensation).

3 Convergent Validation Study

In order to evaluate bed sensor measures of physiology as a valid means for probing the emotional wellbeing of an individual, there is a need for alternate measures of emotional states. Self report measures such as questionnaires or interviews are still the most authoritative means for taking these measures despite shortcomings such as the possibility of using response strategies, conforming to social desirability and the more fundamental problem of whether emotional states are sufficiently accessible to an individual consciousness (suppressed emotions are per definition not consciously accessible). The questionnaires that were used in our research as convergent validation measures are the "Positive And Negative Affect Schedule" (PANAS) [11] and the "Affect Grid" [9]. Both these questionnaires assume a two dimensional model of affect. However, where the PANAS assumes independence between positive and negative affect dimensions, the Affect Grid puts these on opposite ends of the *hedonism* dimension and adds *arousal* as the second dimension. Furthermore, we investigated the effects of Stress and Worry by adapting questionnaires used in previous research [8].

Correlations were predicted between the Heart Rate and Respiration measures on the one side and the Negative Affect (PANAS), Hedonism, Arousal, Worry and Stress on the other. Higher negative affect, lower hedonism, higher arousal, higher worry and stress are all hypothesized to be predictive of increases in Heart Rate and Respiration and a decrease in Heart Rate Variability (HRV).

4 Results and Conclusion

The initial results over three subjects (measured 23 days each) show that of the expected correlations between physiological measures provided by the bed sensor and the questionnaire measures only the correlations between Negative Affect and Respiration ($r = 0.41$, $p < 0.001$), between Stress and Respiration ($r = 0.29$, $p < 0.05$) and between Worry tension and Respiration ($r = 0.27$, $p < 0.05$) were significant. These correlations were calculated on the z-scores of the individual subjects as to compensate for individual differences in baseline physiological activity and style for filling out questionnaires. Nocturnal breathing frequency correlates positively with the reported negative affect, stress and tension associated with worry reported the night before.

At this point, we conclude that there are indications that physiological measures taken with a non-intrusive bed sensor correlate with emotional wellbeing; a prevalence of negative emotions, stress as well as tension associated with worry during the day, lead to significantly higher respiration rates during the night. This opens up possibilities for the automated non-intrusive measurement of emotional wellbeing.

Acknowledgement. This work was supported by the SenterNovem Point One project 'IPTV Portal to Wellness', part of the ITEA-2 project AmIE.

References

1. Anttonen, J., Surakka, V.: Emotions and heart rate while sitting on a chair. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 491–499. ACM, New York (2005)
2. Baumeister, R.F., Bratslavsky, E., Finkenauer, C., Vohs, K.D.: Bad is stronger than good. Rev. Gen. Psychol. 5, 323–370 (2001)
3. Brosschot, J.F., van Dijk, E., Thayer, J.F.: Daily worry is related to low heart rate variability during waking and the subsequent nocturnal sleep period. Int. J. Psychophysiol. 63, 39–47 (2007)
4. Cosmides, L., Tooby, J.: Evolutionary psychology and the emotions. In: Lewis, M., Haviland-Jones, J.M. (eds.) Handbook of Emotions, 2nd edn., pp. 91–115. Guilford, NY (2000)
5. Glaser, R., Rice, J., Sheridan, J., Fertel, R., Stout, J.C.: Stress-related immune suppression: Health implications. Brain Behav. Immun. 1, 7–20 (1987)
6. Horlings, R., Datcu, D., Rothkrantz, L.J.: Emotion recognition using brain activity. In: Proceedings of the 9th international Conference on Computer Systems and Technologies and Workshop For PhD Students in Computing, vol. 374. ACM, New York (2008)
7. Pieper, S., Brosschot, J.F.: Prolonged stress-related cardiovascular activation: is there any? Ann. Behav. Med. 30, 91–103 (2005)
8. Pieper, S., Brosschot, J.F., van der Leeden, R., Thayer, J.F.: Cardiac effects of momentary assessed worry episodes and stressful events. Psychosom. Med. 69, 901–909 (2007)
9. Russell, J.A., Weiss, A., Mendelsohn, G.A.: Affect Grid: A single-item scale of pleasure and arousal. J. Pers. Soc. Psychol. 57, 493–502 (1989)
10. Wang, H., Prendinger, H., Igarashi, T.: Communicating emotions in online chat using physiological sensors and animated text. In: CHI 2004 Extended Abstracts on Human Factors in Computing Systems, pp. 1171–1174. ACM, New York (2004)
11. Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect:The PANAS scales. J. Pers. Soc. Psychol. 54, 1063–1070 (1998)

Using a Dynamic Model to Simulate the Heuristic Evaluation of Usability

Nuria Hurtado[1], Mercedes Ruiz[1], and Jesús Torres[2]

[1] Department of Computer Languages and Systems
University of Cadiz
C/ Chile nº1 - 11003 – Cadiz (Spain)
{nuria.hurtado,mercedes.ruiz}@uca.es
[2] Department of Computer Languages and Systems
University of Seville
Avda. Reina Mercedes s/n - 41012 - Seville (Spain)
jtorres@lsi.us.es

Abstract. Among usability inspection methods, heuristic evaluation, or expert evaluation, is considered the most used and well-known usability evaluation method. The number of evaluators and their expertise are essential aspects that affect the quality of the evaluation, the cost that its application generates, and the time that it is necessary to spend. This paper presents a dynamic simulation model to analyze how different configurations of evaluator team have an effect upon the results of the heuristic evaluation method. One of the main advantages of using a dynamic simulation model is the possibility of trying out different decisions before carrying them out, and change them during the simulation of the evaluation process.

Keywords: Usability, usability evaluation, heuristic evaluation, dynamic simulation modeling.

1 Introduction

Heuristic evaluation or expert evaluation of usability is one of the most used and well-known usability evaluation methods. It was originally developed by Nielsen and Molich [6] and it is possible to find numerous references about this evaluation method and how to carry it out [1][4][9]. The number and expertise of evaluators are essential aspects that affect the evaluation results in a heuristic evaluation. Knowing both, the number and expertise, it is possible to estimate the number of problems that can be found using the mathematical model presented by [8]. These two factors also affect the costs that the method incurs. But there is an additional key factor: time.

Using dynamic modelling and simulation will also allow us to obtain final values of these key factors, but the additional advantage is that it allows us to have visibility over time during the application of usability evaluation. Thus, it is possible to experience different decisions before carrying them out, as well as to introduce changes in the key input parameters during the simulation process. The simulation model can

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 912–915, 2009.

provide information about the results of the method when combinations of evaluators with different levels of expertise and dedication are used, thus allowing us to customize the model to the reality of an organization or a specific project.

This paper presents a dynamic simulation model to analyze the effects that the changes in number, expertise and dedication of evaluators have on the key factors: the quality of the evaluation results, the cost, and the time spent.

2 Dynamic Simulation Model of Heuristic Evaluation

The scheme proposed by Kellner et al. [3] and the methodology for model building by Martinez and Richardson [5] have been followed to build this model:

1. Model Purpose and Model Scope: The purpose of the model determines key questions to address, as well as the model scope. The purpose of the developed model is to analyze how the changes in *number, expertise* and *dedication* of evaluators have an effect upon the behaviour of the method over time. This effect will be assessed through *time*, *cost* and *quality* indicators. The scope of this model, in this case, is limited to modelling the behaviour of the *problem detection* phase of the heuristic evaluation method, which results in a single set of detected usability problems.

2. Output variables and Input Parameters: The output variables are the information items that must be known in order to be able to answer the key questions for which the model was created. In the model developed these indicators will be represented through the evolution of the following main output variables over time:

Usability problems found: This one represents the number of different usability problems found after analysis and aggregation of independent evaluations. It will hence determine evaluation *quality*.

Accumulated variable cost: This one represents the variable *cost* implied by implementing the method.

Time: The behaviour of *time* will be influenced by the dynamic nature of the simulation model.

The main input parameters of the model will make it possible to configure different simulation scenario:

Number of evaluators E(i): this parameter is subdivided into three other parameters depending on the level of *expertise*, making it possible, as well, to change their time *dedication*, hence variation in evaluator team composition is allowed for during simulation. *E1*: It represents the number of novice evaluators (level 1). *E2*: It represents the number of evaluators with expertise in usability (level 2). *E3*: It represents the number of evaluators with expertise in usability as well as in the mastering of the specific system (level 3).

Usability problems to be found: This one represents the total number of estimated usability problems depending on the size and type of system to be evaluated.

Single evaluation capacity E(i): This parameter is also subdivided into E1, E2, and E3 and it will represent the amount of usability problems that an evaluator can sort out per time unit (hour).

Single cost E(i): This parameter is also subdivided into E1, E2, and E3 and it represents the cost (in Euro) per time unit (hour) of an evaluator.

Single problems proportion E(i): This parameter is also subdivided into E1, E2, and E3 and it represents the average rate of problems that a single evaluator may be able to find. It may be estimated using Nielsen´s mathematical model [8] or obtained from other similar studies, or from previous project metrics collections.

3. Model Conceptualization and Model Formulation: The concepts involved in a system dynamics model are normally represented in a causal loop diagram. In this step we have used casual loop diagrams to collect the cause and effect relationships found in the heuristic evaluation method. Once the model has been conceptualized, its formalization as a mathematical model follows. We have used stock and flow diagrams [2] in this step since they help us to represent the structure of the model and provide a bridge to simulation modelling by facilitating the assignment of equations to this structure and the identification of relationships between the variables involved. These diagrams are made up by three main elements: level variables, flow variables, and auxiliary variables. In the model developed three level variables have been considered: *usability problems found*, *accumulated variable cost* and *usability problems to be found.* Two flow variables have been added in the model: *problem detection rate* and *cost increase rate*. The first one represents the flow between *usability problems to be found* and *usability problems found*. The second one increases the level variable *accumulated variable cost* as the simulation progresses. Auxiliary variables are the rest of elements in the process that have an influence upon it. It is also necessary to have constants or input parameters that represent external influences on the model. Once the model has been developed and behaviour-governing equations have been introduced, simulation can be carried out by setting different scenarios.

4. Model Assessment: There are two main aspects that must be taken into account to carry out model assessment: model verification and model validation [10]. Model verification assesses that the model implementation is error free and, that it is a right representation of the intended logical behaviour of the system under study. Model validation assesses that the model helps solve the end user's problems within the context of study. We have carried out different tests to validate and verify the structure and behaviour of the model developed with satisfactory results.

5. Model Simulation and Analysis of Results: The model has been implemented using the Vensim® simulation environment. First of all, to simulate the model we have set the values for the main input parameters: *usability problems to be found, problem single proportion, single capacity, and single variable cost*. Data from various projects published in the related literature have been used to set up the base case of our study [7][9]. Then, the effect of the evaluators can be studied by setting several experiments with different scenarios, each one having different values for the input parameter: *number of evaluators E(i)*. For instance, it is possible to compare different scenarios in a simulation using different number of evaluators of the same level of expertise and dedication. In other instances, the simulation model can also provide information about the behaviour of the method when combinations of evaluators with different levels of expertise or dedication are used. The changes in the input parameters allow us to customize the model to the reality of an organization or a specific project. The simulation results of the different experiments carried out is represented in graphs of evolution over time for the result variables. The analysis of these graphs allows us to help decision making in the usability evaluation.

3 Conclusions

This work illustrates the application of the modelling and simulation techniques to the heuristic evaluation method. The model has been conceptualized and formalized under the system dynamics approach. The implementation of the simulation model makes it possible to add visibility over time during the application of usability evaluation. It allows us to analyse the effect that changes in the number, expertise and dedication of evaluators have upon the results of cost, time and quality, making it possible to try out different policies to manage usability evaluation. The model provides a tool to help decision making for any organization that intends to implement usability evaluation methods. Our next tasks will be aimed at the development of simulation models of other usability evaluation methods.

Acknowledgements. The authors wish to acknowledge the Interdepartmental Commission for Science and Technology and the European Regional Development Fund (Spain), for subsidizing this research through projects TIN2007-67843-C06-03 and TIN2007-67843-C06-04.

References

1. Daly-Jones, O., Bevan, N., Thomas, C.: Handbook of User Centered Design. IE2016 INUSE Deliverable D6.2.1 Version 1.31. Serco Usability Services, National Physical Laboratory, Teddington, Middx, UK (2001)
2. Forrester, J.W.: World Dynamics. Productivity Press, Cambridge (1973)
3. Kellner, M.I., Madachy, R.J., Raffo, D.M.: Software Process Simulation Modeling: Why? What? How? The Journal of Systems and Software 46(2), 91–105 (1999)
4. Lindgaard, G.: Usability Testing and System Evaluation. Chapman & Hall, London (1994)
5. Martinez, I.J., Richardson, G.P.: Best Practices in System Dynamics Modeling. In: Hines, H., Diker, V.G. (eds.) Plenary paper: Proceedings of the 29th International Conference of the System Dynamics Society, Atlanta, GA, USA (2001)
6. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People, Seattle, Washington, United States, pp. 249–256 (1990)
7. Nielsen, J.: Finding Usability Problems through Heuristic Evaluation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM CHI 1992, Monterey, CA, USA, pp. 373–380 (1992)
8. Nielsen, J., Landauer, T.K.: A Mathematical Model of the Finding of Usability problems. In: Proceedings of the ACM INTERCHI 1993 Conference, Amsterdam, The Netherlands, pp. 206–213 (1993)
9. Nielsen, J.: Heuristic Evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods, pp. 25–62. John Wiley & Sons, New York (1994)
10. Sterman, J.D.: Business Dynamics. In: Systems Thinking and Modeling for a Complex World. McGraw-Hill, New York (2000)

Using Avatars for Improving Speaker Identification in Captioning

Quoc V. Vy and Deborah I. Fels

Ryerson University, 350 Victoria Street, Toronto, Ontario M5B 2K3, Canada
{qvy,dfels}@ryerson.ca

Abstract. Captioning is the main method for accessing television and film content by people who are deaf or hard-of-hearing. One major difficulty consistently identified by the community is that of knowing who is speaking particularly for an off screen narrator. A captioning system was created using a participatory design method to improve speaker identification. The final prototype contained avatars and a coloured border for identifying specific speakers. Evaluation results were very positive; however participants also wanted to customize various components such as caption and avatar location.

Keywords: inclusive design, speaker identification, avatars, captioning.

1 Introduction and Background

Film and television represent a significant method of cultural distribution. The ability to access such content is important for all people. Captioning is the main method used by people who are deaf or hard-of-hearing to access this content. It is the verbatim transcription of audio content using text descriptions and symbols. The most popular form of captioning is on television called Closed Captioning (CC) in North America or Subtitling in Europe. In North America, Line 21 CC appears as white mono-spaced text on a black background and is usually located near the bottom of screen. Text may be either as all uppercase or, more recently, as mixed-case lettering. The most common method for speaker identification in CC is using text descriptions consisting of two chevrons, speaker's name or function, followed by a colon (e.g., >>ANNE:).

Text descriptions are ineffective as they require prior knowledge (e.g., names of characters, especially when off-screen) and additional cognitive effort to associate a name and other visual indicators (e.g., lips moving) to the speaker. Among the few studies that have been carried out to address this issue and other non-speech information (NSI) are [1] and [2]. Harkins et al. [1] recommended using explicit descriptions for NSI and King [2] found that colour use for speaker identification did not improve much when compared to placement of captioning near the speaker.

In this paper, a graphical solution to speaker identification for captioning is described along with user comments and reactions from a *formative pilot study*. Four participants (two deaf, two hard of hearing) were used for prototype development and evaluations. Twenty casual hearing observers were added during final prototype evaluation for additional comments and feedback.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 916–919, 2009.

2 Needs and Requirements Analysis

The proposed system was developed using a participatory design (PD) method and involved individuals who were deaf, hard-of-hearing or hearing. We used an activity and mapping technique [3] to gather and understand user's setting and needs. The activity consisted of asking two deaf participants (one male, one female) to watch a favourite television show with CC at home.

A Diagnostic Mapping was produced from the activity and a follow-up interview revealed some common themes. For example, both users complained that CC was missing non-speech information, such as speaker information. Another common theme was a preference for italicized text and brackets to indicate narration and sound effects. A Virtual Mapping was then created to find some possible solutions to the issues identified. For example, using images and symbols to indicate different NSI and using the black bars, found on standard 4:3 screens with widescreen 16:9 content, to accommodate the additional information. Results from mappings indicated that deaf users rely heavily on captioning to be accurate and of sufficient quality to represent or indicate audio information that they would otherwise unable to obtain.

The proposed system is not designed to be used with existing captioning technology found on television as there are some technical limitations. It would be better applied in a digital implementation such as on a computer or the Internet.

3 Prototype Development and Iterations

A paper-based prototype was drawn using a "pencil before pixel" [4] design from a crude mock-up of the system using sticky notes on a screen. Screenshots or "avatars" of characters were placed adjacent to the captioning to visually identify the speaker, together called a "captioning panel". According to Law of Proximity [5], placement of the captioning panel was associated with the location of characters on screen to further aid identification of speakers. Some reactions from deaf participants were that they were excited and thought that this design was "different", "great", and "helpful". Furthermore, they liked the use of avatars and found that avatars helped indicate "who was talking" and improve "their understanding" of content.

The next step was to create image-based prototypes using actual content from a particular movie, in this case Transformers [6]. In this iteration, graphical and coloured elements were introduced to provide redundancy to further distinguish between speakers. For example, a coloured border matching the character's primary wardrobe colour surrounded its corresponding avatar. Both deaf participants liked the coloured border and thought it assisted in their ability to identify the correct speaker.

Although both participants were positive about the avatar and colour border, there was also considerable divergence in deaf participants' expressed needs and preferences. For example, participants wanted to place the avatars and there respective captions in different top/bottom locations and left/right order. As a result, user preferences were implemented which allowed viewers to change the location of the captioning panels, the order of avatars and captions, the size of avatars and text used for captioning, and the transparency of caption background.

4 Initial Evaluation and Discussion of Final Prototype

For the final prototype evaluation, we wanted to gain some additional perspectives from the hard-of-hearing and hearing communities. Two hard-of-hearing (HOH) participants were added to the participant pool and 20 hearing individuals (twelve females and eight males) acted as casual observers. All participants (deaf and HOH) and hearing observers were shown a 4:45 minute video clip of the Transformers movie with the final prototype (see Figure 1).

In Figure 1, four characters are shown on the screen, but only two of them are speaking. The captioning panel is located at the bottom of the screen and depicts the avatars of characters who are speaking, along with their names, a coloured border matching their respective wardrobes, and their corresponding dialogue located to the right of the avatar. The position of the dialogue and avatar is relative to the position of that character on the screen (e.g., Samuel is located on the left side of the screen while Optimus is on the right and mostly off-screen).

All participants thought that using avatars to represent who is speaking and having a multi-level display for captioning was innovative and an improvement over existing text-based methods. They also liked having the name of speaker below the avatar, as well as the use of italicized text and brackets for indicating sounds effects and narration. Participants liked having avatars regardless of ordering, but they preferred that the avatars be located on the left-side of the corresponding captions.

There were a variety of comments from deaf and HOH groups and some unexpected differences. In particular, some graphical elements were helpful for some, but not others. For example, the coloured border was helpful for deaf participants, but not for some of the HOH participants or casual observers. While studies by [2], [7] and [5] found that using colour for speaker identification was not helpful, this study shows that it may be beneficial and desirable by deaf users. Reasons for not liking the coloured border may be due to the use of other methods for identifying speakers such as the image or "avatar" and location of captions.

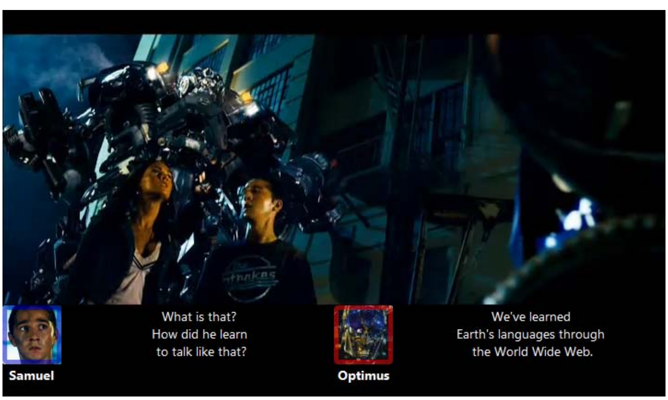


Fig. 1. Screenshot depicting avatars, coloured borders and positioning for speaker identification. Text captions represent the dialogue occurring in this scene.

In this prototype, captions were maintained on screen as long as possible to maximize reading time. However, this caused some difficulty for some deaf participants as the captioning did not always match the onscreen visuals such as lips movements. All participants did not know where to look and were unable to follow along easily. A common suggestion was to highlight the current dialogue similar to karaoke. Highlighting word per word as in karaoke could interfere with reading efficiency as it forces people to read at that particular rate. A more effective implementation might be highlighting the entire captioning instead of individual words. Nonetheless, further research regarding optimizing reading time with speaking time is required.

Some participants from deaf and HOH groups found that they were overwhelmed with the amount of information available on screen. They initially found the screen “too busy” and the captions to appeared “too fast” making it difficult to read. This was caused by the overlapping of multiple dialogues being displayed simultaneously and for the extended duration. As a result, this increased the information that viewers had to absorb in a short time. However, after subsequent views participants were no longer "overwhelmed" as much. It seems that participants were too accustom to conventional CC that they required time to learn this new system and overcome automatic behaviours and expectations for reading the existing style of CC.

Further research and formal user studies are required to determine the effects on perceptual load, the readability of captions, the ability to see and understand video content, and the enjoyment levels. The ability to change various sizes, order and locations of avatars and captioning is a good start in finding an optimal and improved method of access to cultural content for people who are deaf or hard of hearing.

Acknowledgements. Funding for this research was provided by a SSHRC Community-University Research grant. We gratefully acknowledge the participants in the study who provided their time and effort throughout this research. Finally, we thank Dr. Andrew Clement and Dr. Eric Harley for their advice and support.

References

1. Harkins, J.E., Korres, E., Singer, B.R., Virvan, B.M.: Non-speech information in captioned video: A consumer opinion study with guidelines for the captioning industry. Gallaudet Research Institute, Washington (1995)
2. King, C., Lasasso, C., Short, D.: Digital Captioning: Effects of Color-Coding and Placement in Synchronized Text-Audio Presentations. In: ED-MEDIA 1994, pp. 329–334 (1994)
3. Keld Bødker, F.K.: Participatory IT design: Designing for business and workplace realities. MIT Press, Cambridge (2004)
4. Baskinger, M.: Pencils before pixels: a primer in hand-generated sketching. Interactions 15(2), 28–36 (2008)
5. Quinlan, P.T., Wilton, R.N.: Grouping by proximity or similarity? Competition between the Gestalt principles in vision. Perception 27(4), 417–430 (1998)
6. Spielberg, S. (Producer), Bay, M. (Director): Transformers [Motion Picture]. United States of America: DreamWorks (2007)
7. Rashid, R., Vy, Q., Hunt, R., Fels, D.I.: Dancing with words. International Journal of Human-Computer Interaction 24(5), 505–519 (2008)

Biometrics in Practice: What Does HCI Have to Say?

Lynne Coventry[1], Graham I. Johnson[2], Tom McEwan[3], and Chris Riley[2]

[1] Johnston and Johnston Ltd.
[2] NCR Labs, Dundee
christopher.riley@ncr.com
[3] Edinburgh Napier University

Abstract. This panel discusses biometric technologies from different perspectives in HCI in order to arrive at a coherent response that the community can give in this area. The challenges of doing this will also be of interest in terms of HCI's general influence on topical issues.

Keywords: Biometrics; public influence of HCI.

1 Introduction

Biometrics pose a challenge to HCI, in that they expose many of the contradictions inherent in defining our discipline. For at least ten years our community has debated a range of issues – from the minutiae of the capture technology itself to the wider scale of socio-technical analysis – yet the HCI community finds it difficult to communicate best practice to the outside world. In part this is because of the different lenses through which a topical and provocative subject can be viewed – cultural context of use, accuracy of authentication, interaction design, user habituation, human-centeredness v mandatory use. This panel debates what the HCI community can offer to policy makers, manufacturers and deplorers of biometric systems and in so-doing sheds light on how the HCI community can be more influential in topical debates. The panellists represent a range of stakeholders: Lynne Coventry has many years experience evaluating biometrics and usability, Karen Renaud has experience working with numerous different authentication systems. Tom McEwan researches human-centred innovation and seeks to help the UK's HCI community give a coherent response to topical issues such as biometrics and Chris Riley is well placed to represent the views of systems integrators and deplorers of biometrics. Finally, Linda Sorensen can give a perspective from NPL, which has advised on biometrics for many years. Each panellist will speak on their recent practical experiences with biometrics, contrast this with HCI theory and address the following questions:

- What are the characteristics of an acceptable biometric authentication system?
- Taking biometrics as an example, has HCI anything to offer beyond the traditional domains of systems and interface design?
- Are there areas that the HCI community regard as outside the field, e.g. the creation of information media designed to make biometrics acceptable?

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 920–921, 2009.

- What ethical contradictions are there in using human-centred methodologies to design a mandatory biometric system?

Biometric authentication is the process of identifying an individual through measurable characteristics of their behaviour, anatomy or physiology and is becoming increasingly pervasive in modern society. However, biometrics are a contentious form of personal identification, described as an ethically challenging technology which undermines personal privacy [1]. When studied, people voice concern about the technology or the way the way it will be used [2, 3], and with some justification. The use of biometrics automates the identification process, removing control from users, eliminating interpersonal trust. The characteristics used during biometric authentication, such as fingerprints or iris patterns, cannot simply be re-issued if they become compromised.

The HCI community has increasingly turned its attention to the usability and acceptability of biometrics and security systems, for example in the design of the biometric user interface [4], the overall biometric system [5] and the context within which biometrics are used [6]. A recent workshop on the usability of biometric systems [7] is further evidence of an increased focus on biometrics from the HCI community. Given the prevalence and impact of public data security breaches, making biometrics acceptable would seem to be a hard sell. Should the HCI community help with this sell? Can human-centred methods help design with mandatory biometric systems that appear to have no direct benefit for those required to use them (for example the US-VISIT programme and the proposed UK Identity Card scheme). Is user is only really at the centre of the design process, when speed and accuracy of the system are primary design goals?

References

1. Alterman, A.: "A piece of yourself": Ethical issues in biometric identification. Ethics and Information Technology 5, 139–150 (2003)
2. BioSec Consortium.: Report on results of first phase usability testing and guidelines for developers (2004)
3. Riley, C., Buckner, K., Johnson, G.I., Benyon, D.: Culture & Biometrics: Regional differences in the perception of biometric authentication technologies. Journal of AI & Society (2009)
4. Theofanos, M., Michaels, R., Scholtz, J., Morse, E., May, P.: Does Habituation Affect Fingerprint Quality? Paper presented at the 24th SIGCHI Conference of Human Factors in Computing Systems, Canada (2006)
5. Briggs, P., Oliver, P.: Biometric daemons: authentication via electronic pets. Paper presented at the 26th SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy (2008)
6. Heckle, R., Patrick, A., Ozok, A.: Perception and Acceptance of Fingerprint Biometric Technology. Paper presented at the Symposium on Usable Privacy and Security (SOUPS), Pittsburgh (2007)
7. Herman, M.: Objectives and Outcomes. In: The NIST International Workshop on Usability and Biometrics, Washington, June 23-24 (2008)

Demarcating User eXperience

Virpi Roto[1,2]

[1] Tampere University of Technology, Human-Centered Technology, Korkeakoulunkatu 6, 33720 Tampere, Finland
[2] Nokia Research Center, P.O.Box 407, 00045 Nokia Group, Finland
virpi.roto@nokia.com

Abstract. This panel discusses the scoping of user experience as a research field. User experience is a crossing point of several disciplines, each of which tends to define user experience from their own perspective. The distinguished panelists from academia and industry represent the different perspectives to user experience: Traditional human-computer interaction, Psychology, Cognitive psychology, and Design. The goal of the panel is to get one step closer to a shared understanding of the concept of user experience.

1 Goal

The concept of user experience (UX) is widely used but understood in many different ways. The multidisciplinary nature of UX has provoked several definitions and perspectives to UX, each approaching the concept from a different viewpoint. UX is seen as a holistic concept covering all aspects of experiencing a phenomenon, but we are facing the point where UX becomes a too broad concept to be useful in practice. Practitioners have difficulties to understand the concept and to improve UX in their work, and researchers rather use some other term to make their research scope clear. Instead of dismissing the whole term of user experience, we believe UX can be defined to cover a certain set of cases and to make its scope clear.

There have been several activities that aim to get closer to a shared understanding of the vague concept of UX, and this panel builds on those activities. Two workshops [1,2], a Special Interest Group session [3], and an online survey [4] have been conducted to collect information on how UX professionals understand UX. In this panel, we aim to proceed from the phase of gathering broad understanding to the phase of narrowing down the concept of UX. We start to demarcate the field of UX.

2 Topics

The panelists first have a few minutes for presenting their position and perspective to UX. They are guided to prepare for the most intriguing and widely interesting questions about UX demarcation, for example, the ones below. The audience is encouraged to pose questions to panelists as well.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 922–923, 2009.

Potential topics for discussion:

- If a university provides a study program on user experience, what do you think it should be about?
- If I am searching for a User Experience Manager to work at Nokia, what do you think the work will be about?
- What is the difference between UX and experience in general?
- What is the difference between UX and usability?
- How should we evaluate UX (from your perspective to UX)?

3 Panelists

The prestigious group of five panelists from academia and industry represent the different perspectives to user experience as follows.

1. Nigel Bevan, Professional Usability Services, United Kingdom Perspective: Usability, ISO definition for UX (http://www.nigelbevan.com)
2. Jettie Hoonhout, Senior Scientist, Philips, Netherlands Perspective: Industry needs (http://www.linkedin.com/in/jettiehoonhout)
3. Kristina Höök, Professor, Stockholm University, Sweden Perspective: Affective interaction (www.sics.se/~kia)
4. Ilpo Koskinen, Professor, University of Arts and Design, Finland Perspective: Design, Co-experience (http://www2.uiah.fi/~ikoskine/)
5. Gitte Lindgaard, Professor, Carleton University Ottawa, Canada Perspective: Hedonic+pragmatic UX (http://www.carleton.ca/~glindgaa)

Panel moderator: Virpi Roto, Visiting Researcher at Technical University of Tampere and Principal Scientist at Nokia Research Center, Finland.
(http://research.nokia.com/people/virpi_roto)

References

1. Law, E.L.-C., Hvannberg, E.T., Hassenzahl, M.: Proceedings of the workshop on Towards a Unified View of UX, in conjunction with NordiCHI 2006, Oslo, Norway, October 14 (2006)
2. Law, E.L.-C., Vermeeren, A., Hassenzahl, M., Blythe, M. (eds.): Proceedings of the workshop on Towards a UX Manifesto, in conjunction with HCI Conference, Lancaster, UK, September 2 (2007)
3. Law, E.L.-C., Roto, V., Vermeeren, A.P.O.S., Kort, J., Hassenzahl, M.: SIG on Towards a shared definition of user experience. In: Proceedings of ACM CHI 2008, Florence, Italy, April 2008, pp. 2395–2398 (2008)
4. Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J.: Understanding, Scoping and Defining User eXperience: A Survey Approach. In: Proceedings of ACM CHI 2009, Boston, MA, USA (April 2009)

Mobility, Emotion, and Universality in Future Collaboration

Mark Chignell[1] , Naotsune Hosono[2], Deborah Fels[3], Danielle Lottridge[1], and John Waterworth[4]

[1] Dept. of Mechanical & Industrial Engineering, University of Toronto
[2] Oki Consulting Solutions, Tokyo
[3]School of Information Technology Management, Ryerson University
[4] Department of Informatics, Umea University

Abstract. The Graphical user interface has traditionally supported personal productivity, efficiency, and usability. With computer supported cooperative work, the focus has been on typical people, doing typical work in a highly rational model of interaction. Recent trends towards mobility, and emotional and universal design are extending the user interface paradigm beyond the routine. As computing moves into the hand and away from the desktop, there is a greater need for dealing with emotions and distractions. Busy and distracted people represent a new kind of disability, but one that will be increasingly prevalent. In this panel we examine the current state of the art, and prospects for future collaboration in non-normative computing requirements. This panel draws together researchers who are studying the problems of mobility, emotion and universality. The goal of the panel is to discuss how progress in these areas will change the nature of future collaboration.

Keywords: Future Interfaces; Mobile Computing; Perceptual Interfaces; Emotional Design; University Interfaces; Cross-Cultural Interfaces.

1 Panel Description

The future belongs to networked computing with multiple computing devices per person each being used in different ways and contexts. As devices become more flexible and more varied, there are increasing prospects for accommodating the needs of a wide range of people in many different contexts. In this panel we seek to explore the prospects for future online collaboration focusing on the issues of mobility, emotion, and universality. Our aim is to encourage debate concerning the common principles that emerge when we link non-normative forms of interaction into future collaborative interfaces.

The topics discussed in this panel are shown below.

- Mobile Devices (Chignell). Chignell will argue that mobile devices are likely to be the dominant interfaces for future collaboration. He will also discuss the disabilities implied by mobile interaction and the design principles that

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 924–925, 2009.

emerge in this context. In making this argument he will discuss a series of research projects carried out at the University of Toronto that have examined communication and collaboration within healthcare teams and between medical workers and patients using handheld devices.

- Perceptual Interfaces (Waterworth). Waterworth has a particular interest in the extent to which people feel *present* within a medium when interacting with or through a device. The ultimate in mediated presence has been thought of as total immersion in a virtual world - with the assumption that a virtual world is an alternative to the physical world, and in competition with it. This is no longer a viable view. In his view, the future won't be so much about interaction in either mediated *or* physical reality, but about being in unconstrained situations in which both are blended and presence levels are adjusted dynamically and automatically. For this to work, media devices will need to be sensitive to both the situational context of their use, and the state of their users.
- Emotional Design (Lottridge). Ms. Lottridge will discuss her Ph.d work at the University of Toronto on assessment of emotions during interaction as a means to bring the construct of emotional interface design into practical use. In addition to reviewing emotional assessment methods she will also review psychometric constructs such as emotional bandwidth, as a way to quantify the ability to experience and understand emotion. Varying emotional bandwidth capacities, and individual differences in emotional reporting, reveal disabilities or inequalities between people that should be measured in order to be accommodated. The broader implications for the use of emotion in interface design will also be discussed.
- Universal Interfaces (Fels). Dr. Fels will discuss how future interfaces can be designed to be more inclusive, and accommodating, with respect to the requirements of people with a range of abilities and disabilities. She will discuss the principles of universal design and how they can be applied in electronic media. Examples of inclusive media and sign language-based interfaces will be demonstrated.
- Universal and cross-cultural Interfaces (Hosono). The WWC (WorkWel Communicator) system has been developed by Dr. Hosono's group to address problems such as social isolation and reduced options for synchronous communication facing physically challenged persons in their efforts to fulfill their work duties. WWC is a multiple site voice system offering a realistic meeting experience for teleworkers with physical disabilities, while having a simplified interface that requires less effort to use (including simple conference control using a ten-key pad and voice out for dysphonea). This presentation will discuss how lessons learned in developing WWC can be applied to future collaboration systems for people with disabilities.

Designing Interaction for Next Generation Personal Computing

Giorgio De Michelis[1,2], Marco Loregian[1,2], Claudio Moderini[1,3], Patrizia Marti[4], Cesare Colombo[5], Liam Bannon[6], Cristiano Storni[6], and Marco Susani[7]

[1] ITSME srl, viale Sarca 336-F, 20126 Milano, Italy
{giorgio.demichelis,marco.loregian,claudio.moderini}@itsme.it
[2] University of Milano-Bicocca, viale Sarca 336-F, 20126 Milano, Italy
{gdemich,loregian}@disco.unimib.it
[3] Domus Academy, Via Watt 27, 20143 Milano, Italy
claudio.moderini@domusacademy.it
[4] Communication Science Dept., University of Siena, Via Roma 56, 53100 Siena, Italy
marti@unisi.it
[5] CEFRIEL, Politecnico of Milano, Via Fucini 2, 20133 Milano, Italy (Italy)
colombo@cefriel.it
[6] University of Limerick, Ireland
{liam.bannon,cristiano.storni}@ul.ie
[7] Motorola Mobile Devices, Chicago, USA
marco.susani@motorola.com

Abstract. Over two decades of research in the field of Interaction Design and Computer Supported Cooperative Work convinced us that the current design of workstations no longer fits users' needs. It is time to design new personal computers based on metaphors alternative to the desktop one. With this SIG, we are seeking to involve international HCI professionals into the challenges of designing products that are radically new and tackling the many different issues of modern knowledge workers. We would like to engage a wider cross-section of the community: our focus will be on issues of development and participation and the impact of different values in our work.

1 Introduction

With the emergence of new interaction paradigms, such as cloud, mobile and ubiquitous computing, and the evolution of personal computing on the Web, e.g., due to the success of Social Computing (Web 2.0) and to the evolution of Semantic Web technologies, it is now clear that *the desktop metaphor is obsolete*. Actually, research literature sustaining this claim is vast and nearly as old as the metaphor itself – which was developed by Alan Kay and colleagues at the Xerox PARC in the late '70s and reached the market in 1984 with the Apple Macintosh.

Researchers in the various fields of computer science (computer supported cooperative work, ubiquitous computing, knowledge management, cooperative information systems, etc.) and design (interaction design, human computer interaction, etc.) have tackled the problem of finding better ways to support users in interacting with

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 926 – 927, 2009.

computers from many perspectives, bringing forth new ways to look at work practices, better understanding of the relations between language and action, and a closer attention to knowledge sharing within collaborative teams and communities of practice. The focus has been for long time on the shared services that could be provided to people: only in the last years, researchers and practitioners started to pay attention to the way personal workstations could be redesigned in order to go beyond the desktop metaphor and become better tools to help users to manage the complexity of their manifold activities.

This SIG aims at fostering a cross-cultural discussion on how to design interaction for next-generation personal computing, especially taking into consideration the aspects of situatedness of actions.

2 Attendance and Format of Discussions

This SIG is for two groups of attendees – those with some experience of working in this field, e.g., who already tried to develop innovative interactive systems, who want to compare and share practice, and others with little or no experience, who are interested in learning more and possibly starting new projects and finding collaborations.

The SIG will start from a brief introduction of the *itsme* project[1], in which most of the organizers are involved, and on the motivations and features of the design approach. The metaphor of *stories and venues* will be introduced and analyzed as a case of alternative to the desktop metaphor. As already happened in previous seminars and tutorials, this introduction will start a discussion on issues such as:

- The role of the Web in the design of personal computing for the future, with particular attention to social and semantic aspects;
- The role between research and industry in developing anr releasing innovative products on the market.

The duration of a SIG will not allow for a very detailed sharing or debate on personal experiences; therefore, we see this as an important beginning to ongoing process. We plan to continue the conversation online, actively involving the participants in the *itsme* community, in which the issues of open innovation are of foremost importance.

[1] http://www.itsme.it

Postgraduate Studies in the Field of HCI

Teija Vainio[1], Veikko Surakka[2], Roope Raisamo[2], Kari-Jouko Räihä[2], Poika Isokoski[2], Kaisa Väänänen-Vainio-Mattila[1], and Sari Kujala[1]

[1] Unit of Human-Centered Technology, Tampere University of Technology, Korkeakoulunkatu 6, 33101 Tampere, Finland
{teija.vainio,kaisa.vaananen-vainio-mattila, sari.kujala}@tut.fi
[2] Tampere Unit for Computer Human Interaction
{veikko.surakka,rr,kjr,poika}@cs.uta.fi

Keywords: Education, HCI, postgraduate studies.

1 Introduction

In September of 2007, the Tampere Unit for Computer Human Interaction (TAUCHI) at the University of Tampere and The Unit of Human-Centered Technology (IHTE) at the Tampere University of Technology initiated a joint effort to increase collaboration in the field of human-technology interaction (HTI). One of the main aims was to develop higher quality education for university students and to carry out joint internationally recognized HTI research. Both research units have their own master and postgraduate students while the focus of education is at IHTE on usability and human-centered design of interactive products and services whereas TAUCHI focuses on human-technology interaction developing it by harmonizing the potential of technology with human abilities, needs, and limitations. Based on our joint analysis we know now that together TAUCHI and IHTE are offering an internationally competitive master's program consisting of more than 40 basic, intermediate and advanced level courses. Although both units are partners in the national Graduate School in User-Centered Information Technology (UCIT) led by TAUCHI we have recognized a clear need for developing and systematizing our doctoral education.

One of the main aims in our collaboration in the field of education is to create a doctoral program that meets internationally recognized needs for education. On the other hand the aim is to provide more extensive and at the same time more systematized postgraduate studies for both national and international students in the area of HTI. At the same time we hope to make our education internationally recognized and provide possibilities for our own students to include international fertilization and studies in their university degrees. Finding reasonable educational practices can be challenging. This results, for example, from the cases that different educational institutions have developed a variety of conventions in their curricula. In addition, different sources of funding can create challenges for collaborative efforts. Our basic and profound notion is that a need for developing high-quality researcher skills is a must in the multidisciplinary field of HTI. Therefore, we would like to invite conference

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 928–929, 2009.

participants to share and discuss postgraduate HTI studies in order to identify future needs and develop educational approaches that satify these needs. By doing this we also look future possibilities to collaborate internationally.

2 Interest and Relevance of the SIG to INTERACT 2009 and Assumed Attendee Background

We are aiming to collect ideas and opinions to such questions as, has HTI contributed to improved IT use in practice, are existing methods and tools for the design of usable systems sufficient, and are they well connected to the kind of HTI education that is required in the future.

The purpose of this SIG is to share and discuss experiences of designing international postgraduate studies and to provide an open forum to discuss topics of interests. Furthermore, the aim is to identify future possibilities and challenges that should be addressed by the HCI community.

The following topics are suggested to be discussed:

- What are the main challenges of designing international postgraduate HTI studies?
- What kind of education and courses a HTI doctoral student would need in order to have a high quality doctoral education. Please think both research and business skills.
  - Do postgraduate HTI studies differ from other areas of education?
- What are the concrete opportunities for collaboration?

We invite all interested conference participants to attend this SIG. Particularly those conference participants,

- who are educators,
- who are instructors in education, and
- who are or plan to be postgraduate HTI students may find this SIG interesting.

3 Format of Discussion

The first part of the SIG will focus on developing a vision of structured doctoral education. The idea is to have brief introductions to the first two key topics followed by informal discussions and groupwork of existing curricula and experiences of participants (about 30 minutes each). The second part will investigate needs and possibilities for collaboration.

At the conclusion of this SIG, we hope to be able to answer our key questions. In addition, we plan to submit a summary of the SIG's conclusions to SIGCHI Bulletin or Interactions.

Advanced Perceptual User Interfaces: Applications for Disabled and Elderly People

Francisco J. Perales López

Computer Graphics, Vision and Antificial Inteligence Group,
Mathematics and Computer Science Department, UIB
Crta. Valldemossa Km 7.5,
07122 Palma de Mallorca, Spain
Paco.perales@uib.es

Abstract. The research of new human-computer interfaces has become a growing field in computer science, which aims to attain the development of more natural, intuitive, unobtrusive and efficient interfaces. This objective has come up with the concept of Perceptual User Interfaces (PUIs) that are turning out to be very popular as they seek to make the user interface more natural and compelling by taking advantage of the ways in which people naturally interact with each other and with the world. PUIs can use speech and sound recognition and generation, computer vision, graphical animation and visualization, language understanding, touch-based sensing and feedback (haptics), learning, user modeling and dialog management.

Keywords: human-computer interfaces, computer vision thecniques, natural interaction, perceptual user interfaces.

1 Introduction

These new interfaces can be used in different scenarios (cars, domotic houses…), but a more important issue, are the systems' potential users. PUIs offer assistive technology for people with physical disabilities, which can help them to lead more independent lives and to any kind of audience they contribute to new and more powerful interaction experiences.

Of all the communication channels through where interface information can travel, computer vision provides a lot of information that can be used for detection and recognition of human's actions and gestures, which can be analyzed and applied to interaction purposes.

When sitting in front of a computer and with the use of webcams, very common devices nowadays, heads and faces can be presumed to be visible. Therefore, system's based in head or face feature detection and tracking, and face gesture or expression recognition can become very effective human-computer interfaces. Of course, difficulties can arise from in-plane (tilted head, upside down) and out-of-plane (frontal view, side view) rotations of the head, facial hair, glasses, lighting variations and cluttered background Besides, when using standard USB webcams, the provided image resolution is very poor and it has to be taken in account.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 930–931, 2009.

Different approaches have been used for non invasive face/head-based interfaces. For the control of the position some systems analyze facial cues such as color distributions, head geometry or motion. Another works track facial features or gaze including infrared lighting. To simulate the user's events it is possible to use facial gesture recognition. In this paper we consider as facial gestures the atomic facial feature motions such as eye blinking, winks or mouth's opening. Other systems contemplate the head gesture recognition that implies overall head motions or facial expression recognition that combines changes of the mentioned facial features to express an emotion

In this short course, we present an introduction to the main ideas and recent works of non invasive HCI, VBI and PUI in particular oriented to e-Accessibility for disable people. The main work done recently by Computer Graphics & Vision Group and Artificial Intelligence from Computer and Mathematics Department at UIB will be presented. Some videos will be presented in a practical way and real demos using testing software will be used.

1.1 Objectives, Intended Audience and Planning

The main objectives are to show the new advances and techniques in HCI in particular VBI and their special applications to elderly users and disable people. The UE is very interested in the FP7 research program to improve the quality of live for all elderly persons and increase the social relations between older people. So the new paradigms of HCI will be a key point of research in this near future.

This area includes a multidisciplinary research collaborations son the potential users of this short course are very broad. From technician in electronics and computer science to physiologist and social career are potential professional that could be interested in this topics.

The course could be adapted to the level of audience and the time needed to explain the concepts, but four hours sessions are adequate in general. The talks are organized in one hour session. Finally one practical session could be prepared using SINA and Coldiesis software.

Combining Requirements and Interaction Design through Usage Scenarios

Hermann Kaindl

Vienna University of Technology
Gußhausstr. 27-29, A-1040 Vienna, Austria
kaindl@ict.tuwien.ac.at

Abstract. When the requirements and the interaction design of a system are separated, they will most likely not fit together, and the resulting system will be less than optimal. Even if all the real needs are covered in the requirements and also implemented, errors may be induced by human-computer interaction through a bad interaction design and its resulting user interface. Such a system may even not be used at all. Alternatively, a great user interface of a system with features that are not required will not be very useful as well.

Therefore, we argue for combined requirements engineering and interaction design, primarily based on usage scenarios. However, scenario-based approaches vary especially with regard to their use, e.g., employing abstract use cases or integrating scenarios with functions and goals in a systematic design process. So, the key issue to be addressed is how to combine different approaches, e.g., in scenario-based development, so that the interaction design as well as the development of the user interface and of the software internally result in an overall useful and useable system. In particular, scenarios are very helpful for purposes of usability as well.

Keywords: Interaction design, usage scenarios, requirements engineering, user interfaces, usability.

1 Tutorial Goals

This tutorial is targeted towards people who are supposed to work on the interaction design or the requirements in systems development, e.g., interaction designers, user interface developers, Web designers, requirements engineers, or project managers. Whatever the roles of the tutorial participants actually are in their daily work, they should get a better understanding of "other" viewpoints and tasks and, in particular, a common approach. The overall goal of presenting this proposed tutorial is to teach how requirements engineering and interaction design relate and how they can be usefully combined. This can be important for creating better interactive systems in the future.

2 Key Learning Outcomes

In this tutorial, participants learn about combined (concurrent and intertwined) requirements engineering and interaction design. In particular, participants understand

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 932–933, 2009.

how scenarios and use cases can be utilized both for requirements engineering and interaction design, though with different emphasis on the level of detail. They also understand the additional need to specify the functional requirements for the system to be built, even in the context of object-oriented (OO) development. Overall, they gain a better understanding of early systems design.

3 CV of the Presenter

Hermann Kaindl joined the Institute of Computer Technology at the Vienna University of Technology in Vienna, Austria, in early 2003 as a full professor. Prior to moving to academia, he was a senior consultant with the division of program and systems engineering at Siemens AG Austria. There he has gained more than 24 years of industrial experience in software development and human-computer interaction. He has published four books and more than a hundred papers in refereed journals, books and conference proceedings. He is a *Senior Member* of the IEEE, a *Distinguished Scientist* member of the ACM, a *Fellow* of IARIA and a member of the INCOSE, and is on the executive board of the Austrian Society for Artificial Intelligence.

References

1. Kaindl, H.: A Practical Approach to Combining Requirements Definition and Object-Oriented Analysis. Annals of Software Engineering 3, 319–343 (1997)
2. Kaindl, H.: A Design Process Based on a Model Combining Scenarios with Goals and Functions. IEEE Transactions on Systems, Man, and Cybernetics (SMC) Part A 30, 537–551 (2000)
3. Kaindl, H.: Adoption of Requirements Engineering: Conditions for Success. In: Proceedings of the Fifth IEEE International Symposium on Requirements Engineering (RE 2001), invited State-of the-Practice Talk, Toronto, Canada, August, pp. 156–163. IEEE, Los Alamitos (2001)
4. Kaindl, H.: Is Object-oriented Requirements Engineering of Interest? Requirements Engineering 10, 81–84 (2005)
5. Kaindl, H.: A Scenario-Based Approach for Requirements Engineering: Experience in a Telecommunication Software Development Project. Systems Engineering 8, 197–210 (2005)
6. Kaindl, H., Jezek, R.: From Usage Scenarios to User Interface Elements in a Few Steps. In: Proceedings of the Fourth International Conference on Computer-Aided Design of User Interfaces (CADUI 2002), Valenciennes, France, May, pp. 91–102. Kluwer Academic Publishers, Dordrecht (2002)
7. Kaindl, H., Kramer, S., Kacsich, R.: A Case Study of Decomposing Functional Requirements. In: Proceedings of the Third International Conference on Requirements Engineering (ICRE 1998), Colorado Springs, CO, April, pp. 156–163. IEEE Computer Society Press, Los Alamitos (1998)
8. Kaindl, H., Kramer, S., Hailing, M.: An Interactive Guide Through a Defined Modelling Process. In: People and Computers XV, Joint Proceedings of HCI 2001 and IHM 2001, Lille, France, September, pp. 107–124. Springer, London (2001)

Design Patterns for User Interfaces on Mobile Equipment

Erik G. Nilsson

SINTEF ICT, Postboks 124, Blindern, N-0314 Oslo, Norway
egn@sintef.no

Abstract. The objective of this tutorial is to enhance the participants' skills in designing user interfaces for mobile equipment, including adaptive and context sensitive user interfaces and multimodal interaction. Through a combination of lectures and practical exercises, a collection of patterns addressing issues regarding designing user interfaces on mobile devices is presented. The patterns address typical challenges and opportunities when designing user interfaces that are to run on PDAs and SmartPhones – both challenges connected to characteristics of the equipment and connected to tasks to which designing suitable user interfaces is challenging. The tutorial is intended for user interface designer, systems developers, and project leaders that work with or plan to work on development of applications on mobile devices. The tutorial requires basic knowledge of user interface design in general, and basic understanding of challenges connected to designing user interfaces on mobile devices.

1 Description of Tutorial

The objective of the tutorial "Design patterns for user interfaces on mobile equipment" is to enhance the participants' skills in designing user interfaces for mobile equipment, including adaptive and context sensitive user interfaces and multimodal interaction.

The tutorial presents a collection of patterns [1] addressing issues regarding designing user interfaces on mobile devices that we have developed in two research projects conducted in the period 2004-2009 in cooperation with a number of partners developing application and development tools for mobile applications.. The patterns address typical challenges and opportunities when designing user interfaces that are to run on PDAs and SmartPhones – both challenges connected to characteristics of the equipment and connected to tasks to which designing suitable user interfaces is challenging. The patterns in the collection are structured in three main groups:

1. Utilizing screen space
2. Interaction mechanisms
3. Design at large.

The suggested solutions are based on existing design practice, and the description includes examples of good solution and pros and cons of different approaches.

The tutorial contains a mix of presentations and exercises. The exercises are performed in small groups using paper prototyping. The tutorial is intended for user

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 934–935, 2009.

interface designer, systems developers, and project leaders that work with or plan to work on development of applications on mobile devices. The tutorial requires basic knowledge of user interface design in general, and basic understanding of challenges connected to designing user interfaces on mobile devices.

The participants at the tutorial will learn about design challenges and opportunities that are specific and important when designing user interfaces on mobile devices. They will obtain general knowledge about approaches for overcoming the challenges and exploiting the opportunities, in addition to specific knowledge about solving the most important problems.

Participation in the tutorial will be beneficial for the Interact audience because it both gives a structured overview of the challenges and opportunities connected to user interface design on mobile devices, and gives practical solutions to the most important challenges, including adaptive and context sensitive user interfaces and multimodal interaction.

Tutorial documentation includes a comprehensive research report documenting the patterns collection, in addition to important choices regarding platforms, user interface style and deployment strategies, overview of main differences between mobile and stationary user interfaces, and important issues from platform style guides. Furthermore, the tutorial documentation includes an extensive bibliography. When solving the exercises, the participants will also be using an online patterns collection for user interfaces on mobile devices developed by the presenter in cooperation with colleagues at SINTEF ICT.

The tutorial will be presented by **Erik G. Nilsson,** a Senior Research Scientist at SINTEF ICT. He has been working with model-driven systems development at SINTEF since 1984, with a focus on user interface development the last fifteen years. The last eight year he has performed research on user interface development for mobile equipment, with a focus on adaptation and exploiting context information. He has been and is the project leader for two Norwegian research projects (UMBRA and FLAMINCO) that develop design patterns and evaluation methods for user interfaces on mobile equipment. The patterns and methods are developed in co-operation with Norwegian development and consulting companies focusing on mobile technology. Nilsson has authored/co-authored publications at international refereed journals and conferences on re-engineering, systems integration, user interface design, mobile user interface design and model-based user interface development. He has also been instructor at a large number of courses on user interface design and development for Norwegian companies and organizations, and given presentation on numerous industry oriented seminars.

Reference

1. Nilsson, E.G.: Design patterns for user interface for mobile applications. Adv. Eng. Softw. (in press, 2009)

Eye Tracking in Human-Computer Interaction and Usability Research

Tommy Strandvall

Tobii Technology, Karlsrovagen 2D, 182 53 Danderyd, Sweden
tommy.strandvall@tobii.com

Abstract. The objective of the tutorial is to give an overview on how eye tracking is currently used and how it can be used as a method in human computer interaction research and especially in usability research. An eye tracking system records how the eyes move while a subject is completing a task for example on a web site. By analyzing these eye movements we are able to gain an objective insight into the behavior of that person.

Keywords: eye tracking, eye movements, usability, human-computer interaction, method, methodology.

1 Tutorial Overview

The objective of the tutorial is to give an overview on how eye tracking is used and can be used as a method in human computer interaction research and especially in usability research. By analyzing eye movements collected by an eye tracking system we are able to gain an objective insight into the behavior of a person and present the results as powerful visualizations or eye movement metrics and statistics.

During the tutorial we will present a range of eye tracking metrics and visualizations and explain how these can be interpreted in the context of interface design and usability evaluation. Examples of such eye movement metrics are; Number of Fixations, Fixation Duration, and Time to First Fixation. Common ways of visualizing eye movements in a usability or HCI setting are; Heat Maps, Gaze Plots or Scan Paths, and Gaze Replay Videos. Using these eye tracking metrics and visualization together with existing research methods can reveal insights previously unavailable to researchers and practitioners. Additionally it allows us to develop new approaches and methodologies. Some new methodologies where eye tracking plays an important role have already been developed like in the case of the Post-Experience Eye-Tracked Protocol (PEEP) [1].

The tutorial will moreover include hands-on eye tracking tests and analysis exercises. This tutorial is appropriate for researchers, practitioners and students. Participants should be familiar with behavioral research methodology but no prior knowledge of eye movements or eye tracking is required.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 936–937, 2009.

2 Introduction to Eye Movements

An eye tracking system records how the eyes move when a subject is sitting in front of a computer screen. The human eyes are constantly moving until they stop and focus on a point. There are over ten different types of eye movements, of which the most important ones are saccades and fixations. When the eyes focus on a point it is called a fixation and the movements between these fixations are called saccades. When the eye fixates, the stops vary from about 100 to 600 milliseconds and during this stop the brain starts to process the visual information received from the eyes. The length of a fixation is usually an indication of information processing or cognitive activities as this is when the brain interprets the visual information from the eyes. The human eye has a visual field of about 200º but the highest number of light sensitive cells on the retina are located in the part called fovea, which is the only point in our eyes where we are able to see a sharp and colorful image of the world around us. This area is fairly small and covers only about 1 - 2 degrees of our vision (which is about the size of a thumbnail at an arm length's distance). It is only from these cells our brain can receive detailed visual information. An eye tracker records these movements and the location of the foveal vision when the eyes fixate. [2]

It is possible to move our covert attention (the attention of our mind) around the entire visual field when our eyes are at rest. Thus it is possible to move our attention around without eye movements. However, our foveal vision is usually a valid measure for determining the target of our covert attention as our brain can process very little information from complex stimuli from the area outside the fovea (such as objects on a website or software interface) [3]. In such cases it is more efficient for our brain to focus the attention on the fovea rather than on our peripheral vision, as the brain needs to process blurry visual information requiring more effort to interpret than the visual information from the fovea.

This is why it is possible to tell something about human behavior by just following eye movements, especially the fixations, as we know that we can only see something clearly when we fixate on, or near, an object.

References

1. Ball, L., Eger, N., Stevens, R., Dodd, J.: Applying the post-experience eye-tracked protocol (PEEP) method in usability testing. In: Interfaces, Informs, Cincinnati, vol. 67, pp. 15–19 (2006)
2. Rayner, K.: Eye Movements in Reading and Information Processing: 20 Years of Research. Psychological Bulletin 124, 372–422 (1998)
3. Pieters, R., Wedel, M.: Informativeness of eye movements for visual marketing: six cornerstones. In: Wedel, M., Pieters, R. (eds.) Visual Marketing: From Attention to Action. Lawrence Earlbaum, New York (2008)

HCI in the Era of Ambient Media – And beyond 2009 INTERACT Tutorial

Artur Lugmayr

Tampere University of Technology (TUT), POB. 553
Korkeakoulunkatu 1, FIN-33101 Tampere, Finland
artur.lugmayr@tut.fi

Abstract. According to McLuhan, "the medium is the message" – but what means interaction and what is the medium in the age of ubiquitous and pervasive computation – when the medium is 'in' daily objects? Ambient media are media that are embedded throughout our natural environment – location based services, context awareness, emotional responsive interfaces, touch and gesture based interfaces, haptics and biometrics, sensor perception, mobile devices, and smart data mining are the technological enabler for smart media environments. The latest trends from emotional computation, affective computation, and tangible media lay the foundations for this new and exciting form of media existing far beyond screen concepts and mouse based interaction metaphors. The tutorial trains participants in the basic technologies as tools for the design of new interactive 'ambient' environments. It presents case studies and latest research results in the field of ambient media, ranging from ambient assisting living, user experience design, user contributed content, and mobile services. After the tutorial the participants should understand the principles of ambient media with its underlying concepts and methods, especially emphasizing human-computer-interaction. As roundup, the tutorial presents a more visionary viewpoint to the future of media technology: the use of biological metaphors for interactive environments (biomedia).

Keywords: ubiquitous computation, ambient media, pervasive computation.

1 Objectives and Schedule

The goal of the tutorial is to train participants in the basics of ambient media especially focusing on the human-computer-interaction viewpoint. The tutorial is designed for a general audience with interest in a newly emerging media environment and its possibilities.

Monday Afternoon, 24th August 2009
13:30-15:00 Part 1: Introduction, Concepts Overview, Technology
15:00-15:30 Coffee break
15:30-17:00 Part 2: Case Studies, Design Guidelines, Outlook in the Future

Please visit http://www.cs.tut.fi/~lartur for further material and information.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 938–939, 2009.

2 Focus Points of the Tutorial

The tutorial covers the following topics in further depth:

- case-studies of existing ambient media services
- basic concepts and technologies of ambient media
- location based services, mobile interaction, and smart environments
- user experience and interaction design guidelines
- ambient content production and creation
- natural and intuitive interaction methods
- context awareness and intelligent behavior modeling
- proactive and emotional responsive system designs
- ambient services and business models
- ambient social networks.

Introduction to Social Network Analysis

Panayiotis Zaphiris[1] and Chee Siang Ang[2]

[1] Department of Multimedia & Graphic Arts, Cyprus University of Technology, Lemessos, 3603, Cyprus
[2] Centre for HCI Design, City University London, EC1V 0HB, UK
pzaphiri@gmail.com, jimmybbq@gmail.com

Abstract. Social Network analysis focuses on patterns of relations between and among people, organizations, states, etc. It aims to describe networks of relations as fully as possible, identify prominent patterns in such networks, trace the flow of information through them, and discover what effects these relations and networks have on people and organizations. Social network analysis offers a very promising potential for analyzing human-human interactions in online communities (discussion boards, newsgroups, virtual organizations). This Tutorial provides an overview of this analytic technique and demonstrates how it can be used in Human Computer Interaction (HCI) research and practice, focusing especially on Computer Mediated Communication (CMC). This topic acquires particular importance these days, with the increasing popularity of social networking websites (e.g., youtube, myspace, MMORPGs etc.) and the research interest in studying them.

Keywords: Social Network Analysis, Social Computing, Computer Mediated Communication.

1 Objectives

This tutorial provides an overview of Social Network Analysis (SNA) and demonstrates through theory and practical case studies how it can be used in HCI (especially computer-mediated communication and CSCW) research and practice. This topic becomes even more important these days with the increasing popularity of social networking websites (e.g. YouTube, Facebook, MySpace etc.) and social computing. As people increasingly use online communities for social interaction, new methods are needed to study these phenomena. SNA is a valuable contribution to HCI research as it gives an opportunity to study the complex patterns of online communication.

Social network theory views a network as a group of actors who are connected by a set of relationships. Actors are often people, but can also be nations, organizations, objects etc. Social Network Analysis (SNA) focuses on patterns of relations between these actors. It seeks to describe networks of relations as fully as possible. This includes teasing out the prominent patterns in such networks, tracing the flow of information through them, and discovering what effects these relations and networks have on people and organizations. It can therefore be used to study network patterns of

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 940–941, 2009.

organizations, ideas, and people that are connected via various means in an online environment.

Upon completion of this tutorial, attendees will:

1. Be able to understand the basics of social network analysis, its terminology and background (part 1)
2. Be able to transform communication data to network data (part 1)
3. Know practically how social network analysis (SNA) can be applied to HCI (especially CMC) analysis (part 2)
4. Get familiar with the use of standard SNA tools and software (part 2)
5. Be able to derive practical and useful information through SNA analysis that would help design an innovative and successful online community. (part 2)

2 Content and Benefits

The content of the tutorial is divided into two parts, each of which is structured in small groups to maximize the interaction among participants:

1. Part 1: Introduction to Social Network Analysis. Benefits: you will be exposed to the introduction of SNA, get familiar with the terminology and definitions of SNA.
2. Part 2: Practical uses of social network analysis (SNA). Benefits: Through a series of interactive exercises, a number of case studies will be demonstrated and discussed. Case studies will draw from diverse areas (e.g. use of SNA to study age differences in CMC, use of SNA in universal design and research). Ways of using SNA to study new forms of CMC such as MMORPGs, Wikis, blogs etc. will also be discussed.

3 Presentation Format

We will deliver the tutorial content in various presentation formats. The presentation of lecture material will include several different media. The presentation will be accompanied by videos and images to demonstrate examples. Furthermore, we will present popular SNA tools in order to provide the audience with hands on experience of using SNA software. Especially the more practical parts of the tutorial will be held in an interactive atmosphere. Discussions and exercises will give attendees the opportunity to bring in knowledge and questions from their own research and work area. As SNA can be applied in a variety of different research and work areas, we will allow for time to discuss and elaborate on the possibilities of SNA to the research or working areas of the participants.

4 Target Audience

We welcome practitioners and academics interested in computer mediated communication, universal design, especially researchers and practitioners who are interested in domains that social network analysis can be applied.

Key Issues in Planning and Making Sense of International Field Research

Susan M. Dray and David A. Siegel

Dray & Associates, Inc., 2007 Kenwood Parkway, Minneapolis, Minnesota 55405 USA
firstname.lastname@dray.com

Abstract. More and more companies are doing international research as they design products and services for people around the world. These studies are particularly complex and challenging however, and there are, to date, relatively few resources to guide the people doing them. This tutorial uses scenarios and discussion to help practitioners learn how to approach, plan, manage, and interpret international field studies.

Keywords: Fieldwork, ethnography, user research, international research, user experience.

1 Introduction

Doing international field research is highly rewarding when it is done well, but it is fraught with challenges and complexities that can overwhelm even the most seasoned user researchers.

This tutorial focuses on key issues by posing scenarios that have been constructed to illustrate common challenges. We discuss and dissect these with participants. We also share "tips and tricks" that we have learned and found useful when doing international field research.

2 Key Areas That Can Make or Break International Field Research

2.1 Making the Strategic Business Case

Often, strategic arguments are most effective in making the case for international field research. Such arguments can focus on the strategic importance of international markets, risks of mismatches of product concepts and local conditions, or other factors of particular importance in your company.

2.2 Preventing an Overloaded Research Agenda

During the planning of an international user research project, user experience professionals often must struggle to keep the study from being overloaded with topics that

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 942–943, 2009.

either can dilute the research or can threaten to make it so hopelessly complex that it is unwieldy at best, impossible at worst.

2.3 Team Composition and Skills

International field research typically requires cross-cultural and often cross-functional teams. It is important to include local team members who are familiar with the culture and context. It can be equally important to include "outsiders" who are not embedded in the local culture, and people with skills in cross-cultural interpretation. Since they are be aware of the contrasts between observations and expectations, they are be able to help "message" the findings to be useful by the design teams "back home."

2.4 Dealing with Challenges of Sampling and Recruiting

The challenges of sampling and recruiting are complex. Criteria and screeners need to be localized and contextualized – as well as translated. You often must also deal with sampling bias that can be built into the recruiting process. The best way to deal with these issues is to treat the recruiting process as a core part of the research.

2.5 Interpreting Findings

Inevitably, international field researchers must make sense of "surprises" in the field. While these "surprises" are really the whole point of the research, they can present complex interpretive challenges. Sometimes, teams resort to stereotyping of the participants or jumping to simplistic conclusions about the findings. Being aware of some of the common pitfalls of interpretation and having strategies for evaluating interpretive hypotheses can help prevent these problems.

Measuring the Subjective User eXperience

Maurits Kaptein

Eindhoven University of Technology / Stanford University
Groesbeekseweg 124
6524DM Nijmegen, The Netherlands
m.c.kaptein@tue.nl

Abstract. Measuring the subjective user experience is a challenging task. In this tutorial we will demonstrate how psychological constructs can be divided in separate variables, each measured by its individual questionnaire items. The tutorial will address the analysis of the questionnaire data to estimate its validity and reliability. Analysis will be demonstrated using SPSS.

Duration: 3 hours total.

1 Measuring the Subjective User Experience

Recently you could have read the following: "*The evaluated prototype led to a higher sense of social connectedness. This was especially true for the 3D interface*". At first glance this seems positive: Social connectedness sounds like something you would want, and if this is provided by the evaluated prototype than perhaps we should pursue developing it as a product. However, several questions remain: 1. What exactly is social connectedness? 2. How is it operationalized in this study? 3. Has it been measured reliably? 4. Validly? And as such is the initial gut feeling of a positive result actually justified?

This tutorial addresses just these questions. We will show how psychologists tackle the problems of measuring the subjective user experience. We will address hypothetical constructs (such as social connectedness), variables and items. We will address the different types of validity and reliability. Furthermore, we will address the phrasing and wording of questionnaire items [4]. Finally we will address the quantitative analysis of the results of an administered questionnaire. Common problems associated with using existing questionnaires are also discussed.

Questionnaire development starts with the identification of the hypothetical construct one intends to measure, for example *social connectedness*. After identifying the construct one has to determine which underlying variables would properly reflect social connectedness: *bond with your social network*, *the sense of belonging*, *the feeling of being in touch...* After identifying the variables, these have to be measured using different items: the actual questions shown to participants [2].

Next we focus on *validity* and *reliability*. Validity reflects the extent to which your questions measure what you intend to measure. The tutorial first addresses construct validity and then addresses content validity. Furthermore, internal and external

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 944–945, 2009.

validity, ecological validity, temporal validity and face validity are addressed. Reliability reflects the extent to which the measurement reflects the actual score a participant has on the hypothetical construct and thus the extent to which error is omitted from the measurement. We will address test-retest reliability, split-half reliability, and inter-rater reliability [1].

We will continue with a hands-on approach to writing questionnaire items. We will address the informed consent and the special layout of the questionnaire. Furthermore, question wording and inability or unwillingness to answer are addressed. Finally we address the effects of the order of questionnaire items.

At this point in the tutorial it should be clear to participants how to start building up a questionnaire based on psychological literature, evaluate the validity and reliability of previously reported questionnaires, and design ones own items and questionnaire. The remainder of the tutorial will focus on two quantitative techniques to analyze the results of the administered questionnaire.

First off all we will focus on so-called reliability analysis – the systematic analysis of inter item correlations to be able to conclude whether the phrased items indeed reflect one variable. From the concept of item correlation we proceed to inter-item correlations and to scale analysis. Cronbach's Alpha [3] is presented as a measure for scale reliability and interpretations are given. Furthermore, examples are provided of methods to increase scale reliability. Analysis will be supported by practical examples using SPSS.

Secondly we will look at Principal Component Analysis [5]. The purpose of this technique – to identify latent variables in an item set – is explained. Furthermore we give practical examples on determining the number of components (working with Eigenvalues and Scree plots), interpreting the component solution (working with component loadings), and rotating the component solution. Within the section on principal component analysis we will also address confirmative versus explorative analysis and the similarities and differences between component and factor analysis.

At the end of the tutorial respondents will have had a thorough theoretical overview of the process of questionnaire development. Participants will be able to start developing their own questionnaires, and they will be able to correctly judge whether conclusion derived from questionnaire usage by others are reliable and valid. Finally, participants will gain experience in analyzing questionnaire data to evaluate both its reliability and validity. (See http://www.shortadress.com/interact)

References

1. Carmines, E.G., Zeller, R.A.: Reliability and Validity Assessment. Sage, Thousand Oaks (1979)
2. Converse, J.M., Presser, S.: Survey Questions: Handcrafting the Standardized Questionnaire. Sage, Thousand Oaks (1986)
3. Cortina, J.M.: What is coefficient alpha? An examination of theory and applications. Journal of applied psychology 78, 98–104 (1993)
4. Malhotra, N.K.: Marketing research. Pearson Education, London (2008)
5. Tipping, M.E., Bishop, C.M.: Probabilistic Principal Component Analysis. Journal of the Royal Statistical Society 61, 611–622 (1999)

Methods and Tools for Ethical Usability

Iordanis Kavathatzopoulos[1], Agata Kostrzewa[2], and Mikael Laaksoharju[1]

[1] Uppsala University, Department of Information Technology – Human-Computer Interaction, Box 337, 751 05 Uppsala, Sweden
[2] Uppsala University, Department of Business Studies, Box 513, 751 20 Uppsala, Sweden
iordanis@it.uu.se, Agata.Kostrzewa@fek.uu.se, Mikael.Laaksoharju@it.uu.se

Abstract. The objectives of the tutorial are to provide knowledge of basic ethical, psychological and organizational theories that are relevant to consider ethical aspects during design and use of IT systems; knowledge and skills about handling and solving ethical problems in connection with design and use of IT-systems; and skills in using questionnaires, surveys, interviews and the like in connection with software development and IT-use. It contains lectures, workshop and exercises; use of special tools to identify and consider IT ethical issues during planning, construction, installation and use of IT systems; and group exercises where the participants train their ethical skills on IT ethical conflicts and problems. Intended participants are system developers, purchasers, usability experts, academics, HCI teachers.

Keywords: Ethics, Usability, Tools, Methods, Skills.

1 Ethical Competence for IT Systems

Ethical problems caused by the use of IT may have very serious consequences. On the other hand, it is beyond any doubt that the anticipation of ethical problem situations and their proper solution can facilitate design, implementation and efficient use of IT.

Competence to handle ethical problems in a satisfying way is very important for any IT customer, designer or user. But these professionals, just as other people, are often lacking knowledge about what is the right thing to do in concrete IT systems construction situations. Applicable ethical guidelines are extremely difficult to formulate and therefore ethical guidelines are not any option. What would be possible though is to adopt a practical method to identify and solve ethical problems connected to IT use and design. Special tools and working methods can be applied to identify relevant values and stakeholder interests, and to integrate them into the design of the IT system under development.

This tutorial will offer the participants a basic knowledge of ethical and psychological decision-making theories as well as train skills to use certain tools and methods in handling ethical issues. The focus will not be on moral philosophical normative aspects but on individual and organizational ethical problem-solving and decision-making processes. The participants will be trained to use the tools and methods on representative cases as well as on own system development projects.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 946–947, 2009.

There are many moral aspects to the introduction of new technology and there are different approaches to deal with these. Here we do not use normative theory. We only suggest tools and methods that can aid people, who are not informed about ethical theories, when making decisions related to ethical aspects. Our tools can increase the level of ethical competence before a decision is taken, by simply describing all relevant values and aspects for all involved parts in a structured way. By iteratively considering how each possible action or decision affects each possible value for each involved person, company, organisation, etc, a broader and more complex view of the moral dilemma is achieved. The output from these tools work both as a decision support, but also as a kind of documentation for future reference, for continuous dialog and for argumentation reasons. If someone later questions a decision then the documentation can work to explain it, or if additional aspects of the dilemma are revealed then the documentation can be extended with this.

The primary purpose is not to generate answers to particular problems, but rather to acquire a comprehensive view of the problem at hand and also to document the information on which one base the decision making. No matter what the final decision is, our method summarises all aspects considered during the decision-making process. A successful inclusion of moral aspects in IT systems analysis, decision making and decision application can easily fail. The cause of this failure may be found in the way thinking, problem solving and decision making are performed by, for example, a system developer. People use different ways to handle moral problems. Psychological theory and research differentiate between two different moral functions, heteronomy and autonomy, which decide a person's ability to handle moral problems.

Decision makers, developers, users and purchasers, need high ethical competence and confidence in working with ethical issues during IT systems development. That is necessary in order to solve moral problems, and to make moral decisions in accordance with relevant values, principles and interests. Spontaneous subjective reactions to moral issues may give solutions to problems, which probably satisfy one's moral feelings. However, with such more or less emotional reactions only, the relevant factors of the particular moral problem are certainly not fully taken into account. What is needed is a psychological approach to ethical competence and skills. It appears then that the process or the way one tries to identify relevant ethical aspects is what really matters. Suitable tools and working methods that support this way of handling ethical usability issues are necessary.

Our tools and methods, such as EthXpert, support the adoption of a holistic approach, of critical reasoning and of systematic analysis of the ethical aspects during concrete IT systems development. By using our tools and methods for the solution of IT ethical problems decision makers get a great amount of help to identify the significant problems, to make them explicit, and to reformulate them in order to be able to work with them. Decision makers, such as system developers or system purchasers, define their own (organizational as well as personal) position, duties, commitments, values and feelings. They also identify and take into consideration the interests, values and needs of all stakeholders. They generate alternative courses of action, and they systematically weigh each of them against all values and interests involved in the situation. By that they can consider all relevant ethical aspects and adopt the ethically most suitable design solutions.

Model a Discourse and Transform It to Your User Interface

Hermann Kaindl

Vienna University of Technology
Gußhausstr. 27-29, A-1040 Vienna, Austria
kaindl@ict.tuwien.ac.at

Abstract. Every interactive system needs a user interface, today possibly even several ones adapted for different devices (PCs, PDAs, mobile phones). Developing a user interface is difficult and takes a lot of effort, since it normally requires design and implementation. This is also expensive, and even more so for several user interfaces for different devices.

This tutorial shows how human-computer interaction can be based on discourse modeling, even without employing speech or natural language. Our discourse models are derived from results of Human Communication theories, Cognitive Science and Sociology. Such discourse models can specify an interaction design. This tutorial also demonstrates how such an interaction design can be used for model-driven generation of user interfaces and linking them to the application logic and the domain of discourse.

Keywords: Discourse modeling, model-driven user interface generation.

1 Tutorial Goals

The main goals are to show a new approach to modeling discourses, and to demonstrate that and to explain how user interfaces can be generated automatically from high-level discourse models, primarily through transformations using a model-driven approach. Ideally, even end users may be able to model discourses and to transform them into their user interfaces.

2 Key Learning Outcomes

In this tutorial, participants learn about modeling discourses using a new approach inspired by human-human communication. They know how modeling discourses and generating user interfaces can be approached systematically. Based on the available tool for automated generation of user interfaces from such models, participants also learn about a new approach to UI prototyping with "perfect fidelity". The result would be identical to the real UI due to this generation process, assuming that the same interaction design according to our approach is taken as input. After a change in the discourse model, the related changes in the UI can be seen more or less immediately.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 948–949, 2009.

References

1. Bogdan, C., Falb, J., Kaindl, H., Kavaldjian, S., Popp, R., Horacek, H., Arnautovic, E., Szep, A.: Generating an Abstract User Interface from a Discourse Model Inspired by Human Communication. In: Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 41), Big Island, HI, USA. IEEE Computer Society Press, Los Alamitos (2007)
2. Bogdan, C., Kaindl, H., Falb, J., Popp, R.: Modeling of interaction design by end users through discourse modeling. In: Proceedings of the 2008 ACM International Conference on Intelligent User Interfaces (IUI 2008), Maspalomas, Gran Canaria, Spain. ACM Press, New York (2008)
3. Falb, J., Kaindl, H., Horacek, H., Bogdan, C., Popp, R., Arnautovic, E.: A discourse model for interaction design based on theories of human communication. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems, pp. 754–759. ACM Press, New York (2006)
4. Falb, J., Kavaldjian, S., Popp, R., Raneburger, D., Arnautovic, E., Kaindl, H.: Fully Automatic User Interface Generation from Discourse Models. In: Proceedings of the 2009 ACM International Conference on Intelligent User Interfaces (IUI 2009), Sanibel Island, Florida, USA. ACM Press, New York (2009); Tool demo paper
5. Falb, J., Popp, R., Röck, T., Jelinek, H., Arnautovic, E., Kaindl, H.: UI Prototyping for Multiple Devices Through Specifying Interaction Design. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4662, pp. 136–149. Springer, Heidelberg (2007)
6. Kaindl, H.: A Design Process Based on a Model Combining Scenarios with Goals and Functions. IEEE Transactions on Systems, Man, and Cybernetics (SMC) Part A 30, 537–551 (2000)
7. Kaindl, H., Jezek, R.: From Usage Scenarios to User Interface Elements in a Few Steps. In: Proceedings of the Fourth International Conference on Computer-Aided Design of User Interfaces (CADUI 2002), Valenciennes, France, May, pp. 91–102. Kluwer Academic Publishers, Dordrecht (2002)
8. Kavaldjian, S., Bogdan, C., Falb, J., Kaindl, H.: Transforming Discourse Models to Structural User Interface Models. In: Giese, H. (ed.) MoDELS 2008. LNCS, vol. 5002, pp. 77–88. Springer, Heidelberg (2008)

Understanding Users in Context: An In-Depth Introduction to Fieldwork for User Centered Design

Susan M. Dray and David A. Siegel

Dray & Associates, Inc., 2007 Kenwood Parkway, Minneapolis, Minnesota 55405 USA
firstname.lastname@dray.com

Abstract. There is increased awareness of the need for design to be driven by deep understanding of users, their activity patterns, processes, needs and external influences--understanding that can only be gained by studying user behavior in the user's context. This requires understanding of how to plan and carry out observational studies of users, which is a new skill for many. In addition, fieldwork is bigger than any one methodology. Therefore, in this tutorial, we will take a fresh and deeper look at fundamental principles, teach a range of techniques, and examine important issues on which methods differ.

Keywords: Fieldwork, ethnography, user research, contextual inquiry, artifact walkthough, naturalistic usability evaluation.

1 Introduction

Observational research differs from other ways of gathering user data and complements other user-centered design (UCD) methods. For instance, a deep understanding of challenges in users' work patterns can help identify user requirements for new tools, or may spark entirely new product concepts. Information about the context of use that is gained from these studies can inform specific design decisions about a wide variety of things including navigation requirements and affordances of new products. In addition, the usage context and dynamics of use can contribute to building of robust scenarios both to design to and to test against.

Doing useful and valid field research is not as simple as just watching users, or talking with them in their environment. It requires ways to approach the inherent challenges of such research, such as dealing with the inherent ambiguity of qualitative data from field studies; risks of anecdotal evidence; balancing observation and inquiry; sampling bias; reducing the risks of premature closure, identifying and controlling (to the extent possible) reactive effects and demand characteristics.

Planning the field research is a critical first step in assuring its validity. Establishing a focus and creating a focus structure are crucial steps in preparing for field research. The focus determines what will be salient for the researcher, and helps guide the researcher in prioritizing the many avenues that can be pursued during the data gathering in the field. Defining the focus is probably the most critical step in determining how you set up the study and whether you will actually bring back useful and relevant information.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 950–951, 2009.

2 Three Data Gathering Techniques

2.1 Contextual Inquiry

Contextual Inquiry depends upon interaction with users in real time, combined with observation of them as they do their tasks, as well as observation of their context. Unlike a conventional interview, contextual inquiry is very non-linear. Opportunities to probe deeply as well as opportunities to broaden the exploration arise unpredictably. Specific inquiry techniques help the researcher not only to elicit samples of behavior and explore contextual dependencies, but also to balance depth with breadth.

2.2 Artifact Walkthrough

The actual interview techniques of Artifact Walkthrough can resemble those of Contextual Inquiry. However, Artifact Walkthroughs apply these techniques in situations where it is difficult to observe the process of interest. For example, the process may be intermittent and difficult to capture in real time, or it might be inappropriate or impossible to interrupt it. Artifact Walkthroughs allow for exploration of these processes retrospectively while grounding the information in concrete evidence of behavior, which helps deal with the limitations of participant recall. They also provide openings into wider exploration of the user’s process. As with actual behavioral observation, they often bring out aspects of the process that the researcher would not have thought to ask about and the user would have thought to mention

2.3 Naturalistic Usability Evaluation

Naturalistic usability evaluation encompasses a range of techniques for evaluating the user's interaction with technology in the user's natural context, based on the user's own goals and materials. It therefore focuses on tasks that are meaningful to each user or that incorporate the user’s own content more easily than laboratory evaluations. They also allow you to study both usability and utility conjointly and to explore both discovery and task performance.

2nd Workshop on Design for Social Interaction through Physical Play

Tilde Bekker, Janienke Sturm, and Emilia Barakova

Dept. of Industrial Design, Eindhoven University of Technology
Den Dolech 25612AZ Eindhoven, The Netherlands
{M.M.Bekker,J.Sturm,E.I.Barakova}@tue.nl

Abstract. We aim to stimulate social interaction by designing and creating interactive objects for physical play for diverse user groups, such as children, elderly or people with special needs. With this workshop we aim to bring researchers and practitioners together to share and explore issues and opportunities for technology-enhanced physical play for stimulating face-to-face social interaction (as opposed to virtual interaction through a computer). The focus of this workshop is on sharing theories that are valuable for the design and research of products and applications in this field.

Keywords: Social interaction, play, interactive technologies, theory.

1 Theme

Social interaction between people is important for many reasons. For example, children develop many skills just by playing and communicating with other children, such as cooperating, sharing and other cognitive and motor skills. Social interaction is also extremely important for the well-being of elderly people, who tend to get lonely when socializing becomes more difficult because of their age. Unfortunately, because of recent technological advances children spend an increasing amount of time behind their computer playing games on their own. Moreover, an increasing number of people interact through media such as email and instant messaging in addition to or sometimes even at the expense of face-to-face communication.

At the same time, however, new technologies provide interesting novel opportunities for entertainment and interaction. For example, nowadays a multitude of sensors can be used to detect the behavior or physical condition of users. This information can then be used to control the performance of objects or to provide feedback in the form of light, music, etc. In this way, face-to-face interaction can be enhanced and stimulated in a playful and fun way, for children and for elderly, but also for user groups with social impairments, such as autistic children.

Physical play offers many opportunities to enhance and stimulate real social interaction (as opposed to virtual interaction through a computer). For example, one of the main reasons why playing football is so popular among both children and adults is the fact that they can play with and against each other. Playing a game is also a popular

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 952–953, 2009.

means to stimulate autistic children and elderly people to socialize and have fun. Combining technology-enhanced objects that promote social interaction with dedicated games and scenarios may enhance the experience of using those embodied objects in social settings.

Design for social and physical play is a young, multidisciplinary area of research that is closely related to various other fields, such as tangible interaction, robotics, computer games, pervasive games, and exertion games.

2 Objectives

With this workshop we aim to bring researchers together to share and explore issues and opportunities for technology-enhanced physical play for stimulating face-to-face social interaction (as opposed to virtual interaction through a computer).

The first workshop on Design for Social Interaction through Physical Play was held at the Fun and Games conference in Eindhoven, The Netherlands in 2008 and focused on sharing and identifying common interests. Researchers with various backgrounds, such as robotics, industrial design, social psychology, etc. had fruitful discussions about many of the issues that we encounter in our work. One of the conclusions of the workshop was that there is a need for sharing theories that can be used for the design and research of products and applications in this field.

For that reason, our focus for this workshop is on sharing theories about, for example, learning, play or the use of technology. We will discuss their use and applicability to our field of design research. For example, participants from the game design area were quite interested in theories about social connectedness, and how these would inform the design of social games.

Among the expected outcomes of the workshop are an overview of valuable theories that can be used in this area of research and how they inform evaluation methods and metrics. Also, we plan to select the best position papers for publication in a special issue of a relevant international journal.

References

1. Barakova, E., van Wanrooij, G., van Limpt, R., Menting, M.: Using an emergent system concept in designing interactive games for autistic children. In: Proceedings of the 6th international Conference on interaction Design and Children, IDC 2007 (2007)
2. Bekker, T., Sturm, J., Barakova, E.: Design for Social Interaction through Physical Play. To appear in Special Issue of Personal and Ubiquitous Computing (2008)
3. Sturm, J., Bekker, T., Groenendaal, B., Wesselink, R., Eggen, B.: Key issues for the successful design of an intelligent interactive playground. In: Proceedings of the 7th international Conference on Interaction Design and Children, IDC 2008 (2008)

4th Workshop on Software and Usability Engineering Cross-Pollination: Usability Evaluation of Advanced Interfaces

Regina Bernhaupt[1], Peter Forbrig[2], Jan Gulliksen[3], and Janet Wesson[4]

[1] ITC&S-Center, University of Salzburg, Sigmund-Haffner-Gasse 18
A-5020 Salzburg, Austria
Regina.Bernhaupt@sbg.ac.at
[2] University of Rostock, Computer Science Department, Albert-Einstein-Str. 21
D-18051 Rostock, Germany
Peter.Forbrig@uni-rostock.de
[3] Dept. for Human Computer Interaction, Uppsala University, Box 337
SE-751 05 Uppsala, Sweden
Jan.Gulliksen@hci.uu.se
[4] Nelson Mandela Metropolitan University, Dept. of Comp. Sc. & Inf. Syst., PO Box 77000
6031 Port Elisabeth, South Africa
Janet.Wesson@nmmu.ac.za

Abstract. The usability evaluation of advanced interfaces such as intelligent, adaptive and context-aware interfaces presents several challenges. These interfaces need to be evaluated over an extended time period and used in their actual context of use, rather than in a laboratory setting. Usability evaluation methods for traditional interactive systems have been extensively researched and are well understood. These methods include heuristic evaluation, usability testing and field studies. The goal of this workshop is to bring together people who are engaged in the design and evaluation of advanced interfaces across different disciplines. The objective is to exchange ideas and techniques relevant to the usability evaluation of advanced interfaces and to establish guidelines for the evaluation of such interfaces that cross-pollinate the different disciplines.

Keywords: Usability Evaluation, Advanced Interfaces, Guidelines.

1 Introduction

HCI experts, software engineers and usability engineering are affected by a mutual influence that we call “cross-pollination”. Examples are task specifications, design patterns and life cycle models. These examples were invented in one field and later on adapted in a new context. New developments in intelligent and adaptive environments and mobile computing require new solutions, especially for usability evaluation methods [1-4]. The key attribute of advanced interfaces is that they need to adapt to time, location and usage which makes them very difficult to evaluate using standard techniques [5-8].

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 954–956, 2009.

The workshop will focus on how to integrate and extend traditional evaluation methods in order to optimally evaluate the usability of advanced interfaces in their specific context of use [9, 10]. Experts in HCI, software and usability engineering need to learn from each other to facilitate and encourage this convergence.

The workshop aims to be a forum for sharing ideas about potential and innovative ways to cross-pollinate the expertise among the different communities and to show examples, which can stimulate industrial software development. Additionally it should provide a forum that will help to grow a community of interest in this area.

2 Structure of the Workshop

2.1 Goals and Topics

The goals of this workshop are to provide HCI specialists, software engineers and usability specialists from industry and research institutions an opportunity to discuss both the state-of-the art and the cutting edge practice in usability evaluation.

Topics of interest include, but are not limited to, the usability evaluation of the following advanced interfaces and interactive systems:

• Adaptive Interfaces • Context-aware Interfaces • Human-Robot Interfaces • Intelligent Interfaces	• Location-aware Interfaces • Mobile Interfaces • Novel Interfaces

2.2 Participation

The workshop is the official workshop of IFIP working group 13.2 "Methodologies for User-Centered Systems Design". http://wwwswt.informatik.uni-rostock.de/IFIP_13_2/
It expects HCI specialists, software and usability engineers from academia and industry as participants.

2.3 Workshop Activities and Dissemination

Participants have to prepare a position paper of 4 to 10 pages which will be reviewed. Selected papers will be published on the workshop web site (http://wwwswt.informatik.uni-rostock.de/EVAL/) and will be presented during the one day workshop. The outcome of the workshop will be a white paper presented on the web site of the workshop.

2.4 Workshop Organizers

Regina Bernhaupt
ITC&S-Center, University of Salzburg
Sigmund-Haffner-Gasse 18
A-5020 Salzburg, Austria
Regina.Bernhaupt@sbg.ac.at

Jan Gulliksen
Dept. for Human Computer Interaction
Uppsala University, Box 337,
SE-751 05 Uppsala, Sweden
Jan.Gulliksen@hci.uu.se

Peter Forbrig
Department of Computer Science
University of Rostock, Albert-Einstein
Str. 21,18051 Rostock, Germany
pforbrig@informatik.uni-rostock.de

Janet Wesson
Nelson Mandela Metropolitan University
Department of Computer Science &
Information
Systems, PO Box 77000
6031 Port Elisabeth, South Africa
Janet.Wesson@nmmu.ac.za

References

1. Chin, D.N.: Empirical Evaluation of User Models and User-Adapted Systems. User Modelling and User Adapted Interaction (11), 181–194 (2001)
2. Jones, S., Jones, M., Marsden, G., Patel, D., Cockburn, A.: An Evaluation of Integrated Zooming and Scrolling on Small Screens. International Journal of Human-Computer Studies, 2005 (65), 271–303 (2005)
3. Kjeldskov, J., Graham, C.: A Review of Mobile HCI Research Methods. In: Chittaro, L. (ed.) Mobile HCI 2003. LNCS, vol. 2795. Springer, Heidelberg (2003)
4. Masthoff, J.: The Evaluation of Adaptive Systems. In: Patel, N.V. (ed.) Adaptive Evolutionary Information Systems, pp. 329–347. Idea Group Publishing (2002)
5. Brusilovsky, P., Karagiannidis, C., Sampson, D.: The Benefits of Layered Evaluation of Adaptive Applications and Services. In: Bauer, M., Gmytrasiewicz, P.J., Vassileva, J. (eds.) UM 2001. LNCS, vol. 2109. Springer, Heidelberg (2001)
6. Gena, C.: Methods and Techniques for the Evaluation of User-Adaptive Systems. The Knowledge Engineering Review 20(1), 1–37 (2005)
7. Kaikkonen, A., Kallio, T., Kekäläinen, A., Kankainen, A., Cankar, M.: Usability Testing of Mobile Applications: A Comparison between Laboratory and Field Testing. Journal of Usability Studies 1(1), 4–16 (2005)
8. Tamminen, S., Oulasvirta, A., Toiskallio, K., Kankainen, A.: Understanding Mobile Contexts. Journal of Personal and Ubiquitous Computing (8) , 135–143 (2004)

Culture and Technologies for Social Interaction

Qinying Liao[1], Susan R. Fussell[2], Sheetal K. Agarwal[3], Arun Kumar[3], Amit A. Nanavati[3], Nitendra Rajput[3], and Yingxin Pan[1]

[1] IBM China Research Lab. Bldg. 19, Zhongguancun Software Park, 100094 Beijing, P.R. China
{liaoqy,panyingx}@cn.ibm.com
[2] Cornell University, 336 Kennedy Hall, Ithaca, PA 14853 USA
suefussell@gmail.com
[3] IBM India Research Lab., 4, Block - C, Institutional Area, Vasant Kunj, New Delhi 110070 India
{sheetaga,kkarun,namit,rnitendra}@in.ibm.com

Abstract. While social activities and user requirements clearly vary across cultures, we are far from having a systematic understanding of culture's role in the design, development and use of social technologies. This workshop will bring together a mix of HCI researchers and practitioners, social scientists and psychologists who are interested in areas of social technologies and culture, in order to (a) examine the design and use of technologies for social interaction in different cultures to date; (b) explore various viewpoints on the key issues for future research; (c) outline various approaches and identify some basic guidelines for understanding cultural impacts in building social technologies as well as user interfaces and (d) foster new collaborations in the community.

Keywords: Culture, social interaction, social networking sites, online community, social technology, social software.

Recent years have witnessed an increase of social collaborative technologies worldwide, such as discussion groups, blogs, wikis, and social networking sites, resulting in the rapid development of online communities, where people communicate, share information and keep in touch with each other. Designing user interfaces for these technologies requires intensive attention on understanding social interactions among users and acknowledging demographic diversity. In today's global marketplace, these differences may reflect world-wide cultures. Despite its importance, however, most behavioral research on computer-mediated social interactions and many of the technologies developed to support these activities have to date been rooted in Western cultures.

This workshop attempts to bring together researchers, designers and developers with interests in understanding interactions between culture and social technologies. Through collaborative tasks and group discussions, we aim to (a) examine the design and use of technologies for social interaction in different cultures; (b) explore various viewpoints on the key issues for future research; (c) outline various approaches and identify some basic guidelines for understanding cultural impacts in building social technologies and user interfaces; and (d) foster new collaborations in the community.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 957–958, 2009.

Previous research has indicated that culture does indeed influence acceptance of technology [2]. As for technologies that support social activities, such as social networking sites, understanding the impact of culture involves understanding the social behaviors themselves. Think about a social recommendation site where you find information about restaurants. Do you want to see the rankings first or read about the comments or look for the background information of the restaurant to assess its credibility? Different cultures may look for different data to make decisions.

Social research suggests that societies vary along a number of important dimensions that might be expected to influence how and why people use social technologies and what features might be most desirable and/or useful for them. For example, Kavan and Fussell et. al. [3] found that multi-party chat, audio-video chat and emoticons were much more popular in Asia than in North America, reflecting the differences between collectivistic, high-context cultures and individualistic, low-context cultures. Ardichvili et.al. [1] found that modesty requirements as well as high degree of competitiveness among employees were serious barriers to information sharing in China through online communities of practice, but not in Brazil and Russia.

These efforts, however, have to date been scattered and we are far from having a systematic understanding of the concept and role of culture in the use and design of online communities and supporting technologies. To cope with these challenges, this workshop will address multiple issues: (1) the first will be identification of culture-independent and culture-dependent social interactions and technologies, to bring some order to the diverse studies on cultural aspect of social technologies and explore key issues for future research; (2) secondly, we will collectively examine theoretical and methodological challenges, such as the cultural factors to be considered during the design of social software, models required to identify and evaluate the culture's role in dynamic social interactions, measurement instruments and research design that can be developed to permit direct, cross-cultural comparisons, methods that cross-culture theories can be translated into usable features for social software, etc.; (3) we will also consider various approaches by which design recommendations for social interaction can be specified via the dialogues among technologies (developers), user interfaces (designers) and usage patterns (researchers).

Productive collaboration will be emphasized in the workshop through collaborative tasks and group discussions. By the end of the workshop, we will produce a poster for the INTERACT conference as well as a report for publication. A final outcome goal is to strengthen relationships among researchers in the area of culture and human-computer interaction that may lead to new research collaborations. The major target will be the INTERACT, CHI, CSCW communities.

References

1. Ardichvili, A., Maurer, M., et al.: Cultural influences on knowledge sharing through online communities of practice. Journal of Knowledge Management, 10 (2006)
2. Evers, V., Day, D.: The role of culture in interface acceptance. In: Howard, S., Hammond, J., Lindgaard, G. (eds.) Human Computer Interaction: INTERACT 1997, Sydney. Chapman and Hall, Boca Raton (1997)
3. Kayan, S., Fussell, S.R., et al.: Cultural Differences in the Use of Instant Messaging in Asia and North America. In: CSCW 2006, Extended Abstracts (2006)

Design and Evaluation of e-Government Applications and Services (DEGAS'2009)

Marco Winckler[1], Monique Noirhomme-Fraiture[2], Dominique Scapin[3], Gaëlle Calvary[4], and Audrey Serna[4]

[1] University Toulouse 3, Institute of Research in Informatics of Toulouse (IRIT)
118 route de Narbonne, F-31062 Toulouse, France
winckler@irit.fr

[2] Institut d'Informatique, Facultés Universitaires Notre-Dame de la Paix (FUNDP)
Rue Grandgagnage 21, B-5000 Namur, Belgium
mno@info.fundp.ac.be
http://www.info.fundp.ac.be/~mno/

[3] INRIA Paris - Rocquencourt Research Centre, Domaine de Voluceau-Rocquencourt, B.P. 105, Le Chesnay 78153, France
Dominique.Scapin@inria.fr

[4] University Joseph Fourier, Laboratory of Informatics of Grenoble (LIG)
385, rue de la Bibliothèque, BP 53, 38041 Grenoble, France
Gaelle.Calvary @imag.fr

Abstract. The main goal of this workshop is to bring researchers and practitioners together to explore the issues and challenges related to the development of usable and accessible user interfaces for e-Government applications using innovative Information and Communication Technology (ICT). This workshop is the second in a series of workshops organized at IFIP TC 13 Interact conference focused on User Interfaces for e-Government applications. The present edition addresses the emergence of ubiquitous platforms and the multiple access points to e-Government applications. In particular, we are concerned by case studies, theories, applications, and design and evaluation methods for ubiquitous e-Government applications that are committed with the universal access for citizens. DEGAS 2009 is officially supported by the IFIP WG 13.3 on HCI and disability and the IFIP WG 13.2 on Methodologies for User-Centered Systems Design.

1 Workshop Background

The increasing use of the Web as a software platform together with the advance of technology has promoted the Web platform as a suitable alternative for delivering information and services to citizens. However, the adoption of services provided to citizens depends upon how such applications comply to the users needs. The development and implementation of e-government involves consideration of its effects including environmental, social, cultural, educational, consumer issues, among others. On one hand, e-Government software is mandated to follow very strict requirements in terms of evolving regulation, use of legacy technologies, confidentiality protection, and technical constraints related to the management. On the other hand, the design of

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 959–960, 2009.

e-Government applications must consider the impact on the diversity of users in terms of age, language skills, cultural diversity, literacy, and information technologies practice. Some major constraints underline the importance of investment on the User Interface (UI) design and evaluation of e-Government applications:

- The ever growing number of users of e-Government applications calls for a universal access to e-Government applications. Usability has become one of the major challenges for large adoption of many e-services provided to citizens, in particular those suffering from some kinds of disability or having some literacy barriers (e.g. illiterate users, immigrants).
- E-Government applications present several advantages for both front office users (e.g. citizens, associations, companies and so on) and back office people (e.g. government employees, administrative clerks) as they reduce costs of information transfer and treatment. Thus front office and back office users are two sides of the same coin. Whilst universal access should be provided to front office users, usability for back office users should not be neglected as some usability problems could cause errors and/or losses of data that might compromise the quality of the whole system.
- Public administration should ensure multiple access points to e-Government applications allowing home access via Internet broadband, computer-based kiosks, as well as mobile platforms. Nowadays the large variety of computing systems (e.g. low-weight desktop/notebook computers, cell phones, Personal Digital Assistants - PDA, smart phones) has created a milestone for cost-effective development and fast delivery of multi-target interactive systems.
- Citizens are faced with a large set of e-services (e.g., tax declaration, documents requests such as passports, procedures involving health care insurance, etc) which are provided by different governmental agencies. The availability of such e-services generates lots of personal data. As a result, privacy issues as well as management concerns (e.g., sharing data among different applications, recovering personal information from different applications) emerge.
- Bad design can have huge impact not only on the adoption of user interface by users but also compromise the validity of democratic processes. So that, accessibility had become a mandatory requirement for any e-Government initiative.

2 Scope and Objectives

This workshop is intended for anyone (researchers and practitioners) who is concerned with the design and/or evaluation of accessible and/or usable UIs. This includes representatives from administrations, academia (e.g., lecturers in HCI), and policy-making organizations. We aim at triggering a discussion on the topics related to the design and evaluation of ubiquitous e-Government applications, requirements and constraints for the development of e-Government applications, user experience with e-Government services, user involvement into the development process, universal access, policies for implementing accessibility and usability culture into government agencies. Expected outcomes for this workshop include a review of trends and currently available solutions for the design and evaluation of e-Government UIs and a research agenda for further developments in the field.

Designing for Naturally Engaging Experiences

D. Browning[1], M. van Erp[2], M. Bødker[3], N. Bidwell[4], and Truna Aka J. Turner[5]

[1] Discipline of IT, JCU, Townsville, Australia
david.browning@jcu.edu.au
[2] hAAi, voorhaven 25, 3025 HC Rotterdam, Netherlands
marlyn.van.erp@haai.nl
[3] CBS, Center for Applied ICT, Howitzvej 60, 2. Copenhagen, Denmark
mb.caict@cbs.dk
[4] ICT4D, UCT, Cape Town, South Africa
nicola.bidwell@jcu.edu.au
[5] CRC for Interaction Design. QUT (Brisbane), QLD AU
truna@acid.net.au

Abstract. This full day workshop explores how insights from artefacts, created during data collecting and analysis, are translated into prototypes. It is particularly concerned with getting closer to people's experience of shaping a design space. The workshop draws inspiration from data-products resulting from interactions in natural, unbuilt places with the intention of supporting both those with work integrating understandings of such experiences into design and those interested in the way material provokes ideas and inspiration for design.

Keywords: Interaction Design, Place, Representation.

1 Introduction

This workshop explores translating understandings, arising from products created by collecting data about people's experience of places, into prototypes; and, in particular, ways to enable getting close to 'what it feels like' to shape a design space that conveys a sense of 'being there'.

Experience-centred design and designing for affective computing often includes video or photographic records of interactions, data that can 'extend' or deepen our time in the field. Our interactions during interpretation of such data form a bricolage that illuminates the underlying logics of the social practice under review. What we see, as designers, is reflexive, grounded in a cultural perspective from which we describe our interpretation of practices, and the cultural perspective of our audience. Watching a video of a participant's visit to a natural place might cue reflection on the interaction as a 'view from somewhere', but it also has the potential as a 'view from somewhere else'. So, the recorded data itself is a site of continued interaction. Importantly, this interaction adds to the original material and inspires design.

Consider the example of drawing upon recordings of a visitor's familiar experience of a local beauty spot to inspire designing a technology to mediate others' experiences of that or another special places. We have found that Egocentric POV video of a wilderness park in the north Australian tropics depicts some of the memories and

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 961–962, 2009.

meanings visitors associate with it and the ways in which people recreate and augment these meanings in subsequent visits. Whilst digital technology might never transfer the sense of embodiment in a place to those who have no in situ experience, it can be used to considerably augment the otherwise often flat and sterile depictions of place, that are found, for example, in tourist guides.

2 Aims and Objectives

The workshop will explore how designers get from the products of data collection (artefacts, recollections and ideas) to a prototype, focusing on reflexive, ongoing questioning of the values and preconceptions brought into play when understanding the design space. We will invite participants to present their own material and potential design metaphors that will feed into collaborative group design exercises

2.1 Specific Goals/Outcomes

To discuss using new media to depict sensory data of place, particularly in natural places and the impact on experience-centred design; To consider the challenges in informing design with these insights; To explore the remediation produced by practical interaction with participants' materials; To evolve and paper prototype new interaction designs that portray or mediate experiences of places and afford the addition of new experiences; To share insights about data collection in natural places.

2.2 Workshop Program (Full Day)

Prior to the workshop (i) researchers are asked to submit a short (3 pages) paper, describing their use of video and/or photographs and/or audio; and (ii) practitioners to detail their interest and processes (1 page). Please include media examples.

We start with all participants presenting brief examples of their own data. A collaborative review of design methods using situated media will explore using reflective inquiry to illuminate values participants draw upon when understanding design spaces. Hands-on work begins by collectively exercising our design muscles to encourage flexible thinking and inspiration. Group work starts when participants interact with media material and produce initial design responses. In the afternoon, groups will evolve new artefacts and paper prototypes. Lastly, we will reconvene to discuss prototypes and facilitate reflection on the processes unfolding in design and the role of the media in getting close to 'what it feels like' to shape a design space.

3 Organisers

David Browning – natural places interaction researcher working in Australia
Marlyn van Erp – Rotterdam design studio practioner with the Dutch Government amongst her clients.
Mads Bødker – Associate Professor at the Center for Applied ICT, CBS, Copenhagen.
nicola bidwell – Senior Research Fellow in interaction design working in South Africa
truna – games and interaction design researcher and practioner in Australia.

Ethics, Roles and Relationships in Interaction Design in Developing Regions

Anxo Ceriejo-Roibas[1], Andy Dearden[2], Susan Dray[3], Phil Gray[4], John Thomas[5], and Niall Winters[6]

[1] Vodafone Group, [2] Sheffield Hallam University, [3] Dray & Associates,
[4] University of Glasgow, [5] IBM, [6] London Knowledge Lab
anxo-ceriejo-roibas@vodafone.com, a.m.dearden@shu.ac.uk,
susan.dray@dray.com, pdg@dcs.gla.ac.uk,
jcthomas@us.ibm.com, n.winters@ioe.ac.uk

Abstract. A workshop to explore ethical and organisational issues associated with interaction design efforts conducted in developing regions of the world. The workshop will discuss the challenges of conducting this type of work in a way that can bring sustainable benefits to people living in developing regions.

Keywords: International development, ethics, capacity-building, cross-cultural design.

1 Workshop Overview

Recent years have seen a growing interest in the contribution that effective interaction design might make in improving the lives, livelihoods, and life opportunities of people in developing regions (Dray et al, 2003). General workshops on interaction design in developing regions have been held at: CHI 2007, HCI 2007, CHI 2008, DIS 2008, PDC 2008. A panel (Dearden et al., 2007) and two special interest group meeting (Best, et al., 2007, Smith et al., 2007) were held at Interact 2007. These general discussions have raised a variety of important questions, with organisational and ethical issues being particularly salient.

The contribution of interaction design researchers is typically in the context of projects where novel interactive products, systems or services that are designed for (and with) people and organisations in these regions. Participants in such projects hope that their efforts will lead to long term benefits both for the particular communities that they work with, and for other similar groups. However, development is a complex social and political process. Previous eras of technological intervention have had mixed consequences, often generating unexpected side-effects, sometimes disadvantaging the supposed beneficiaries, and occasionally including exploitation. Current debates emphasise the need for external actors to partner with local organisations, to build capacity.

In Interaction Design for Developing Regions, authors have repeatedly stressed the need for researchers who are based in other countries to work with local partners. This raises many issues and questions:

- How do we ensure that interaction design for developing regions is sustainable in local contexts?

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 963–964, 2009.

- How do projects select partners to work with? How can roles be negotiated to create successful systems and to build local capability?
- Participatory approaches are widely regarded as necessary. However, authors have critiqued both the reality of participation and the evidence for its effectiveness (Cooke & Kothari, 2001). What kinds of participation should be supported, and is participatory design always the right approach?
- Projects involve interactions between organizations with very different purposes, principles, cultures and interests. What are the challenges of relationships between local, commercial, and research organisations?
- How does the environment of funding, commercial, personal and institutional interests impact on our ability to make meaningful contributions? What are appropriate roles for individuals, institutions, government, commercial companies and NGOs in this setting?
- Development practices in other fields have much longer traditions and have developed their own ethical frameworks. What can we draw from existing debates about partnership working in development to inform future practice?

In this workshop we shall bring together field experience of interaction design and reflections from development studies to examine these complex issues.

2 Who Should Attend?

The workshop will attract interaction designers working on projects in developing regions, researchers and practitioners based in developing regions, and researchers working in the field development studies, particularly those examining development informatics. We shall be inviting participation particularly from members of IFIP Working Group 9.4 on the Impact of Computers in Developing Countries. We shall be seeking funds from external sources to enable researchers and practitioners from developing regions to attend the workshop.

References

1. Best, M., Dearden, A., Dray, S.M., Light, A., Thomas, J.C., Buckhalter, C., Greenblatt, D., Krishnan, S., Sambasivan, N.: Sharing Perspectives on Community-Centered Design and International Development. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4663, pp. 696–697. Springer, Heidelberg (2007)
2. Cooke, B., Kothari, U. (eds.): Participation: The New Tyranny? Zed Books, London (2001)
3. Dearden, A., Dunckley, L., Best, M., Dray, S.M., Light, A., Thomas, J.: Socially Responsible Design in the Context of International Development. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4663, pp. 694–695. Springer, Heidelberg (2007)
4. Dray, S., Siegel, D., Kotzé, P.: Indra's Net: HCI in the Developing World. Interactions 10(2), 28–37 (2003) (Special Issue on HCI in the Developing World)
5. Smith, A., Joshi, A., Liu, Z., Bannon, L., Gulliksen, J., Baranauskas, C.: Embedding HCI in Developing Countries: Localizing Content, Institutionalizing Education and Practice. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4663, pp. 698–699. Springer, Heidelberg (2007)

Human Aspects of Visualization

Achim Ebert[1], Alan Dix[2], Nahum Gershon[3], and Margit Pohl[4]

[1] German Research Center for Artificial Intelligence (DFKI), Kaiserslautern, Germany
ebert@dfki.de
[2] Lancaster University, Lancaster, UK
alan@hcibook.com
[3] The MITRE Corp., McLean, VA, USA
gershon@mitre.org
[4] Vienna University of Technology, Vienna, Austria
margit@igw.tuwien.ac.at

1 Goals and Issues

The quantity of data available in our modern information society is ever-growing. It is becoming unfeasible for any one person to oversee, much less understand the overwhelming amount of unstructured, multidimensional information he or she is confronted with every day. The human visual system, the "mind's eye", relies in large part on the eye and on the processing and the interpretation of the information processed by the brain. Visualization utilizes both.

Interactive visualization methods can increase cognitive resources by providing an additional, external visual resource to the human memory. They may reduce the amount of searching and ease the recognition of patterns as well as enhance understanding of relationships in large amounts of data and information. In addition, they provide a medium that enables the user to have a representation of information that he or she can quickly and easily modify, restructure or consider from a different perspective. This ability to manipulate the data is of extreme importance especially for analytical reasoning and sense-making. Virtually all sciences and many industries rely on this ability to identify methods and models, which can turn data into reliable and provable knowledge.

Research in HCI and visualization has produced numerous forms of visualization and methods for interacting with visualizations of large data sets. There is longstanding design advice for both static and dynamic visualization, and many user studies of particular visualization and interaction methods. However, systematic understanding of the interaction issues with visualization and the human cognitive and perceptual processes involved is perhaps less well developed. Knowledge about human perception can help to define a more systematic framework informing the design of Information Visualization tools.

Most visualizations involve some level of computation, from proportions in pie charts, to time series analysis in sales forecasting. Sometimes this is for data reduction, sometimes to reveal new or different aspects of data that are hard to see with the raw data. In complex domains, such as business strategy, environmental monitoring, network management, and security, more complex forms of processing are being used in conjunction with visualization; this area where visualization and complex data

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 965–966, 2009.

manipulation meet has become known as visual analytics. Even simple forms of data manipulation such as averages or clustering have issues of interpretation (e.g. visualizing uncertainty, loss of outliers), but as more complex statistical or machine learning techniques are involved how can users make sense of the derived data or control the parameters of the underlying computational processes? On the other hand, appropriate visualizations have the potential to help users make sense of these complex algorithms.

In addition, the increasing graphical processing power of PCs has fuelled a powerful demand for larger and more capable display devices whilst, the proliferation of mobile devices requires visual applications on very small screens. With these small- or large-scale displays, assumptions of the normal desktop metaphor no longer apply and mice and keyboards, designed for desktop interaction, become unsatisfactory; all calling for novel interactions devices and techniques.

2 The Workshop

This workshop, co-organized by IFIP WG TC13.7 on "HCI and Visualization" and the European VisMaster Coordinated Action, seeks to survey and expand our understanding of the cognitive and perceptual issues of interactive visualization and visual analytics and to bring together researchers interested in these issues. It will outline the research required to understand which aspects of analysis most closely match human capabilities, and how interactive visual support should be designed and adapted to make optimal use of human capabilities, in terms of information perception and processing. This is a challenging agenda and will need to take into account many factors including user role, tasks, collaborations, interests and previous knowledge.

Topics of interest include: models of visualization and interaction, perceptual aspects (visual discrimination, gestalt, etc.), visualization of machine learning, visualizing uncertainty, studies of use of visualization systems, distributed cognition and visualization, social and collaborative aspects of visualization and interaction, organizational and political context of visualization, display technology, interactive mobile visualization, usability and evaluation aspects.

The workshop goals require the cooperation of scientists from different disciplines addressing aspects of visualization and analysis. This will enable a mapping between foundational theories in the different research areas to support ongoing collaboration between researchers in these fields. This will include those already working at the intersection of HCI, Visualization and Visual Analytics and also related areas such as data mining, data management, perception and cognition.

Further information both before and after the workshop can be found on the workshop website at http://www.hciv.de/interact09.

Innovation for an Inclusive Future

Mark Springett[1], Mark Rice[2], Alex Carmichael[2], and Richard Griffiths[3]

[1] Interaction Design Centre, Middlesex University Town Hall,
Hendon, London, NW4 4BT, UK
m.springett@mdx.ac.uk
[2] School of Computing, University of Dundee Dundee, DD1 4HN
{mrice,acarmichael}@computing.dundee.ac.uk
[3] School of Computing, Mathematical & Information Sciences, University of Brighton Watts Building, Moulsecoomb, Brighton, BN2 4GJ, UK
r.n.griffiths@brighton.ac.uk

1 Workshop Objectives

This workshop will focus on setting the agenda for research, practice and policy in support of inclusive design for third generation computer-based products. The next generation of technology represents an unprecedented opportunity to improve the quality of life for groups of users who have previously faced exclusion, such as those with impairments and older citizens. At the same time it risks creating a greater digital divide and further exclusion. How we approach design for this new generation will determine whether or not the third wave will provide positive advances towards an inclusive digital world. We therefore need to put forward both a rationale for inclusive design and provide pointers towards technical development and design practice in support of inclusion. It is our belief that there is not only a strong moral case for design for inclusion but also significant commercial incentive, which may be key to persuading influential players to focus on inclusion. Therefore one of our key objectives is to describe and promote the advantages of designing ‘in from the edges’ of the user population rather than designing for a notional ‘average’ user.

The gradual convergence of TV and PC-based services, and the greater interactivity of TV, affords the opportunity to re-think how we can best interact and design services, suitable devices and integration into domestic living spaces and with divergent lifestyles. This convergence and increased capability for service design facilitates creative innovation in which the space of possible applications can be explored. Advances include, but are not limited to, more inclusive provision of entertainment and communication applications associated with interactive TV and Web 2.0. There is a further agenda afforded by the genesis of innovative solutions for independent living, lifelong learning and social inclusion that exploit the potential of TV, mobile, satellite and network technologies.

Yet despite this potential there is an unresolved paradox in the development of digital interactive television and other information appliances within a private, domestic sphere. For example, the replacement of specialized analogue by generalized digital technology has brought about a step change in the flexibility with which broadcast services are consumed. There is consequently an (in principle) improvement in accessibility for disabled audience members. In practice however, there is an increasing

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 967–968, 2009.

exclusion of segments of the audience. For example, for people with visual impairments, there is an increased need to read text and interpret graphical displays. For people with motor disabilities there is increased need to make fine manipulations of complex remote controls. For people with cognitive disabilities there is an increased need to make sense of a complex communications technology; one by which they may enter into contractual obligations with a range of programme, service and product providers.

These issues are a reflection in the push towards supporting the youth entertainment market and other markets where early adopters are likely to provide a healthy return from those investing. Consequently, as with previous technology advances there are also groups that are vulnerable to being ignored or left behind as technology and applications advance – despite the potential/promise this new technology can provide. Design has typically targeted the 'ideal user' with the effect that citizens who are dissimilar to that profile are effectively excluded from design thinking. This can be seen as a significant factor in the growing 'digital divide', where the gap between technology adopters and those left behind increases as technology advances.

Design for inclusion potentially rides more easily than ever before with the goals of mainstream end-user technology innovation and development. There is enormous scope for innovation both for enhancement to existing services and exploration of the space of possible new services. However, the conditions and attitudes must be right for this to become a reality. For example, enablement of better inclusion for digital TV services requires development of open and standardised television architectures. This would facilitate the development of a secondary market in accessible technology, perhaps through the adoption of solutions already available in the personal computer market, and perhaps of wholly new solutions specialised for television. This approach has been recognised in the EU FP7 research programme with a strand dedicated to 'Open Systems Reference Architectures, Standards and ICT Platforms for Ageing Well' (Objective ICT-2009.7.1 ICT & Ageing, Target Outcome b.)[1].

Reference

1. European Commission – Information Society and Media: Information and Communication Technologies, Work Programme 2009-2010 (2008)

Interplay between Usability Evaluation and Software Development (I-USED 2009)

Silvia Abrahão[1], Kasper Hornbæk[2], Effie Law[3], and Jan Stage[4]

[1] ISSI Group, Universidad Politécnica de Valencia
c/ Camino de Vera s/n, 46022 Valencia, Spain
sabrahao@dsic.upv.es
[2] University of Copenhagen, Denmark
Universitetsparken 1, DK-2100 Copenhagen OE, Denmark
kash@diku.dk
[3] Institut TIK, ETH Zürich, Switzerland
law@tik.ee.ethz.ch
[4] Aalborg University, Department of Computer Science, Selma Lagerlofs Vej 300
DK-9220 Aalborg East, Denmark
jans@cs.aau.dk

Abstract. This workshop is aimed at bringing together researchers and practitioners from the Human-Computer Interaction (HCI) and Software Engineering (SE) fields to determine the state-of-the-art in the interplay between usability evaluation and software development and to generate ideas for new and improved relations between these activities. The aim is to base the determination of the current state on empirical studies. Presentations of new ideas on how to improve the interplay between HCI & SE to the design of usable software systems should also be based on empirical studies.

Keywords: software development, user interface development, usability assessment, usability evaluation methods, empirical studies.

1 Introduction

Software development is highly challenging. Despite many significant successes, several software development projects fail completely or produce software with serious limitations, including (1) lack of usefulness, i.e. the system does not adequately support the core tasks of the user, (2) unsuitable designs of user interactions and interfaces, (3) lack of productivity gains or even reduced productivity despite heavy investments in information technology [2] [6] [5].

Broadly speaking, two approaches have been taken to address these limitations. The first approach is to employ evaluation activities in a software development project in order to determine and improve the usability of the software, i.e. the effectiveness, efficiency and satisfaction with which users achieve their goals [3]. To help software developers' work with usability within this approach, more than 20 years of research in Human-Computer Interaction (HCI) has created and compared techniques for evaluating usability. The second approach is based on the significant advances in techniques and methodologies for User Interface (UI) design that have been achieved in the last decades (e.g., participatory design and user-centered design). However, the

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 969–970, 2009.

Software Engineering (SE) community has recognized that usability does not only affect the design of user interfaces but the software system development as a whole. Specifically, efforts are focused on explaining the implications of usability for requirements gathering [4], software architecture design [1], and the selection of software components [7].

The interplay between these two fields, and between the activities they advocate to be undertaken in software development, have been limited. Integrating usability evaluation at relevant points in software development with successful and to-the-point results has proved difficult. Research in HCI and SE has been done mainly independently of each other with no in substantial exchange of results and sparse efforts to combine the techniques of the two approaches. As a result, the state of industry practice is still quite immature and even in some cases the basic principles and currently available Usability Evaluation Methods (UEMs) are not understood.

The goal of the second edition of the workshop is to bring together researchers and practitioners from the HCI and SE fields to determine the state-of-the-art in the interplay between usability evaluation and software development and to generate ideas for new and improved relations between these activities. The aim is to base the determination of the current state on empirical studies. Within this focus, topics of discussion include, but are not limited to:

- Which artifacts of software development are useful as the basis for usability evaluations?
- How do the specific artifacts obtained during software development influence the techniques that are relevant for the usability evaluation?
- In which forms are the results of usability evaluations supplied back into software development (including the UI design)?
- Do existing UEMs deliver the results that are needed in user interface design?
- How can usability evaluation be integrated more directly in UI design?
- How can UEMs be applied in emerging techniques for UI design?
- How can usability evaluation methods be integrated to novel approaches for software development (e.g., model-driven development, agile development).

References

1. Bass, L., John, B.: Linking Usability to Software Architecture Patterns through General Scenarios. Journal of Systems and Software 66(3), 187–197 (2003)
2. Brooks Jr., F.P.: No Silver Bullet: Essence and Accidents of Software Engineering. IEEE Computer 20, 10–19 (1987)
3. ISO, Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11: Guidance on Usability (1998)
4. Juristo, N., Moreno, A.M., Sanchez-Segura, M.: Guidelines for Eliciting Usability Functionalities. IEEE Transactions on Software Engineering 33(11), 744–758 (2007)
5. Landauer, T.K.: The trouble with computers: usefulness, usability, and productivity. MIT Press, Cambridge (1995)
6. Mathiassen, L., Stage, J.: The Principle of Limited Reduction in Software Design. Information Technology & People 6(2-3), 171–185 (1992)
7. Perry, D., Wolf, A.: Foundations for the Study of Software Architecture. ACM Software Eng. Notes 17(4), 40–52 (1992)

New Challenges for Participation in Participatory Design in Family, Clinical and Other Asymmetrical, Non-work Settings

Olav W. Bertelsen[1] and Per-Olof Hedvall[2]

[1] Dept. of Computer Science, Aarhus University, Aabogade 34, DK-8200 Aarhus N, Denmark
olavb@cs.au.dk
[2] CERTEC, Lund University, P.O. Box 118, SE-221 00 Lund, Sweden
per-olof.hedvall@certec.lth.se

Abstract. Participatory design (PD) has taken as its ideal that designers and users should engage in an equal language game. When we apply PD in contexts where some of the users involved are weak, ill, or have impairments, this assumed equality can no longer be an ideal. The workshop explores new ideals for participatory design in non-work settings with highly heterogeneous user constellations.

Keywords: Participatory Design, Non-Work Setting, Asymmetry, Heterogeneity.

1 About the Workshop

Participatory design has the point of view that designers and users should engage in an equal language game. The ambition has been that equality in design is possible to achieve by creating certain conditions for the design activities. Whether or not this ideal has ever been met is a big question, but when PD operates in contexts where some of the involved users are weak, ill, or have impairments, we are clearly confronted with a situation where this assumed equality is no longer an ideal that can be reached. Examples of such contexts are:

Premature newborns. Grönvall et al. [1] have worked with the design of IT enhanced environments for premature infants. The doctors and nurses are at work and have clear perspectives of the therapeutic strategies they want the environment to support; the parents are in the middle of a severe life crisis of which they probably do not have any previous experience; the child is fighting for survival. They are all users but can be involved in design in very different ways.

Phobic patients. Bering [2] designed a mobile device to support cognitive therapy for phobic patients. The fragility of the patients made it difficult to involve them, and any exposure to the triggers of phobic reaction would interfere with their therapeutic programs.

Families that have children with complex communication needs. Children with these needs are often challenged by asymmetrical terms for communication and

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 971–972, 2009.

interaction in everyday family life [3]. This affects the whole family. Thus, even when the child is be the potential “user” of a design, the entire family needs to be engaged in the design process.

Erotic intimacy. As pointed out by Bertelsen and Petersen [4], the increased presence of information technology in the private sphere has become a limiting factor for erotic intimacy today. Designing solutions for this problem is a challenge to established methods in participatory design.

In early participatory design, it made sense to take a power struggle perspective and aim for user-empowerment in the design of better computer support. As contexts change, however, other perspectives should be introduced to maintain participatory design as a realistic pragmatic design approach. Classical activities developed in participatory design, such as mock-ups, organizational games and other activities that rely on breakdowns in the flow of simulated action as a source of shared sense making may not be useful. Instead new activities such as diary studies, fictitious users, therapists as mediators, various forms of probing, and completely new forms of participation have to be invented.

In developing an understanding of participatory design in these contexts, new and old theoretical frameworks may be useful. Activity theory has been used to make sense of early participatory design [5] and may have potential in this context. Actor network theory, distributed cognition and communities of practice are other examples that can be revisited. This “new participatory design” is not only a challenge to participatory design but is also a potentially transformative resource for some therapeutic, family and care-oriented domains. Furthermore, it contributes to the fields of accessibility and impairments [6].

The workshop seeks to approach new participatory design by inviting position papers, informative design cases, techniques, methods and theoretical analyses. The workshop will be organized as a mixture of presentations, discussions and various breakout activities depending on the number of participants and the submitted work. Proceedings from the workshop will be edited and published after the meeting [7].

References

1. Grönvall, E., Marti, P., Pollini, A., Rullo, A., Bertelsen, O.W.: Palpable Time for Heterogeneous Care Communities. In: Proc. Crit. Comp. ACM Press, New York (2005)
2. Bering, P.F.: Cognitive-behavioral Therapy using Mediating Instruments. Unpublished master’s thesis, Department of Computer Science, Aarhus University (2005)
3. Beukelman, D.R., Mirenda, P.: Augmentative & Alternative Communication – supporting children & adults with complex communication needs. Brookes, Baltimore (2005)
4. Bertelsen, O.W., Petersen, M.G.: Erotic Life as a New Frontier in HCI. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796. Springer, Heidelberg (2007)
5. Bødker, S.: Through the Interface. LEA (1991)
6. Hedvall, P.-O.: Xings for augmented family Communication. Full paper and poster presented at ISCAR 2008, San Diego (2008)
7. `http://www.daimi.au.dk/~olavb/NewPD/`

New Sociotechnical Insights in Interaction Design

José Abdelnour-Nocera[1] and Anders I. Mørch[2]

[1] Centre for Internationalisation and Usability, Thames Valley University
St Mary's Road, London, UK, W5 5RF
Jose.abdelnour-nocera@tvu.ac.uk
[2] InterMedia, University of Oslo
P.O. Box 1161 Blindern, N-0318 Oslo, Norway
anders.morch@intermedia.uio.no

Abstract. New challenges are facing interaction design. On one hand because of advances in technology – pervasive, ubiquitous, multimodal and adaptive computing – are changing the nature of interaction. On the other, web 2.0, massive multiplayer games and collaboration software extends the boundaries of HCI to deal with interaction in settings of remote communication and collaboration. The aim of this workshop is to provide a forum for HCI practitioners and researchers interested in knowledge from the social sciences to discuss how sociotechnical insights can be used to inform interaction design, and more generally how social science methods and theories can help to enrich the conceptual framework of systems development and participatory design. Position papers submissions are invited to address key aspects of current research and practical case studies.

1 Introduction

One of the biggest challenges for HCI and CSCW is addressing the ongoing tensions created by the gap between social requirements and the affordances of technical design [1]. The translation of social knowledge into design decisions is not a simple problem, but one that requires a redefinition of disciplinary boundaries and the subject and object of interaction design. Addressing this socio-technical gap requires a fresh look at how diverse areas of the social sciences explore and conceptualize the relation between people, society and technology under the rubric of 'sociotechnical'. While organizational studies of technology adoption have a well-defined conceptual framework known as sociotechnical systems theory with established principles [e.g. 2] the situation is not the same for interaction design research. While perspectives like ethnography, ethnomethodology [3] and activity theory [4] have problematized and stimulated the design of interactive systems, the potential contribution of other social science perspectives such as American Pragmatism [e.g. 5,6] or the Sociology of Technology [e.g. 7] are beginning to appear, providing new experiences to extend the existing work.

This workshop aims to gather researchers in HCI and CSCW who work on expanding the connection between social science research and interaction design practice

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 973–974, 2009.

under the umbrella 'sociotechnical'. The challenge is to achieve some level of 'translation' between the two domains of discourse in order to improve the design of interactive systems. This workshop is the second in a series of workshops that started in London in 2008 as an event jointly organized by the Interactions and Sociotechnical specialist groups of the British Computer Society. The workshop led into a special issue in the International Journal of Sociotechnology and Knowledge development in 2009 [8].

2 Goals and Topics

The goal is to enable new translations from the social sciences to interaction design. The topics include, but are not limited to, the following areas: overview of related work in HCI and CSCW; critiques of earlier approaches to design; related work on sociotechnical design (e.g. participatory design, organizational informatics); actionable recommendations and guidelines for the conception, design and evaluation of interactive systems as 'social proxies'; improved methods for the gathering and elicitation of 'social requirements'; identifying socially responsible policies for interaction design; interaction design for web technology and social networking (e.g. web 2.0); using theories from the social sciences to inform interaction design; understanding participatory design as a sociotechnical endeavor in software engineering (e.g. agile methods; end-user development). It is hoped this workshop will foster dialogue between academics in different disciplines (e.g. HCI, CSCW, Sociology, Psychology, Anthropology, Marketing, Software Engineering, Ergonomics, Education, Information Systems) interested in interdisciplinary research in interaction design.

References

1. Ackerman, M.: The intellectual challenge of CSCW: the gap between social requirements and technical feasibility. Hum.-Comput. Interact. 15, 179–203 (2000)
2. Mumford, E.: The ETHICS approach. Commun. ACM 36, 82 (1993)
3. Dourish, P., Button, G.: On "Technomethodology": Foundational Relationships Between Ethnomethodology and System Design. Human-Computer Interaction 13, 395–432 (1998)
4. Kaptelinin, V., Nardi, B.A.: Acting with Technology: Activity Theory and Interaction Design. MIT Press, Cambridge (2006)
5. Barnes, S.: The Contemporary Relevance of George Herbert Mead's Social Psychology and Pedagogy. Philosophical Studies in Education 33, 55–63 (2002)
6. Mørch, A.I.: From Pragmatism to Interaction Design: A Sociotechnical Design Space. International Journal of Sociotechnology and Knowledge Development, Special issue on Sociotechnical Systems and Interaction Design 1(2), 8–22 (2009)
7. Abdelnour-Nocera, J., Dunckley, L., Sharp, H.: An Approach to the Evaluation of Usefulness as a Social Construct using Technological Frames. International Journal of Human Computer Interaction, Special Issue about Field Research Methods 22(1), 157–177 (2007)
8. Abdelnour-Nocera, J., Eason, K., Beale, R.: New Sociotechnical Insights in Interaction Design. International Journal of Sociotechnology and Knowledge Development 1, i–iii (2009)

Team Meetings within Clinical Domains – Exploring the Use of Routines and Technical Support for Communication

Kristina Groth[1,2], Ann Lantz[1], Eva-Lotta Sallnäs[1], Oscar Frykholm[1], and Anders Green[1]

[1] HCI, CSC, Royal Institute of Technology, Stockholm, Sweden
[2] CLINTEC, Karolinska Institutet, Stockholm, Sweden
{kicki,alz,evalotta,frykholm,green}@csc.kth.se

1 Introduction

Today, it is common that a team of clinicians, from different disciplines, instead of one single doctor, care for a patient. This is especially true when it concerns more complicated diseases in highly specialised health care. Going from one doctor to a team of doctors raises new dimensions/problems/issues when deciding about the diagnosis and how to treat the patient. Instead of one person deciding, based on the information given from others, a group of people need to agree on a decision. How do the participants during such decision meetings argue for their experience and skill? What kind of technologies are available and how do they support the communication in the meeting? Måseide (2006), for example, focuses on how different forms of evidence influence and regulate the judgements and decisions of medical practitioners during such meetings. Groth et al. (2008), for example, focuses on the technology used during such meetings, with a focus on audio, video, and images.

It has been shown that this kind of multi-disciplinary team meetings increase patient safety (Burton et al. 2006) and that multi-disciplinary site-specialist teams improve processes and outcomes of care (Morris et al. 2006). When comparing operations for liver metastases, Burton et al. reported that decisions made at such meetings showed a higher precision when judging the part of the liver that needs to be removed during the operation.

The objective of this workshop is to gather people with an interest in different aspects on the meeting process in clinical domains. The organisers of the workshop are involved in a project working with technical support in the meeting process at a gastro clinic focusing on highly specialised care. The aim of our project is to explore the use of multi-modal technologies in advanced medical processes such as meetings where physicians from different disciplines meet to decide how to treat a patient.

One focus in our project is to find possibilities to provide relevant information about the patient case discussed during the meeting. For example, if an identified tumour behaves in a specific way, then this together with the patient's general status might be of interest to find similar, previous cases. Cross-referencing of patients, treatment and outcome is today more or less only based on individual physicians' experience and remembrance, not formalised or easily available for others. Another

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 975–976, 2009.

focus in our project is to see how haptic feedback can be used in a discussion where radiologist and surgeons discuss a patient case (cf. Anderlind 2008). Radiologists are used to look at radiology pictures and to communicate the radiological diagnosis. The information discussed when looking at the radiology pictures could be further enhanced through other modalities, like touch, in order to better understand the implications for treatment.

The focuses of our project may serve as examples of what we find interesting to discuss in this workshop.

2 Issues and Goals

The major focus on the workshop is on team meetings in clinical domains, where several clinicians meet and discuss medical issues. Such meetings may take place both within premises and with participants from remote sites, and may include participants from one or several disciplines. The workshop focuses on the following themes, all in relation to the communication process during team meetings: multimodal interfaces, visualisation of information, participants' interaction with information, haptic interfaces, mediated communication, and other technologies that can support the outcome of the meeting.

Apart from the poster generated as an outcome of the conference, other expected outcomes of this workshop are three-folded: to initiate a network within this area of research, to get a broader perspective of the state of the art research within this area, to initiate future collaborations, organise other workshops or research meetings with the aim of producing prototypes, test scenarios and/or publications.

References

1. Burton, S., Brown, G., Daniels, I.R.: MRI directed multidisciplinary team preoperative treatment strategy: the way to eliminate positive circumferential margins? Br. J. Cancer 94 (2006)
2. Groth, K., Olin, K., Gran, O., Permert, J.: The role of technology in video-mediated consensus meetings Journal of telemedicine and e-health 14(7) (2008)
3. Morris, E., Haward, R.A., Gilthorpe, M.S., Craigs, C., Forman, D.: The impact of the Calman-Hine report on the processes and outcomes of care for Yorkshire's colorectal cancer patients. Br. J. Cancer 95, 979–985 (2006)
4. Måseide, P.: The deep play of medicine: Discursive and collaborative processing of evidence in medical problem solving. Communication & Medicine 3(1), 43–54 (2006)
5. Anderlind, E., Noz, M.E., Sallnäs, E.-L., Bengt, K., Lind, B.K., Maguire Jr., G.Q.: Will haptic feedback speed up medical imaging? An application to radiation treatment planning. Journal of Acta Oncologica 47(1), 32–37 (2008)

Touch Affordances

Karin Slegers[1], Dries De Roeck[1], and Timo Arnall[2]

[1] Centre for User Experience Research, Parkstraat 45, 3000 Leuven, Belgium
{karin.slegers,dries.deroeck}@soc.kuleuven.be
[2] Oslo School of Architecture and Design, Maridalsveien 29, 0175 Oslo, Norway
timo.arnall@aho.no

Abstract. The workshop "Touch Affordances" addresses a concept relevant to human computer interactions based on touch. The main topic is the challenge of applying the notion of affordances to domains related to touch interactions (e.g. (multi)touch screens, RFID & NFC, ubiquitous interfaces). The goals of this workshop are to launch a community of researchers, designers, etc. interested in this topic, to create a common understanding of the field of touch affordances and to generate ideas for new research areas for intuitive touch interactions. The workshop will be highly interactive and will have a creative, generative character.

Keywords: Affordances, HCI, intuitive design, NFC, RFID, (multi)touch interface, touch experience, ubiquitous interface.

1 Introduction

The focus of this workshop is on possible fields and domains where 'touch' as a means of human computer interaction is, or could be implemented. For the purpose of this workshop, the term "touch interaction" is used to refer not only to (multi)touch screens, but also to RFID and NFC related technologies, to gesture –based interactions and ubiquitous interfaces because these technologies share the basic interactions of touching, grasping, making hand movements etc.

The topic of this workshop is "Touch affordances". While on the one hand touch-based interactions require very simple actions, on the other hand, touch-based interfaces often are unfamiliar to users and therefore do not invite them to perform the correct actions. Findings of previous studies suggest that unfamiliarity with touch-related interfaces can indeed result in difficulties with regard to interaction (e.g. Belt, Greenblatt, Häkkilä & Mäkelä, 2006; Peltonen, Kurvinen, Salovaara, et al., 2008). This topic leads to the question of how to design simple yet intuitive touch interfaces.

There have been some initiatives to design intuitive graphics to clarify the intended touch-based interaction to the user (e.g. Arnall, 2006; Saffer, 2008). When talking about "touch affordances" within the context of this workshop, we are looking beyond signage and items on a screen that invite a person to 'touch' it. Instead, touch affordances are envisioned in a much broader way. Areas such as ubiquitous computing, haptics, presence, experience research and human computer confluence play a major role in this.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 977–978, 2009.

2 The Workshop

This workshop aims to 1) bring together people who are active in the topic of the workshop to create a community; 2) come to a shared view on the field of Touch Affordances; and 3) generate ideas for new research areas for innovative (and intuitive) touch interfaces and interactions.

The one day workshop will have a creative, generative character. It will consist of four parts described below. During the workshop, observants will live blog about the workshop. A tag cloud will be generated based on these blogs.

Session 1: Keynote & wrap-up (1.5 hours, plenary session) A keynote presentation will be given by Timo Arnall, who will provide his vision on the topic of the workshop. After the keynote, the organizers of the workshop will give an overview of the submitted position papers.

Session 2: Introduction to creative session & plenary brainstorm (1 hour, plenary session) Challenges and questions of interest identified in Session 1 will be the starting point for a generative group discussion. The main goal of this discussion will be to set up a number of design scenarios for which touch interfaces will be relevant.

Session 3: Creative session – developing lo-fi touch affordance prototypes (1.5 hours, four groups) Participants will be divided into four groups, based on their interest in one of the scenarios identified in Session 2. These groups will each develop a lo-fi prototype of a touch affordance for one of the scenarios. The workshop organizers will provide materials to create such prototypes. Each group will make a visual representation of their prototype.

Session 4: Testing prototypes collaboratively & creation of poster (2 hours, plenary session) Each group presents the prototype that was developed. The intuitiveness of each prototype will be assessed by conducting interaction tests with other workshop participants. Next, the participants will collaborate to create a poster summarizing the main results of the workshop. The poster will include the four prototype representations, as well as a tag cloud based on the bloggers' reports of the workshop. On the poster, the scenarios, the prototypes and each participant will be linked to relevant tags in order to create a visual map of the field of Touch Affordances and its community.

References

1. Arnall, T.: A graphic language for touch-based interactions. In: Workshop Mobile Interaction with the Real World (MIRW 2006) in conjunction with MobileHCI 2006, Espoo, Finlands (2006)
2. Belt, S., Greenblatt, D., Häkkilä, J., Mäkelä, K.: User Perceptions on Mobile Interaction with Visual and RFID Tags. In: MIRW 2006, Workshop W5 @ MobileHCI 2006, Espoo, Finland (2006)
3. Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., Oulasvirta, A., Saarikko, P.: It's Mine, Don't Touch: Interactions at a Large Multi-Touch Display in a City Centre. In: Proc. CHI 2008, Florence, pp. 1285–1294. ACM Press, New York (2008)
4. Saffer: Designing Gestural Interfaces. O'Reilly Media Inc., Sebastopol (2008)

Towards a Manifesto for Living Lab Co-creation

Asbjørn Følstad[1], Petter Bae Brandtzæg[1], Jan Gulliksen[2,3], Mikael Börjeson[4], and Pirjo Näkki[5]

[1] SINTEF, Oslo, Norway
[2] University of Uppsala, Sweden
[3] Royal Institute of Technology, Stockholm, Sweden
[4] CDT at Luleå University of Technology, Sweden
[5] VTT Technical Research Centre of Finland, Espoo, Finland
{asf,pbb}@sintef.no, Jan.Gulliksen@it.uu.se, Mikael.Borjeson@ltu.se, Pirjo.Nakki@vtt.fi

Abstract. There is a growing interest in Living Labs for innovation and development in the field of information and communication technology. In particular there seem to be a tendency that current Living Labs aim to involve users for co-creative purposes. However, the current literature on Living Lab co-creation is severely limited. Therefore an Interact workshop is arranged as a first step towards a manifesto for Living Lab co-creation.

Keywords: Living Labs, co-creation, workshop.

1 Co-creation in Living Labs

Living Labs is an approach to user-centred innovation and development, where the users are involved within a familiar context; preferably their every-day environment. Living Labs has lately generated a great deal of interest in Living Labs within the field of information and communication technology (ICT); particular seen in the near explosive growth in the European Network of Living Labs (ENoLL) which now includes more than one hundred Living Labs across Europe [1].

The concept of Living Labs is still evolving, but an important emerging trend is to see Living Labs as a way to tap into the creative potential of the potential users of the innovation. Instead of being recipients of the outcome of innovation and development, users may be engaged in co-creative innovation processes of a Living Lab [2]. Co-creation typically is seen as creative collaboration between users, developers and stakeholders.

ENoLL have been particularly clear about the importance of co-creation in Living Labs [1]. Co-creation may be seen in relation to user-driven innovation and open innovation, even though these concepts are only partially overlapping.

We are only at the beginning of exploring how Living Lab participants may be involved in targeted and beneficial co-creation processes, and the current state-of-the-art is severely limited. A recent literature review showed that co-creation is a Living Lab characteristic reflected in about half the existing literature on Living

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 979–980, 2009.

Labs. However, even though co-creation clearly is an emerging Living Lab trend, there is – apart from a few notable exceptions (e.g. [3] and [4]) – a lack of descriptions of processes and method supporting Living Lab co-creation [5]. The interest in Living Lab co-creation is dramatically increasing, but at present there exist no explicated set of processes and methods to support such co-creation.

2 Workshop Objective and Relevance to the Field

The workshop objective is to explore Living Lab practices for co-creation, and on this basis develop a manifesto for Living Lab co-creation. Currently, the experience and knowledge of Living Lab co-creation seem to reside among Living Lab practitioners rather than researchers, and it is particularly important to engage Living Lab practitioners in the workshop.

The main idea of the workshop organization is to use the preparatory period to conduct an exploration of Living Lab co-creation practices, and then to reach a draft manifesto for Living Lab co-creation at the workshop itself. The draft manifesto will be refined after the workshop, in dialogue with all contributors.

The workshops' relevance to the HCI-field is high due to the increasing importance of Living Labs for user-centered innovation and development. HCI research and practice can offer valuable methods for Living Labs, whereas Living Labs can provide a fresh innovation-centered approach that has not been in the focus in the HCI field traditionally. The workshop ambition to explicate the experiences and knowledge of practitioners may potentially have important implications also for the research conducted in the field of HCI.

In order for the work of both Living Lab practitioners and HCI researchers to be adequately adapted to the requirements of co-creation in Living Labs, it is necessary to create meeting places like this workshop to exchange experiences and develop new knowledge.

References

1. European Network of Living Labs: Open Living Labs, `http://openlivinglabs.eu`
2. Eriksson, M., Niitamo, V.-P., Kulkki, S., Hribernik, K.A.: Living Labs as a Multi-Contextual R&D Methodology. The 12th International Conference on Concurrent Enterprising, ICE 2006(2006)
3. Pierson, J., Lievens, B.: Configuring living labs for a 'thick' understanding of innovation. In: The Ethnographic Praxis in Industry Conference, EPIC 2005 (2005)
4. Näkki, P., Antikainen, M.: Online tools for co-design: User involvement through the innovation process. In: The NordiCHI 2008 workshop, How can HCI improve social media development?, pp. 92–97. Tapir akademisk forlag, Trondheim (2008)
5. Følstad, A.: Living Labs for innovation and development of information and communication technology: A literature review. Electronic Journal of Organizational Virtualness 10, 99–131 (2008) (special issue on Living Labs)

User Experience Evaluation Methods in Product Development (UXEM'09)

Virpi Roto[1,2], Kaisa Väänänen-Vainio-Mattila[2,1], Effie Law[3], and Arnold Vermeeren[4]

[1] Nokia Research Center, P.O.Box 407, 00045 Nokia Group, Finland
virpi.roto@nokia.com
[2] Tampere University of Technology, Human-Centered Technology, Korkeakoulunkatu 6, 33720 Tampere, Finland
kaisa.vaananen-vainio-mattila@tut.fi
[3] University of Leicester, LE1 7RH Leicester, U.K.
elaw@mcs.le.ac.uk
[4] Delft University of Technology, Industrial Design Engineering, Landbergstraat 15, 2628 CE Delft, The Netherlands
a.p.o.s.vermeeren@tudelft.nl

Abstract. High quality user experience (UX) has become a central competitive factor of product development in mature consumer markets [1]. Although the term UX originated from industry and is a widely used term also in academia, the tools for managing UX in product development are still inadequate. A prerequisite for designing delightful UX in an industrial setting is to understand both the requirements tied to the pragmatic level of functionality and interaction and the requirements pertaining to the hedonic level of personal human needs, which motivate product use [2]. Understanding these requirements helps managers set UX targets for product development. The next phase in a good user-centered design process is to iteratively design and evaluate prototypes [3]. Evaluation is critical for systematically improving UX. In many approaches to UX, evaluation basically needs to be postponed until the product is fully or at least almost fully functional. However, in an industrial setting, it is very expensive to find the UX failures only at this phase of product development. Thus, product development managers and developers have a strong need to conduct UX evaluation as early as possible, well before all the parts affecting the holistic experience are available. Different types of products require evaluation on different granularity and maturity levels of a prototype. For example, due to its multi-user characteristic, a community service or an enterprise resource planning system requires a broader scope of UX evaluation than a microwave oven or a word processor that is meant for a single user at a time. Before systematic UX evaluation can be taken into practice, practical, lightweight UX evaluation methods suitable for different types of products and different phases of product readiness are needed. A considerable amount of UX research is still about the conceptual frameworks and models for user experience [4]. Besides, applying existing usability evaluation methods (UEMs) without adaptation to evaluate UX may lead to some scoping issues. Consequently, there is a strong need to put UX evaluation from research into practice.

Keywords: User experience, Evaluation methods, User-centered design, Product development.

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 981–982, 2009.

1 Goals of the Workshop

This workshop will primarily focus on concrete and pragmatic UX evaluation methods, but the UX models developed in academic research will act as ground for the methodological discussion. The aim is to transfer knowledge in both directions: To enable academics to gain knowledge about real problems in evaluating UX, and enable practitioners to get informed and inspired by academics with knowledge about UX factors, models and theories.

The overview of UX evaluation methods and techniques resulting from the previous UXEM'08 workshop [5] and from a SIG session in CHI'09 conference [6] will be used as a starting point. The envisaged main result of the current workshop is a consolidated mapping which shows UX evaluation methods along the product development and product readiness phases. Additionally, challenges and gaps in UX research relevant for UX evaluation will be listed to guide further UX research.

To achieve this, the emphasis in the accepted position papers will be on concrete UX evaluation case studies, but some more theoretical papers will also be accepted. In addition, a list of UX elements and a basic model of product development will be used as a basis for the discussion and group work sessions during the workshop. The questions explored in the workshop cover:

- What experiential evaluation methods are available or under development? What methods from related disciplines, such as arts, industrial design, psychology, sociology, and marketing, could be utilized in UX evaluation?
- What is the applicability of different evaluation methods in different product development phases and product readiness stages?
- How does UX of different application fields – such as mobile consumer services or devices, games, harvester machines, or financial applications – and usage contexts differ? How should the evaluation approaches be adapted for different cases?

Results of this workshop will be applicable as a guideline for UX researchers and practitioners in product development.

References

1. Pine, J., Gilmore, J.H.: The Experience Economy: Working is Theatre & Every Business a Stage. Harvard Business School Press, Boston (1999)
2. Hassenzahl, M.: The thing and I: understanding the relationship between user and product. In: Blythe, M., Overbeeke, C., Monk, A.F., Wright, P.C. (eds.) Funology: From Usability to Enjoyment, pp. 31–42. Kluwer, Dordrecht (2003)
3. ISO 13407:1999, Human-Centred Design Proceses for Interactive Systems. International Standardization Organization (ISO), Switzerland
4. COST294-MAUSE, Workshop on User Experience – Towards a Unified View. In: conjunction with NordiCHI, conference, Oslo, Norway (2006), http://www.cost294.org/
5. Väänänen-Vainio-Mattila, K., Roto, V., Hassenzahl, M.: Now Let's Do It in Practice: User Experience Evaluation Methods in Product Development. In: CHI 2008 Proceedings - Workshops, Florence, Italy, April 6, 2008, pp. 3961–3964 (2008), http://www.cs.tut.fi/ihte/CHI08_workshop/slides/Poster_UXEM_CHI08_V1.1.pdf
6. Obrist, M., Roto, V., Väänänen-Vainio-Mattila, K.: User Experience Evaluation: Do You Know Which Method to Use? A special interest group session. In: Extended abstracts of CHI 2009 conference (2009)

Virtual Teams and Human Work Interaction Design - Learning to Work in and Designing for Virtual Teams

Rikke Orngreen[1], Torkil Clemmensen[2], and Annelise Mark Pejtersen[3]

[1] Assoc. prof., Danish School of Education, Aarhus University
[2] Assoc. prof., Dept. of Informatics, CBS, Denmark
[3] Research Prof., chair of IFIP TC 13

1 Introduction to Theme/Topic

The boundaries and work processes for how virtual teams interact are undergoing changes, from a tool and stand-alone application orientation, to the use of multiple generic platforms chosen and redesigned to the specific context. These are often at the same time designed both by professional software developers and the individual members of the virtual teams, rather than determined on a single organizational level. There may be no impact of the technology per se on individuals, groups or organizations, as the technology for virtual teams rather enhance situation ambiguity and disrupt existing task-artifact cycles. This ambiguous situation calls for new methods for empirical work analysis and interaction design that can help us understand how organizations, teams and individuals learn to organize, design and work in virtual teams in various networked contexts.

2 Workshop Objectives and Expected Outcomes

The objective of this workshop is to raise discussion on both findings from and new methods for studying how people work in and design for virtual teams. The workshop will discuss and present current research on virtual teams as a case of human work interaction design. Topics may include but are not limited to:

- Virtual team cases from public, private and educational institutions
- Formal and informal learning in virtual team design situations
- New methods for empirical analysis and design of collaboration in virtual teams within specific domains and fields of practice
- Studies of virtual team analysis and design activities, e.g. virtual teams performing remote usability testing
- Virtual team design needs with respect to both functionality, usability and pleasurability
- Measures of the time, effort and usefulness spend on ‘normal style’ work processes as organizations localize and reduce the ambiguity of new virtual team processes.

Through the workshop the participants will help increase awareness of the need for deep understandings of the tasks and processes carried out in virtual team working, as well as the cultural, social and organizational context. Both in practice and in research

T. Gross et al. (Eds.): INTERACT 2009, Part II, LNCS 5727, pp. 983–984, 2009.

there is an interest in finding, exploring and using new ways of investigating virtual teams learning processes and to investigate how to analyze and transcend findings to design. .

3 Workshop Organisation and Duration

This full day workshop begins with presentations (by the workshop organizers and workshop participants) creating a common ground for discussions. Prior to the workshop, participants will be asked to come up with interesting dilemmas, paradoxes or possibilities, based on their work in the position papers. These topics and questions will be the foundation of a more detailed round-table discussion facilitated by the organizers. The workshop will end with a thematic view on the day's discussions.

Participants are required to submit position papers (up to 4 pages), based on the theme and topics listed above. Accepted position papers will be made available to the workshop participants prior to the workshop, at the 13.6 HWID website. The participants will be asked to read the position papers in advance and prepare a couple of questions or remarks to one of the papers (pointed out by the workshop organizers), thus ensuring all papers are not only read thoroughly by the workshop organizers and a least one of the participants, but giving a good basis for dialog.

4 Target Audience

The target audience of the workshop is researchers, practitioners and educators, who have experience with virtual teams learning and (re)design processes in human work interaction design. Position papers will be reviewed with respect to their relevance, quality and ability to stimulate discussion. Please send your submission in PDF format, using the standard INTERACT template to: Rikke Orngreen, rior@dpu.dk

5 Key Organisers of the Workshop

The organisers of the workshop are members of the 13.6 HWID working group (see http://ilex.cbs.dk/hwid/ and the event-menu on this site for past initiatives).

Author Index