ACHI 2015

The Eighth International Conference on Advances in Computer-Human Interactions

ISBN: 978-1-61208-382-7

February 22 - 27, 2015

Lisbon, Portugal

ACHI 2015 Editors

Leslie Miller, Iowa State University - Ames, USA
Alma Leora Culén , University of Oslo, Norway

ACHI 2015

Forward

The eighth edition of The International Conference on Advances in Computer-Human Interactions (ACHI 2015) conference was held in Lisbon, Portugal, February 22 - 27, 2015.

The conference on Advances in Computer-Human Interaction, ACHI 2015, was a result of a paradigm shift in the most recent achievements and future trends in human interactions with increasingly complex systems. Adaptive and knowledge-based user interfaces, universal accessibility, human-robot interaction, agent-driven human computer interaction, and sharable mobile devices are a few of these trends. ACHI 2015 brought also a suite of specific domain applications, such as gaming, social, medicine, education and engineering.

The event was very competitive in its selection process and very well perceived by the international scientific and industrial communities. As such, it is attracting excellent contributions and active participation from all over the world. We were very pleased to receive a large amount of top quality contributions.

The accepted papers covered a wide range of human-computer interaction related topics such as graphical user interfaces, input methods, training, recognition, and applications.
We believe that the ACHI 2015 contributions offered a large panel of solutions to key problems in all areas of human-computer interaction.

We take here the opportunity to warmly thank all the members of the ACHI 2015 technical program committee as well as the numerous reviewers. The creation of such a broad and high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and efforts to contribute to the ACHI 2015. We truly believe that thanks to all these efforts, the final conference program consists of top quality contributions.

This event could also not have been a reality without the support of many individuals, organizations and sponsors. In addition, we also gratefully thank the members of the ACHI 2015 organizing committee for their help in handling the logistics and for their work that is making this professional meeting a success.

We hope the ACHI 2015 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the human-computer interaction field.

We also hope that Lisbon provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

ACHI 2015 Chairs

Leslie Miller, Iowa State University - Ames, USA
Katia Sykara, Carnegie Mellon University, USA
Alma Leora Culén , University of Oslo, Norway
Hermann Kaindl, Vienna University of Technology, Austria

ACHI 2015

COMMITTEE

ACHI Advisory Committee

Leslie Miller, Iowa State University - Ames, USA
Katia Sykara, Carnegie Mellon University, USA
Alma Leora Culén , University of Oslo, Norway
Hermann Kaindl, Vienna University of Technology, Austria

ACHI 2015 Technical Program Committee

Andrea Abate, University of Salerno, Italy
Víctor Acinas Garzón, Instalaciones Inabensa S.A. - Madrid, Spain
Eugenio Aguirre Molina, University of Granada, Spain
Andrea Alessandro Gasparini, University of Oslo, Norway
Majed Alshamari, King Faisal University, Saudi Arabia
Florian Alt, University of Munich, Germany
Margherita Antona, Foundation for Research and Technology - Hellas (FORTH) - Heraklion, Greece
Leena Arhippainen, University of Oulu, Finland
Mukesh Barange, École Nationale d'Ingénieurs de Brest, France
Michael Bauer , Heilbronn University, Germany
Marios Belk, University of Cyprus, Cyprus
Peter Berkelman, University of Hawaii-Manoa, USA
Jane Billestrup, Institute of Computer Science - Aalborg University, Denmark
Mark Billinghurst, University of Canterbury, New Zealand
Karlheinz Blankenbach, Pforzheim University, Germany
Clodis Boscarioli, Western Paraná State University (Unioeste), Brazil
Mohamed Bouri, EPFL, Switzerland
Christian Bourret, Université Paris-Est Marne-la-Vallée (UPEM), France
Marius Brade, Technische Universität Dresden, Germany
Lourdes Brasil, University of Brasília at Gama, Brazil
Gerd Bruder, University of Hamburg, Germany
Luis Carriço, University of Lisbon, Portugal
Silvio Cesar Cazella, Federal University of Health Sciences of Porto Alegre (UFCSPA), Brazil
Jessie Y.C. Chen, U.S. Army Research Laboratory - Human Research & Engineering Directorate, USA
Jacek Chmielewski, Poznań University of Economics, Poland
Ryszard S. Choras, University of Technology & Life Sciences - Bydgoszcz, Poland
Thomas Christy, School of Computer Science at Bangor University, UK
Shaowei Chu, University of Tsukuba, Japan
Sharon Lynn Chu, Virginia Tech, USA
José Coelho, University of Lisbon, Portugal
Cesar Alberto Collazos O., University of Cauca - Popayán, Colombia
André Constantino da Silva, Instituto Federal de São Paulo, Brazil
Sandra Costa, University of Minho, Portugal
Lynne Coventry, Northumbria University, UK

Manuel Cruz, Immersion Corp., Canada
Leonardo Cunha de Miranda, Federal University of Rio Grande do Norte (UFRN), Brazil
Urska Cvek, Louisiana State University Shreveport, USA
Dragos Datcu, Netherlands Defence Academy, The Netherlands
Kerstin Dautenhahn, The University of Hertfordshire, United Kingdom
Lucio Tommaso De Paolis, Salento University - Lecce, Italy
Nancy Diniz, Xi'an Jiaotong-Liverpool University, China
John M. Dolan, Carnegie Mellon University, USA
Rolf Drechsler, DFKI - Bremen, Germany
Carlos Duarte, University of Lisbon, Portugal
Ismael Duque, University of Hertfordshire, UK
Martin Ebner, Graz University of Technology, Austria
Anastasios A. Economides University of Macedonia - Thessaloniki, Hellas
Larbi Esmahi, Athabasca University, Canada
Sascha Fagel, Zoobe GmbH, Germany
Michael Farrugia, University College Dublin, Ireland
Karim Fraoua, Université Paris-Est- Conservatoire des arts et métiers, France
Somchart Fugkeaw, Mahidol University - Nakhonpathom, Thailand
Mathias Funk, Eindhoven University of Technology, The Netherland
Miguel Garcia, University of Valencia, Spain
Matjaz Gams, Josef Stefan Institute- Ljubljana, Slovenia
Marta Gatius, Universitat Politècnica de Catalunya, Spain
Cristina Gena, Università di Torino, Italy
Panagiotis Germanakos, University of Cyprus, Cyprus
Irini Giannopulu, Pierre & Marie Curie University - Paris VI, France
Denis Gillet, Swiss Federal Institute of Technology (EPFL), Switzerland
Mehmet Göktürk, Gebze Institute of Technology, Turkey
Luis Gomes, Universidade Nova de Lisboa / UNINOVA-CTS - Monte de Caparica, Portugal
Pascual Gonzalez, University of Castilla-La Mancha, Spain
Bernard Grabot, LGP-ENIT, France
Andrina Granic, University of Split, Croatia
Toni Granollers, University of Lleida, Spain
Vic Grout, Glyndwr University - Wrexham,UK
Jaume Gual Orti, Universitat Jaume I – Castellón, Spain
Celmar Guimarães da Silva, University of Campinas, Brazil
Ido Guy, IBM Research-Haifa, Israel
Maki K. Habib, The American University in Cairo, Egypt
Till Halbach, Norwegian Computing Center / Norsk Regnesentral (NR), Norway
Lynne Hall, University of Sunderland, UK
Kristian J. Hammond, Northwestern University - Evanston, USA
Riad Hammoud, BAE Systems, USA
Lane Harrison, Tufts University, USA
Christoph Hintermüller, Guger Technologies OG – Schiedlberg, Austria
Richard Hoadley, Anglia Ruskin University, UK
Thomas Holz, University College Dublin, Ireland
Ilias Iliadis, IBM Research - Zurich, Switzerland
Jamshed Iqbal, COMSATS Institute of Information Technology (CIIT), Pakistan
Ryo Ishii, NTT Communication Science Laboratories - NTT Corporation, Japan

Dragan Ivetic, University of Novi Sad, Serbia
Jari Due Jessen, Technical University of Denmark, Denmark
Ran Jiao, Beihang University - Beijing, China
Gareth Jones, Dublin City University, Ireland
Hanmin Jung, Korea Institute of Science & Technology Information, Korea
Janusz Kacprzyk, Polish Academy of Sciences - Warsaw / IEEE Fellow, Poland
Hermann Kaindl, Vienna University of Technology, Austria
Ahmed Kamel, Concordia College - Moorhead, USA
Athanasios Kapoutsis, Centre for Research and Technology Hellas | Information Technologies Institute, Greece
Koji Kashihara, Tokushima University, Japan
Hamed Ketabdar, Deutsche Telekom Laboratories - TU Berlin, Germany
Jonghwa Kim, University of Augsburg, Germany
Youngmin Kim, Korea Electronics Technology Institute (KETI), South Korea
Michael Kipp, University of Applied Sciences Augsburg, Germany
Frank Klawonn, Ostfalia University of Applied Sciences - Wolfenbuettel, Germany
Jesuk Ko, Gwangju University, South Korea
Anita Komlodi, University of Maryland - Baltimore County, USA
Heidi Krömker, Technische Universität Ilmenau, Germany
Satoshi Kurihara, Osaka University, Japan
Andrew Kusiak, The University of Iowa, USA
Jannie Lai, CITRIX, USA
Vidas Lauruska, Siauliai University, Lithuania
Soo-Young Lee, KAIST and RIKEN BSI, Korea
Alma Leora Culén , University of Oslo, Norway
Cynthia Y. Lester, Tuskegee University, USA
Tsai-Yen Li, National Chengchi University - Taipei, Taiwan
Yangmin Li, University of Macau - Taipa, Macao
Chin-Teng (CT) Lin, National Chiao-Tung University, Taiwan
Wei-Kai Liou, National Taiwan Normal University, Taiwan
Jean-Luc Lugrin, Universität Würzburg, Germany
Ameersing Luximon, Hong Kong Polytechnic University, Hong Kong
Francisco Madera, Universidad Autonoma de Yucatan - Merida, Mexico
Jacek Mandziuk, Warsaw University of Technology, Poland
Antonio Maria Rinaldi, Università di Napoli Federico II, Italy
Célia Martinie, IRIT | University Toulouse 3, France
Andrea Mason, University of Wisconsin-Madison, USA
Christoph Mayer, Technische Universität München, Germany
Troy McDaniel, Arizona State University, USA
Gerrit Meixner, Heilbronn University, Germany
Daniel R. Mestre, Aix-Marseille University, France
Harald Milchrahm, Technical University Graz, Austria
Leslie Miller, Iowa State University - Ames, USA
Peyman Moghadam, CSIRO, Australia
Yasser F. O. Mohammad, Assiut University, Egypt
Jaime Muñoz-Arteaga, Universidad Autónoma de Aguascalientes, México
Fazel Naghdy, University of Wollongong, Australia
Toru Nakata, National Institute of Advanced Science and Technology (AIST), Japan

Deok Hee Nam, Wilberforce University, USA
Claudia Nass, Fraunhofer Institute for Experimental Software Engineering, Germany
Andreas C. Nearchou, University of Patras, Greece
Roger Ng, Hong Kong Polytechnic University, Hong Kong
Grace Ngai, Hong Kong Polytechnic University, Hong Kong
Radoslaw Niewiadomski, University of Genoa, Italy
Gregory O'Hare, University College Dublin (UCD), Ireland
Zahen Malla Osman, Centre d'Etude et de Recherche en Informatique et Communications (CEDRIC), France
Fakri Othman, Cardiff Metropolitan University (UWIC), UK
Evangelos Papadopoulos, National Technical University of Athens, Greece
Igor Paromtchik, INRIA - Saint Ismier, France
Alexandru Patriciu. McMaster University - Hamilton, Canada
Estelle Perrin, CReSTIC | Reims University – Charleville-Mezieres, France
Jose Pinto, Universidade do Porto, Portugal
Marina Puyuelo Cazorla, Universitat Politècnica de València (UPV), Spain
Didier Puzenat, Université des Antilles et de la Guyane - Pointe-à-Pitre, Guadeloupe
Rafael Radkowski, Iowa State University, USA
Nitendra Rajput, IBM Research - New Delhi, India
Marco Antonio Ramos Corchado, Universidad Autónoma del Estado de México, Mexico
Xiangshi Ren, Kochi University of Technology, Japan
Arcadio Reyes Lecuona, University of Málaga, Spain
Andreas Riener, University Linz - Linz/Donau, Austria
Antonio Maria Rinaldi, Università di Napoli Federico II, Italy
Joel Rodrigues, Intituto de Telecomunicações / University of Beira Interior, Portugal
Agostinho Rosa, Lasseb-ISR-IST / Technical University of Lisbon, Portugal
José Rouillard, Université des Sciences et Technologies de Lille, France
Rebekah Rousi, University of Jyväskylä, Finland
Juha Röning, University of Oulu, Finland
Antonio José Sánchez Salmerón, Universidad Politécnica de Valencia, Spain
Corina Sas, Lancaster University, UK
Munehiko Sasajima, Osaka University, Japan
Dominique Scapin, INRIA - Le Chesnay, France
Shunji Shimizu, Tokyo University of Science, Suwa, Japan
Maria Augusta Silveira Netto Nunes, Universidade Federal de Sergipe (UFS), Brazil
Johanna Silvennoinen, University of Jyväskylä, Finland
Pavel Slavik, Czech Technical University in Prague, Czech Republic
Ross Smith, University of South Australia, Australia
Jaka Sodnik, University of Ljubljana, Slovenia
Lübomira Spassova, Luxembourg's Public Research Centre Henri Tudor - SANTEC, Luxembourg
Dimitris Spiliotopoulos, University of Peloponnese, Greece
Weilian Su, Naval Postgraduate School - Monterey, USA
Tsutomu Terada, Kobe University, Japan
Daniel Thalmann, VRlab/EPFL - Lausanne, Switzerland
Robert J. Teather, McMaster University, Canada
Edgar Thomas, Federal University of Pernambuco | Brazil's Government, Brazil
Hiroaki Tobita, Sony CSL Paris, France
Godfried T. Toussaint, New York University Abu Dhabi, United Arab Emirates

Copyright Information

Table of Contents

UI Delegation: The 3rd Dimension for Cross-Platform User Interfaces

Dagmawi Lemma Gobena
Addis Ababa University
IT Doctoral Program
Addis Ababa, Ethiopia
dagmawi.Lemma@gmail.com

Abel Gomes
University of Beira Interior
Covilhã, Portugal
agomes@di.ubi.pt

Dejene Ejigu
Addis Ababa University
IT Doctoral Program
Addis Ababa, Ethiopia
ejigud@yahoo.com

***Abstract*—Two of the prominent dimensions behind the development of cross-platform UIs are the UI distribution and UI migration. In UI distribution, since UI elements of a given application has to be distributed across more than one device, some UI elements can be even duplicated. In UI migration, the description and construction of UI elements are centralized using a client-server model of computing over a computer network. Thus, we end up having limitations with respect to scalability and maintainability of the computing environment. Also, UI distribution and migration mostly support explicit HCI for interactive systems. However, in ubiquitous computing, implicit HCI is the most desired interaction approach. In this paper, we present the theoretical concept of *UI delegation* as the third dimension that ideally supports implicit HCI and trans-modality by assuring autonomy of the platforms using a peer-to-peer model**.

Keywords – cross-platform UI; multi-platform UI; user interface design; ubiquitous computing.

I. INTRODUCTION

The amalgamation of various technologies to support the needs of new computing models has become prevalent in computing environments like ubiquitous computing. For example, in the ranking system shown in Figure 1, which we developed to rank and produce outcomes for athletes, we have the following: (1) results are redisplayed using a web application; (2) setting to radio frequency identification (RFID) system is made using a GUI application; (3) every athlete wears RFID tag that is uniquely encoded, hence it is possible to create a sort of implicit interaction between the athlete and the system; (4) the RFID reader box is configured using a monochrome display; (5) athletes can receive their results on their mobile phones, or on any other personal device they may use. Such amalgamation of various technologies results in heterogeneous environment.

Nowadays, heterogeneity of personal devices is inescapable. Heterogeneity is one of the characteristics of ubiquitous computing [1], and it is caused by the coexistence of various devices in the same computing environment. Furthermore, the heterogeneity is a result from the diversity of software, users, interaction modalities, and environments.

Nevertheless, Weiser's vision of ubiquitous computing, which demands that computer is an invisible servant [2] [3], has not been achieved yet [4]. With regards to the human-computer interaction (HCI), invisibility of computers can be achieved, partly through implicit HCI (*i*-HCI) [5] and context aware systems. On the other hand, the explicit HCI (*e*-HCI) development for interactive systems requires consideration of capabilities and constraints of diverse platforms and users, in addition to provide interaction modalities in a human fashion (e.g., speech, gesture, etc.)

The platform heterogeneity, together with additional needs of interaction modalities, as well as the proliferation of new technologies, poses unique challenges to designers and developers of user interfaces (UIs). Analyzing user profiles and platform capabilities and constraints in the usability engineering lifecycle [6] is certainly challenging due to the heterogeneity of platforms (devices and software) and users. Therefore, UIs are expected to be cross-platform. That is, a UI that runs on a certain platform (e.g., desktop screen) shall be able to appear on another platform (e.g., small handheld device) without losing its usability.

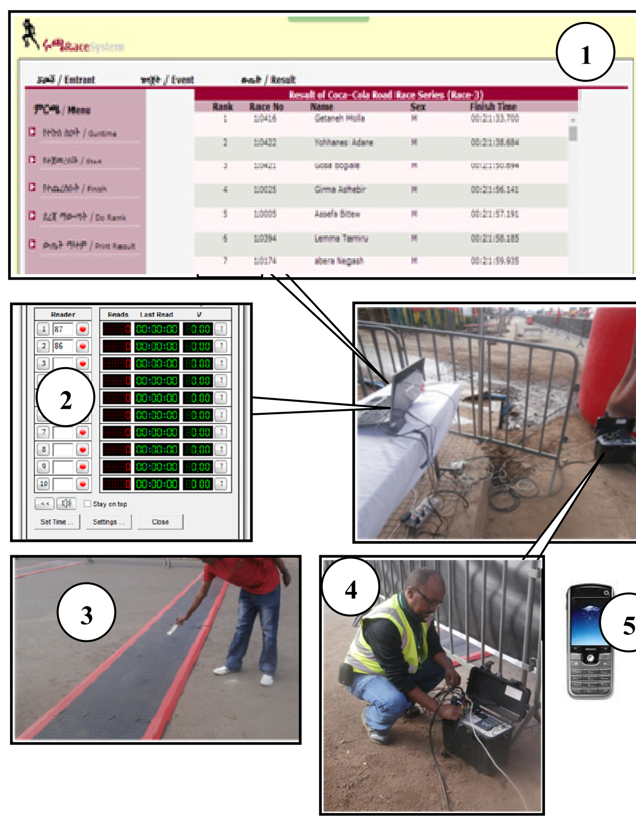


Figure 1. Heterogeneous race system for athletics.

 ISBN: 978-1-61208-382-7

To overcome the challenges of heterogeneity, there are models and theoretical frameworks suggested and developed in the HCI community, in order to sustain the notion of cross-platform UIs. In this paper, by cross-platform UI (respectively, cross-platform interaction), we mean UIs (respectively, interaction modality), no matter whether it is implicit or explicit, through which computers and users interact across various platforms in conformity with capabilities and constraints of users (user profiles) and platforms, without compromising the usability. Thus, cross-platform UIs should consider the context of the systems (e.g., applications, tools, interaction modalities, etc.), devices and users.

Two of the prominent dimensions behind the development of a cross-platform UI are those concerning UI distribution and migration. In the literature, we found that both approaches focus on a particular aspect of the heterogeneity – mostly the device [7]. However, generating UIs in a heterogeneous environment based on a specific context (e.g., device, user, task, interaction modalities, etc.) most likely reduces the usability of the system, which entails several usability issues [8]. Thus, we propose the concept of *UI delegation* as the third dimension to help in the development of cross-platform UIs.

Our motivation is based on three main points. Firstly, in order to automatically (or at design time) generate a cross-platform UI, we have to consider the merger of diversified contexts from the system (i.e., including interaction modalities, web services, etc.), device and user aspects, so as to meet usability requirements [6] [9]. Therefore, we found it important to introduce a different approach that compels the consideration of ternary views (the *system*, *device,* and *user*) in cross-platform UIs, but not in partiality of any of the views. In fact, though it is common practice to consider these three views in UI development, most works and techniques related to cross-platform UIs only focus on one of the views (user, device or system) at the time of automatically generating a specific UI. For example, the pattern based-approach proposed by Lei et al. [10] as well as responsive web development (RWD) focus on screen size adaptation, while Nichols et al. [11] focus on the functionality of the appliances, and Sauter et al. [12] only consider the device type.

Secondly, both distributed and migratory UI concepts are often implemented using the client-server model. While the final UI runs on the client side, the appropriate UI for a specific platform is generated at the server side [13]. Thus, the server is responsible to maintain the UI description [11] [14], to update and preserve the UI state [7], to store a duplicate version of the UI [12], and so forth. But, centralizing the description of capabilities of each platform often imposes limitations to the scalability and maintainability of the environment. Furthermore, in a ubiquitous environment, the peers are desirably autonomous, so that each peer shall be able to generate a UI as per its capabilities and its own autonomy.

Finally, heterogeneity may also be a result of the presence of various sorts of interaction modalities (including *i*-HCI). Nevertheless, most works we found in the literature are about *e*-HCI. Otherwise, despite the fact that the ubiquitous computing aims at invisible UIs [15], interaction modalities in the arena of *i*-HCI are not well covered. It has to be noted that *i*-HCI can be also achieved by using various technologies (i.e., sensors, motion capturing tools, etc.); and this in turn leads to another aspect of heterogeneity. Therefore, we need a new approach to the development of cross-platform UIs.

The purpose of this paper is to provide a new theoretical concept that complements the efforts made so far to support UI development for heterogeneous environments. We consider the problem of heterogeneity as a result from the need of collaboration between platforms (i.e., *device*-and-*system* units) that are owned or controlled by a *human user*. Hence, we focus on the concept of delegation as it is applied in [16] for supporting collaboration between agents in an agent-based environment. Accordingly, we propose the concept of *UI delegation* with autonomy.

Autonomy of nodes that collaborates in environments like ubiquitous computing can be more effective if peer-to-peer approach is followed instead of client-server. Thus, considering the benefits of peer-to-peer model, we present the concept of *UI delegation*. Furthermore, the UI-related data, which are exchanged between the peers, shall be based on a protocol elaborated on a common interface language (CIL).

In the context of the present work, the *UI delegation*:

- insures autonomy of peers to render UI according to its own capabilities, having also into consideration the user capabilities listed in his/her profile;
- takes into account the heterogeneity of interaction modalities, including *i*-HCI;
- includes a protocol that facilitates collaboration as in the peer-to-peer model;
- contributes to the usability of the system in the sense that it provides a comprehensive understanding of the usability concerns related to human, system, and platform views;
- advocates decentralization to attain autonomy, and intends to resolve scalability and maintainability issues that may prevail as a result of centralization.

Our approach is different from other works in three ways. Firstly, it attempts to simultaneously take onboard the *system*, *device* and *user* aspects in the process of generating UI at run time, instead of relying only on one of those aspects. Secondly, it is proposed in the context of a peer-to-peer model, where multicasting is used and collaboration is maintained using a CIL protocol between peers. Thus, it is possible to achieve autonomy of peers and resolve scalability and maintainability issues. Finally, it is different since our concept uses the *system* view to include various interaction modalities, as well as *i*-HCI, instead of limiting the system view to describe software capabilities and constraints.

The rest of this paper is structured as follows. In Section 2, we discuss works related with cross-platform UI development, as well as UI distribution and UI migration. In Section 3, we discuss the concept of *UI delegation*, including the requirements deemed to satisfy this concept. Finally, in

 ISBN: 978-1-61208-382-7

Section 4, we draw relevant conclusions about the concept of UI delegation, discussing what more should be done to materialize it across ubiquitous environments.

II. Literature Review

Model-based UI development (MBUID) is one of the principal approaches that strive in developing UIs that can run on multiple or across heterogeneous platforms. In MBUID, users, data, tasks and functions can be modeled in order to turn them into interaction concepts [11] [17] [18]. The models are then used to guide the UI development, as well as to automatically generate the end UI [10] [11].The Cameleon Reference Framework abstracts models to describe the UI at different levels of abstraction, namely: abstract UI, concrete UI, and final UI [13]. The abstraction in a model signals the list of candidate widgets for the interaction. For example the "choice" concept can be an abstraction of combo box, list box, check box, and radio group [16].

Vanderdonckt [19] classifies the UI design for heterogeneous platforms as per the situation that causes the diversity. Therefore, the UI design may focus on the presence of multiple users or, alternatively, on the usage of multiple monitors, devices, platforms, and displays [20]. In this regard, the UI distribution and migration are followed as general UI development approaches [4] [7] [21] [22] [8]. UI distribution is the concept of spreading UI components "across one or more of the dimensions of input, output, platform, space, and time" [21].

In [8], distributed and migratory UIs are discussed as two independent concepts. Migratory UIs can be in the form of distributed UIs, but they shall enable the user to continue the interaction without losing the state (content) of the UI [7] [8]. In UI distribution, UI elements are distributed across platforms, and, in some cases, this may create duplication of UI elements [8]. For example, in [12], a multi-client (multi-platform) UI is presented using the model-view-controller (MVC) architecture that stores different versions of a webpage (UI) on the server for each predefined platform, and where the controller selects one of the UI versions that most fits a particular platform. But, this approach is prone to maintainability and scalability issues. For example, if a UI element has to be modified or added, such an operation has to be done for each version of the respective UI.

Elmqvist [21] pointed out in that distributed UIs can have multi-device environment and/or interaction modalities aspects, including application and content redirection in addition to UI migration. In due case, usability is a concern when adapting the application interface to another device with different capabilities and constraints. The notion of plasticity of UI is thus presented as another concept to refer the ability of UI to withstand variations across platforms, while preserving its usability [23].

If usability is a concern, then both the platform and user capabilities have to be addressed while generating an UI [6] [9] to be distributed or migrated. However, most works put a focus on the capabilities of one sort of participating entities, mostly the user or platform. For example, Nichols et al. described the interface and function of appliances using a UI description language, which is applied to create "specifications for 33 appliances, including several with more than 100 functional elements" [11]. Thus, only the platform (appliance and application) capabilities are the main consideration for generating the UI. Also, Lei et al. considered that the device context is to adapt UIs across devices with various screen sizes [10]. MARIA [14] was also proposed as a description language to support migratory UIs and to design and develop multi-device UIs by using Web services following the form of *e*-HCI. But, it has to be noted that UIs are regarded as a means of communication between the user and the computing environment, and this should include invisible UIs (or *i*-HCI), in which interaction ideally takes place with no perceived mediation, and in a more real-world interaction style [15]; and such modality is the one that most fits the notion of ubiquitous computing. Some works [24] [25]have attempted to address the personalization of users by automatically generating interfaces that are customized to an individual user profile.

Paternò et al. [8] pointed out that in UI distribution, at least two devices are involved when rendering UI. Despite the main focus could be elsewhere (i.e., user, task, environment, modality, etc.), it is vital to consider the device context in any case. Therefore, we formally consider both the platform (device and system) and user capabilities in our conceptual approach.

Four concerns are discussed in [11][21] focusing on the multiplicity of displays, platforms, operating systems, and users, before proposing their toolkit developed in the peer-to-peer model. The user aspect in [20] considers distributing UI for multiple users, but not on the heterogeneity of users. This differs from our concept of user view, since we consider the user capabilities and constraints as part of the user view, in addition to the number of users to which the UI is migrated. Similarly, Elmqvist [26] introduced a peer-to-peer middleware, which is "tailored for high-performance visualization" [26] within an environment with diversified display sizes, such as tabletop display, wall-mounted display, and mobile devices, but without the *user* view.

In [6], during the usability engineering lifecycle of UI development for interactive system, the *user profile* is mainly about characterizing the user, not only naturally (physically), but also with respect to the cognition model and the psychological makeup of the user. It is important to note that supported human capabilities (what we are proposing to be included in generic terms) is different from the user profile studies that focus on classifying the behavior of specific user groups by culture, experience, knowledge, etc. Such constraints are contextual, but our concept lays on the physical and technical capabilities and constraints of users coexisting in the same environment, or those participating in the UI distribution or migration. For example, if an UI is shared between users who have visual impairment and those who have not, then the process of UI delegation shall consider this situation by creating the UI on the *delegatee* side as per the user profile, which can be different from the user profile on the *delegator* side.

After considering the various solutions, concepts, and theories in the literature, we noted the following gaps:

- The works we found in the literature focus on *e*-HCI and solutions that are most related with interactive systems; however, ubiquitous computing can be smart as well. Thus, it requires *i*-HCI [5].
- Distributed and migratory UIs are generated by considering a particular view, mostly the device capabilities (e.g., screen size), but rarely the user and interaction modalities. Furthermore, the migration or distribution is often between similar modalities [7]. Thus, a generic approach that focuses on the merging of *user*, *device* and *system* is important, so as to support heterogeneity of users and the interaction modalities.
- In the general setting of ubiquitous environments, with which computing and interaction with heterogeneous platforms is carried out on the fly, scrutinizing the platform capabilities and constraints would be endless and impractical due to the numerous options of interaction modalities and technologies, not to mention those that are expected to emerge in the future. Hence, scalability should be one of the prominent considerations to be taken in the new concept of UI delegation.

III. USER INTERFACE DELEGATION

The rationale for the emergence of the UI distribution and migration concepts result from the need for enabling users to continue performing their tasks on the go and pervasively. We consider the concept of *UI delegation* as the third dimension (in addition to UI distribution and UI migration) that sustains the development of cross-platform interfaces, so that interaction can be extended and usability can be improved by sharing capabilities of *delegatee* platform. For example, a **list box** widget can be used "*on behalf of*" **radio button** for implementing "*choice*" concept in the interaction. Similarly, instead of visually reading a text from the screen, it can be converted to audio and played if the capability exists. Thus, **audio listening** can be used "*on behalf of*" of **visual reading** across platforms, so that trans-modality can be achieved after all.

The notion of "*on behalf of*" is driven by the cooperation and collaboration between the *delegator* (i.e., the one that requires the UI to be rendered on a remote platform) and *delegatee* (i.e., the one that renders a UI on behalf of other peer), and this should happen when the *delegator* desires to perform the task but knows there is a better capability on the *delegatee* side. For example, while composing a message the user can type using a keyboard on desktop/laptop more easily and efficiently than using keypad of a smartphone. On the other hand, smartphone may possess the connection and SMS service. Therefore, the notion of "*on behalf of*" exists if the desktop/laptop is delegated only for the purpose of delivering the input modality as per its capability. Considering this example, the idea could be similar to the concept of UI granularity (or, in our case, granularity of the interaction modality) that is manipulated during distribution or migration of the UI (or part of it) as discussed in [27]. However, in *UI delegation,* the UI element is not distributed but created at runtime as per the capability of the platform that renders the UI element – the *delegatee*.

UI delegation, in addition to supporting cross-platform UIs as distribution and migration approaches do, it is also useful to create a merger of capabilities in a certain computing environment. As discussed above, the heterogeneity may occur as the result of the diversity of capabilities owned by *systems* (application and interaction modalities), *devices*, and *users*. The merger of the capabilities can be thus used to extend the *capability domain.*

$$C_1 = \{a, b, c, d\}$$

$$C_2 = \{a, b, x, y\}$$

$$\textit{Therefore}$$

$$C_d = C_1 \cup C_2$$

Figure 2. Representation of capabilities domain

For instance, as shown in Figure 2, let C_1 and C_2 be sets of capabilities (interaction modalities) of platforms P_1 and P_2, respectively, (i.e., those coexisting in the same computing environment). Then, C_d is the *capability domain* we can benefit from using such computing environment. It is apparent that both platforms have shared common capabilities $\{a, b\}$, but each of them also has exclusive capabilities, $\{c, d\}$ for P_1 and $\{x, y\}$ for P_2. Therefore, UI delegation can be applied to enable one platform to use one or more capabilities of another platform. In due process, if the modalities between the two platforms are the same, then it is said that we have mono-modality; otherwise it is trans-modality. Thus, if P_1delegates P_2, then P_2 is running the desired interaction modality *"on behalf of"* the *delegator*. We call P_1 the *delegator* and P_2 the *delegatee*.

A. Theory of Delegation

Castelfranchi et al. relate delegation to agents since it is related with the notion of "*task*" of "*on behalf*" in addition to the need of autonomy and collaboration [15]. Thus, "*task*" of "*on behalf*", autonomy, and collaboration are the three prominent reasons leading to the theory of delegation [15]. Similarly, we use the theory of delegation, but applied to the cross-platform context of HCI.

The notion of "*task*" of "*on behalf*" is discussed above. In HCI, the autonomy is satisfied by letting platforms (peers) to run delegated UI (i.e., the complete or partial version of the UI) in their own capabilities, instead of generating the UI from server side. Finally, the collaboration is met using the communication protocol between peers.

In the literature, we found that the theory of delegation is well presented in works related to agent-based systems. Haddadi develops the theory by taking "*an internal perspective to model how individual agents may reason about their actions*" [26]. This is further developed in [15], where it is stated that "*in delegation an agent A needs or likes an action of another agent B and includes it in its own plan, thus, A is trying to achieve some of its goals through*

B's action". According to Castelfranchi et al., *A* is said to be the "client", while *B* is the contractor [15]. In our context, *A* decides to whom to delegate as an autonomous peer, although *B* may reasonably agree or not to be delegated. We call to *A* and *B* the "*delegator*" and "*delegatee*", respectively.

In spite of not being defined as a theory in [27], the concept of delegation is used with the intent of compensating the low computational performance of small handheld devices by delegating in high performing computers the execution of tasks requiring higher performance computation.

We draw the theory to support cross-platform UIs within peer-to-peer model, such that interaction modalities include the practice of *i*-HCI, as well as the notion of invisible UI. Furthermore, the capabilities used to generate a UI component shall be defined from the *human, device* and *system* views. Thus, in UI delegation, we have the following:

- a peer (*delegator*) shall demand a capability of another peer in the same computing environment;
- all peers are responsible to register and maintain their own capabilities locally, and advertise them when required;
- a peer looking for a capability shall advertise it, and only peers that own such a capability shall respond;
- a *delegator* is in control only before transferring the UI-related information to the *delegatee*; and
- a *delegatee* is in control only while delivering the UI, loosing such control when the UI state is changed as a result of interaction.

In order to maintain the collaboration between peers, and to standardize how capabilities are represented, the peers shall use the CIL that serves as a protocol between peers.

B. The Protocol (CIL)

The UI delegation concept we propose in this paper is meant to fit in a peer-to-peer model that requires decentralizing UI-related information, as well an enabling peer that is intended to serve as *delegatee*. Hence, as in [15], where the agent has to select the task to be run for another agent, the *delegatee* has to invoke some of its own capabilities (i.e., locally stored) that are adequate to deliver the required UI (or part of it) on behalf of the *delegator*. To achieve this requirement, we propose a set of rules governing the UI-related information exchange between peers. In addition, how each peer registers its capabilities locally has to be standardized. Therefore, CIL is conceptualized as the protocol that serves these needs.

In order to apply CIL as a protocol between the peers, it shall play three basic functions as: *syntax and semantics*, *description language*, and *communication rules*.

1) Syntax and semantics

The peers taking advantage of the concept of *UI delegation* shall use a standardized and common way of describing the UI-related information. This includes standardizing the syntax and semantics of the language to be used between peers. Also, it requires a decision about which aspect of UI-related information to be represented using the protocol. As discussed above, one of our main goals is to consider and use the merger of the *system, device*, and *user* contexts during the cross-platform UI development, provided that are deemed important for the UI generation (i.e., at runtime or design time). Therefore, at this stage of our work, we consider the *human*, *system*, and *device* as views to be integrated in the *CIL-definition*. The human and device views can be taken into account for identifying physical and technical capabilities supported and available on each peer. Hence, a *delegator* can use the information to select the *delegatee* that optimally meets the desired capabilities.

On the other hand, the system view is required to define available capabilities related to interaction modalities, available support for *i*-HCI, tools useful to support conversion between modalities (e.g., text-to-speech), and so forth. In particular, the system view covers three broad aspects of the interaction:

- *How interaction is presented:* the presentation of UI can be in the form that the user shall react to (e.g., web form), or implicitly (e.g., ambient display) in which users can be passive in respect to the presentation.
- *How interaction is triggered:* interaction can be triggered as a result of the occurrence of a specific event, command, periodic instance, etc.
- *The type of modality-state:* peer might have the capability to use mono-modality or trans-modality. In trans-modality, peers are capable of converting one modality into a different type of modality (e.g., text-to-speech)

As shown in Figure 3, the *CIL-definition* is the foundation on which UI related-information and messaging are described during the process of *UI delegation*. The *CIL-definition* constructs the syntax and semantics of the CIL in general, which has to be followed by each peer. Also, since the definition can be improved from time to time, it has to be associated with version identifier.

2) Description protocol

Once each peer knows how and what to specify, the protocol can be used to describe the capabilities of each peer, as well as the presentational information of the current UI desired to run on the *delegatee* side.

Each peer shall describe locally the capabilities it supports in accordance to the *CIL-definition* with the *human, device*, and *system* views. Thus, when the *delegator* decides to delegate a peer, the selected peer (*delegatee*) shall present the UI in its own capability as described locally.

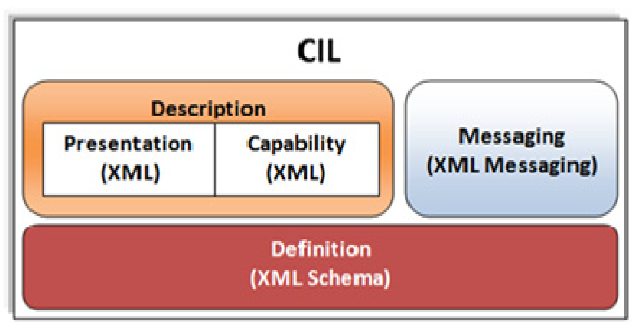


Figure 3. Structure of CIL

Description is also required to communicate and use presentational UI information (i.e., the structure of UI element, the entire UI or interaction modality) useful to create the UI presentation on the *delegatee* side.

The presentation is created using the *CIL-definition* in two stages. In the first stage, the *delegator* has to create the CIL version of the UI intended to run on the *delegatee* side. Then in the second stage, the *delegatee* shall map the presentation (*CIL-description*) in accordance to its capabilities, and generate the new presentation of the delegated UI. More discussion about mapping capabilities is given further ahead.

3) Communication rule

The exchange of UI-related information shall follow a standard that can be understood and interpreted by each peer. Therefore, in addition to creating the *CIL-definition* as rule upon which the *CIL-description* of capabilities and presentation is made, we found it valid to consider a third role within CIL through which peers collaborate: *CIL-Messaging*. *CIL-messaging* is the third role that must be played in three situations:

- when the *delegator* sends a *delegation request* to peers;
- when peers respond to a *delegation request;*
- after the *delegator* selects one of the peers as *delegatee* and, when the described UI or interaction modality is transmitted; and
- when the *delegatee* decides to transmit the UI with its new state back to the *delegator*.

C. UI Delegation Process

The delegation process can be started on-demand or automatically, and there are five important requirements to be fulfilled:

- Describing capabilities
- Creating *delegation request*
- Responding to a *delegation request*
- Selecting and appointing *delegatee* and
- Mapping UI/interaction modality

The *CIL-description* and/or *CIL-messaging* are used to fulfill each of these requirements, while the *CIL-definition* serves as a standard for maintaining consistency and interoperability across peers in the process of messaging, as well as to describe the capabilities and presentation.

1) Describing Capabilities

Local capabilities of each peer can be described by the UI designer, and verified as per the *CIL-definition* (the XML schema).

2) Creating Delegation Request

Delegation request has to be created first by translating the current UI deemed to be delegated into *CIL-description*. During *delegation request*, the *CIL-description* is in a more abstract form only to depict the desired capabilities from user, device, and system points of view.

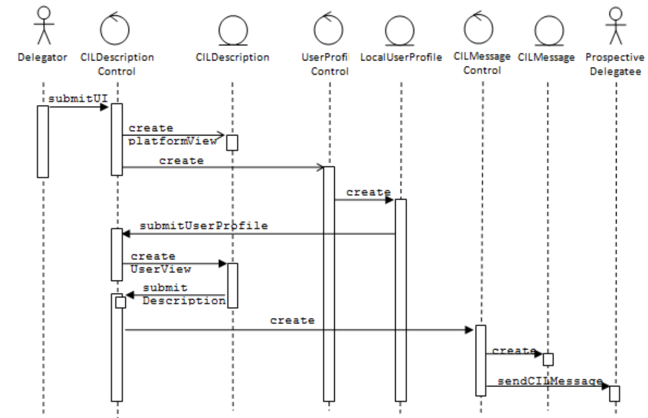


Figure 4. *Delegation request*

In addition, the required user profile shall be used to create the human-being view description. Once the description is done, it is used to define the *delegation request* message using the CIL-messaging format, being then the message sent to the prospective CIL-enabled peer, as shown in Figure 4.

3) Responding to delegation request

Each peer receiving the *delegation request* shall compute the degree-of-matching between the requested capability coming in the CIL message and its own capabilities. The degree-of-matching *M* can be computed using the algorithm (pseudo code) shown in Figure 5. Basically, each element *e* in the *CIL-message* is searched within the local list of capabilities of the prospective *delegatee*. Three situations may occur:

- if *e* is *identical* to a capability of the delegatee (first condition), the prospective *delegatee* is probably similar to the *delegator*. Hence, the value of *M* is incremented by two;
- if *e* is *found* to be similar to one of the capabilities of the delegatee (second condition), the prospective *delegatee* has similar capability but may not be in the same way as in the *delegator* (e.g., browser of different type). Hence, the value of *M* is incremented by one;
- if *e* is not found at all, the value of *M* is decremented by one.

```
Line 1 | Set M →0, i→1
Line 2 | Read Local_capability []
Line 2 | for each element e in CIL_message
Line 2 |      Search for e in Local_capability
Line 3 |      If identical found
Line 4 |           M:=M+2
Line 5 |      Else if found
Line 6 |           M:=M+1
Line 7 |      Else
Line 8 |           M:=M-1
```

Figure 5. Simplified algorithm for calculating degree of matching (*M*).

4) Selecting and appointing delegatee

Once the value for degree-of-matching *M* is received from each prospective *delegatee*, the *delegator* will select the *delegatee* that responds with the largest *M* value. In due process, the presentational *CIL-description* will be sent to the selected *delegatee*.

5) Mapping UI/interaction modality

Once a peer is appointed and receives the presentational *CIL-description*, the *delegatee* will replace the *CIL-description* with its local capability. For example, the description of an HTML tags for *single line text* input tag of a web interface in Figure 6 (a) can be mapped to Figure 6 (b) that of an a multiline *text box* or vice versa. In the mapping process, either part of the description is depicted to fit the local capability, or it could be expanded. However, the major structure descriptor and the state of the widget are maintained as-is (see the bold and underlined part).

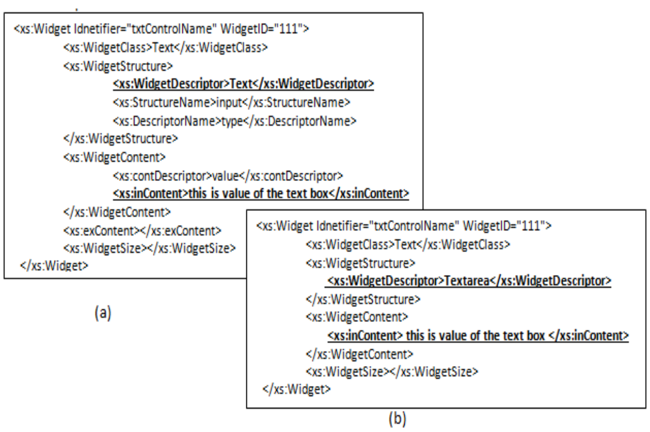


Figure 6. Description of UI concept using different capability

Nevertheless, UI elements that should not change the original structure should not pass through the delegation process. For example, some text inputs (e.g., username) might be required to be just one line.

Therefore, to correctly perform the mapping, a standard has to be followed between the *delegator* and the *delegatee* on how to describe the widgets (or other UI-related information). Hence, the CIL plays important role during the mapping.

IV. CONCLUSION

UI distribution and UI migration are two prominent dimensions that are useful to support the development and usage of cross-platform UIs. These concepts are often applicable if a client-server model is followed and it is not desired to have autonomous platforms. Thus, a dedicated server has to be assigned to orchestrate the distribution or migration. Furthermore, a server of this sort requires higher degree of reliability; otherwise, it can be a point of failure that jeopardizes the entire cross-platform operation. In addition, most works following these approaches focus primarily on a specific context (device, user, system, task, etc.). However, the use of *human*, *system*, and *device* views should be apparent, and shall not be split, so that the usability of the system can be improved even in cross-platform UI development.

Therefore, we draw from the theory of delegation – which is most applicable in agent-based system – the new concept of *UI delegation* as the third dimension in cross-platform UI development. In our work, we propose the UI delegation concept to follow peer-to-peer approach, so as to assure that peers remain autonomous. In due process, protocol for UI-related information exchange is important. Accordingly, we have discussed the notion of CIL together with the process of *UI delegation.*

Therefore, we claimed that if each peer is able to describe and maintain its capabilities and constraints, then new peers can be added easily. In due process, we consider *i*-HCI as interaction modality, which can be defined by the amalgamation of contextual information and intelligent technology. Thus, it is one dimension to be satisfied by the use of *CIL-messaging*, as per the *CIL-description*, which is built from the human, device and system points of view. Furthermore, considering delegation as per the capabilities of the *delegatee* peer would help to define autonomous peers, which are limited by the capabilities and constraints defined at the server.

In order to materialize the concept of *UI delegation,* in the future, more work has to be done to complete standardization of the CIL. It is also important to define a framework for CIL-enabled peers. In due course, though the computational power of small handheld devices is higher than ever, in the future, it is important to address the performance aspect of the delegation process since *delegatee* peers are responsible for mapping the UI description into their context.

REFERENCES

[1] K. Byeong-Ho, "Ubiquitous computing environment threats and defensive measures," International Journal of Multimedia and Ubiquitous Engineering, vol. 2, no. 1, 2007, pp. 47-60.

[2] A. Greenfield, Everyware: The Dawning Age of Ubiquitous Computing. Berkeley, USA: New Riders Publishing, 2006.

[3] M. Weiser, "The computer for the 21st century.," Scientific American, vol. 263, no. 3, 1991, pp. 94-104.

[4] H. Sørensen, D. Raptis, J. Kjeldskov, and M. B. Skov, "The 4C framework: principles of interaction in digital ecosystems," ACM International Conference on Pervasive and Ubiquitous Computing (UbiComp 2014), ACM Press, 2014, pp. 87-97, doi: 10.1145/2632048.2636089.

[5] S. Posland, Ubiquitous Computing: Smart Device, Environment, and Interactions. Chicester, UK: John Wiley& Sons Ltd, 2009.

[6] D. J. Mayhew, The Usability Engineering Lifecycle: a Practitioner's Handbook for User Interface Design. San Francisco, USA: Morgan Kaufmann Publishers, 1999.

[7] S. Berti, F. Paternò, and C. Santoro, "A taxonomy for migratory user interfaces," Interactive Systems. Design, Specifcation, and Verification Lecure Notes in Computer Science, Springer, 2006, pp. 149-160.

[8] F. Paternò and C. Santoro, "A logical framework for multi-device user interfaces," The 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2012), ACM, 2012, pp.45-50, doi:10.1145/2305484.2305494.

[9] J. Nielsen, Usability Engineering. Cambridge, MA, USA: Academic Press, 1993.

 ISBN: 978-1-61208-382-7

[10] Z. Lei, G. Bin, and L. Shijun, “Pattern based user interface generation in pervasive computing,”The 3rd International Conference on Pervasive Computing and Applications (ICPCA 2008), IEEE Press, vol. 1, 2008, pp. 48-53, doi: 10.1109/ICPCA.2008.4783636.

[11] J. Nichols and B. A. Myers, "Creating a lightweight user interface description language: an overview and analysis of the Personal Universal Controller project," ACM Transactions on Computer-Human Interaction (TOCHI), vol. 16, no. 4, 2009, article 17.

[12] P. Sauter, G. Vogler, G. Specht, and T. Flor, "A Model–View–Controller extension for pervasive multi-client user interfaces," Personal and Ubiquitous Computing, vol. 9, no. 2, 2005, pp. 100-107.

[13] G. Calvary et al., "The CAMELEON reference framework. Deliverable 1.1," 2002.

[14] F. Paternò, C. Santoro, and L. D. Spano, "MARIA: A universal, declarative, multiple abstraction-level language for service-oriented applications in ubiquitous environments," ACM Transactions on Computer-Human Interaction (TOCHI), vol. 16, no. 4, 2009, article 19.

[15] K. P. Fishkin, T. P. Moran, and B. L. Harrison, "Embodied user interfaces: towards invisible user interfaces," The IFIP TC2/TC13 WG2.7/WG13.4 7th Working Conference on Engineering for Human-Computer Interaction, IFIP, vol. 22, 1999, pp. 1-18, ISBN:0-412-83520-7.

[16] R. Castelfranchi and C. Falcone, "Towards a theory of delegation for agent-based systems," Robotics and Autonomous Systems, vol. 24, no. 3, 1998, pp. 141-157.

[17] E. G. Nilsson, J. Floch, S. Hallsteinsen, and E. Stav, “Model-based user interface adaptation,” Computer and Graphics, vol. 30, no. 5, 2006, pp. 692-701.

[18] M. Welie, “Task-based user interface design,” SIKS Dissertation Series No. 2001-6, Dutch Graduate School for Information and Knowledge Systems, Vrije Universiteit, 2001.

[19] J. Vanderdonckt, "Distributed user interfaces: how to distribute user interfaces elements across users, platforms and Environments," The 11th Congreso Internacional de Interacción Persona–Ordenador (Interacción 2010), Universidad Politécnica de Valencia, Valencia, Spain, Sept. 2010, pp. 3-14.

[20] J. Melchior, D. Grolaux, J. Vanderdonckt, and P. Roy, "A toolkit for peer-to-peer distributed user interfaces: concepts, implementaiton and applications," The 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2009), ACM Press, 2009, pp. 69-78.

[21] N. Elmqvist, "Distributed user interfaces: state of the art," J. A. Gallud, R. Tesoriero, V. M.R. Penichet (eds.), Distributed User Interfaces: Designing Interfaces for the Distributed Ecosystem, Human-Computer Interaction Series. London, UK: Springer-Verlag, 2011, pp. 1-12.

[22] L. Frosini and F. Paternò, "User interface distribution in multi-device and multi-user environments with dynamically migrating engines," The 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2009), ACM Press, 2009, pp. 55-64.

[23] D. Thevenin and J. Coutaz, "Plasticity of user interfaces: framework and research agenda," M. A. Sasse and C. Johnsson (eds.), The 7thIFIP Conference on Human-Computer Interaction (INTERACT 1999). IOS Press, 1999, pp. 110–117, ISBN: 09673355074274903087.

[24] K. Gajos and D. S. Weld, "SUPPLE: automatically generating user interfaces.," The 9th International Conference on Intelligent User Interfaces (IUI 2004), Funchal, Madeira, Portugal. ACM Press, Jan. 2004, pp. 93-100.

[25] D. Grolaux, P. Roy, and J. Vanderdonckt, "Migratable user interfaces: beyond migratory interfaces," The 1st Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MOBIQUITOUS 2004).Boston, Massachusetts, USA: IEEE Press, 2004, pp. 422-430, doi: 10.1109/MOBIQ.2004.1331749.

[26] S. K. Badam, E. Fisher, and N. Elmqvist, "Munin: apeer-to-peer middleware for ubiquitous analytics and visualization spaces," IEEE Transactions on Visualization and Computer Graphics, vol. 21, no. 2, 2015, pp. 215-228.

[27] A. Haddadi, "Communication and cooperation in agent systems: a pragmatic theory," Lecture Notes in Computer Science, vol. 1056, 1996.

[28] A. Q. Pham, "Privilege delegation and revocation for distributed pervasive computing environments," G. Abraham and B. I. P. Rubinstein (eds.), The 2nd Australian Undergraduate Students’ Computing Conference, 2004, pp. 136-141, ISBN: 0-975-71730-8.

 ISBN: 978-1-61208-382-7

Human Input about Linguistic Summaries in Time Series Forecasting

Katarzyna Kaczmarek
Systems Research Institute
Polish Academy of Sciences
Newelska 6
01-447 Warsaw, Poland
K.Kaczmarek@ibspan.waw.pl

Olgierd Hryniewicz
Systems Research Institute
Polish Academy of Sciences
Newelska 6
01-447 Warsaw, Poland
Olgierd.Hryniewicz@ibspan.waw.pl

Rudolf Kruse
Faculty of Computer Science
Otto-von-Guericke-University Magdeburg
Universitaetsplatz 2
D-39106 Magdeburg, Germany
Rudolf.Kruse@ovgu.de

***Abstract*—Finding an appropriate predictive model for time series and formulating its assumptions may become very challenging task. We propose to represent time series in a human-consistent way using linguistic summaries. Such summaries describe general trends in time series and are easily interpretable for decision makers. The aim of this contribution is to show that the linguistic summaries may be successfully applied to support the analysis and forecasting of time series. Information about trends is first retrieved from experts, and then, processed with soft computing tools. The performance of the approach is verified on the real-world datasets from the M3-Competition. Users are asked to evaluate linguistic summaries that are intuitive and easy for interpretation. This paper shows that human-consistent summaries deliver new knowledge for forecasting.**

Keywords–information retrieval; human-computer interaction; time series and sequence models; Bayesian methods; supervised learning.

I. INTRODUCTION

Practitioners are very often posed to the dilemma of choice between the wealth of mathematic models for forecasting for the imprecise real-world data. For a recent review of competitive forecasting models and methods, see e.g., Gooijer and Hyndman [1]. Within this research, the Box-Jenkins models are adapted. They are simple, and though, have been proven successful in various practical applications [2]–[4].

An important task of the Box-Jenkins time series analysis is the estimation of the unknown variables. One of the potential approach to this estimation is the Bayesian inference, that enables the inclusion of subjective prior information. Following Geweke [5], definitions for the prior probability distributions are usually assumed basing on expert's experience and intuitions, and normal or uniform distributions are often appropriate. However, experts may fail to adequately establish the prior distributions for the unknown variables and models, and then the problem arises. To conclude, the ability to describe the data imprecision in terms of prior probability distributions is one of the main advantages of the Bayesian approach, and at the same time, the main challenge for practitioners, because models may be difficult to understand for non-mathematician experts. Also, the proper selection of prior probability distributions is essential for the satisfactory forecasting performance.

Therefore, we propose to retrieve from experts the information about the expected trends in time series, and then, formulate the prior probability distributions automatically basing on this natural language information. The proposed approach realizes innovative incorporation of linguistic summaries into time series analysis.

The objective of this paper is to present this approach that consists of the human-computer interaction for the information retrieval, and then, its incorporation into the time series analysis and forecasting process. It assumes employing techniques from the following research fields: time series analysis and forecasting, the fuzzy set theory, the time series summarization and pattern mining, classification methods and the Bayesian analysis.

The comparative analysis of the forecasting accuracy is performed on time series from the M3-Competition by Makridakis and Hibon [6]. Experiments prove that the human-consistent summaries deliver new knowledge for forecasting.

The structure of this paper is as follows. Next section introduces basic definitions of the time series analysis. In Section III, the description of the human-computer interaction related to the linguistic summaries is explained. Section IV presents the proposed approach to incorporate the retrieved human-consistent knowledge into the Bayesian forecasting. Numerical results are gathered in Section V. In Section VI, conclusions are presented.

II. PRELIMINARIES: TIME SERIES ANALYSIS

Discrete time series $y = \{y_t\}_{t=1}^{n} \in Y$ is a sequence of observations measured at successive $t \in T = \{1, ..., n\}$ moments and at uniform time intervals, e.g., a sequence of monthly sales for a specific product builds up a time series.

A. Box-Jenkins Methodology

Due to the Box-Jenkins methodology [7], the time series analysis starts with the identification of the probabilistic model that generated the observed time series.

One of the most popular stationary models for forecasting, and though, very successful in applications are autoregressive and moving average (ARMA) models, defined as follows:
ARMA(p,q) [7]

$$\tilde{y_t} = \sum_{i=1}^{p} \phi_i \tilde{y_{t-i}} + \sum_{i=1}^{q} \alpha_i a_{t-i} + a_t; a_t \sim N(0, \sigma^2); \tilde{y_t} = y_t - \mu \quad (1)$$

where $\theta = \{\phi_1, ..., \phi_p, \alpha_1, ..., \alpha_q, \mu, \sigma^2\}$ *is the vector of unknown variables.*

ARMA models are based on the concept of a linear filter assuming that the observations are generated by the sequence $\{a_t\}_{t=1}^{n}$ of values taken by the independent and identically distributed random variable (white noise).

The class of autoregressive and moving average processes is rich, and in many contexts, it is usually possible to find a process or a combination of processes which provide an adequate description to the considered real-life time series data.

The Box-Jenkins approach to the time series analysis is an iterative process. After the identification of the probabilistic model, the following steps are performed: estimation of its parameters, verification techniques and finally, prediction. For further details, we refer to [2].

B. Bayesian Model Averaging

To diminish the risk of selecting one non-adequate model, multiple models may be combined through the Bayesian inference. Clemen and Winkler [8] show that combining various methods for forecasting on average leads to better results than applying individual ones.

The Bayesian model averaging enables to include multiple models and the posterior density $p(y_{n+h}|y, M)$ is a weighted average of the posterior densities of models $\{M_1, M_2, ..., M_J\}$:

$$p(y_{n+h}|y, M) = \sum_{j=1}^{J} p(y_{n+h}|y, M_j)p(M_j|M) \qquad (2)$$

The Bayesian averaging requires defining the prior model probability distributions $p(M_j|M)$. In [9], Ley and Steel show by theoretical and empirical evidence the critical importance of prior assumptions for the Bayesian model averaging. Within the proposed approach, these prior model probability distributions are automatically generated basing on the human-computer interaction.

III. HUMAN INPUT ABOUT LINGUISTIC SUMMARIES

In many domains, it is important to deliver results that are simple and easy to interpret by user. One may provide various forms of human-consistent descriptions of large datasets with the use of data mining and knowledge discovery techniques, and the literature on discovery of different information granules about time series data is extensive [4], [10]–[12]. Linguistic summaries are an example of information granules, and mining for linguistic summaries has also gained a lot of attention in the literature [13]–[15].

Within the proposed approach, we adapt the linguistic summaries in the sense of Yager [16] developed by Kacprzyk et al. [13], and we use the fuzzy set theory as introduced by Zadeh [17] to model the data imprecision.

A. Linguistic Summary Definition

Linguistic summaries describe general trends about the evolution of time series with quasi natural language, e.g., *Most increasing trends are short*.

Linguistic summary [16]
Let $A = \{a_1, a_2..., a_u\}$ denote a finite set of attributes (e.g., dynamics of change, duration). $S = \{l_1, l_2..., l_s\}$ is a finite set of imprecise labels for attributes (e.g., *quickly increasing trends, short trend*). The protoform-based linguistic summary

$$LS : Q\ R\ trends\ are\ P \qquad (3)$$

consists of quantifier Q (e.g., *most, among all*), summarizer P (attribute together with an imprecise label), qualifier R.

The imprecise labels refer to linguistic values of either qualitative or quantitative measurements for attributes (e.g., *low, increasing, short*). The interpretation for imprecise labels is modeled as fuzzy trapezoidal numbers. For further definitions, refer to e.g., [18], [19].

The quality of a linguistic summary is evaluated with **degree of truth (validity)** T due to [20], defined as follows:

$$T(LS) = \mu_Q\left(\frac{\sum_{i=1}^{n} (\mu_R(y_n) \wedge \mu_P(y_n))}{\sum_{i=1}^{n} \mu_R(y_n)}\right) \qquad (4)$$

where $\mu_R(y_n), \mu_P(y_n)$ are the membership functions $\mu_R, \mu_P : \Re \to [0, 1]$ determining the degree to which R, P, respectively, are satisfied for the time series y at the given moment n.

B. Linguistic Summary Retrieval

The following attributes and labels defining trends are considered to build up the linguistic summaries: duration (*short, medium, long*), dynamics (*increasing, constant, decreasing*) and variability (*low, moderate, high*). The resulting set of linguistic summaries may be exemplified by:
Most decreasing trends are medium;
Most trends are constant;
Most trends are decreasing.

If a time series is long, then the linguistic summaries are generated and evaluated automatically, e.g., with the Trend Analysis System [21]. However, at the beginning of the data collection process, if the available time series is very short, then the automatic results may be unreliable. Therefore, experts could be employed to validate the quality of linguistic summaries.

Let $T_E : LS \to [0, 1]$ denote subjectively defined quality evaluation function that maps linguistic summaries to the interval [0,1].

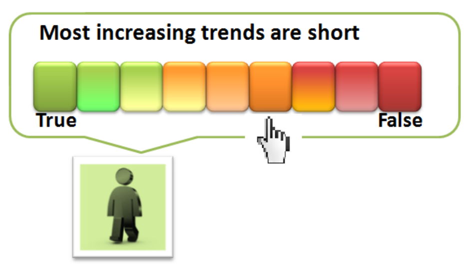


Figure 1. Evaluating quality of linguistic summaries.

As presented in Figure 1, the simple natural language expressions, e.g., *Most increasing trends are short* are presented to the decision maker who points his confidence that this summary is true about the considered time series. The values of T_E are interpreted as the expert's degree of confidence that the linguistic summary is true.

 ISBN: 978-1-61208-382-7

IV. FORECASTING WITH LINGUISTIC SUMMARIES

Algorithm 1 described below presents a high-level description of the proposed *Forecasting with Linguistic Summaries* (F-LS) approach.

Algorithm 1 Forecasting with Linguistic Summaries (F-LS) provides prediction y_{n+1}

Input:
y: $y = \{y_t\}_{t=1}^{n}$, $n \in \{n_{min}, ..., n_{max}\} \subseteq N$, $y \in Y$, where Y is a space of discrete time series
S: set of imprecise labels
$M = \{M_1, M_2, ..., M_J\} \subseteq M$: template probabilistic models where M is a set of stationary autoregressive processes
Output: y_{n+1}
Algorithm:

1: *Defining of imprecise concepts:*
2: build_fuzzy_numbers (S)
3: *Data preprocessing:*
4: **repeat** difference(y) **until** y *is* validated
5: min-max normalization(y)
6: *Supervised learning for the training database:*
7: **while** $i \in J$ **do**
8: T_m^s, C^s = generate k sample time series (M_i, k, m)
9: LI^s = discover_linguistic_summaries (T_m^s)
10: V^s = calculate_degree_of_truth (LI^s)
11: CL = supervised_learning_withSVM (C^s,V^s)
12: *Imprecise knowledge retrieval from humans:*
13: LI^E = create_provisional_linguistic_summaries (y)
14: v^E = calculate_degree_of_truth (LI^E)
15: T^E = expert_evaluation (LI^E, v^E) *EXPERT INPUT*
16: **while** $i \in J$ **do**
17: Sc^{M_i} = estimate_classification_scores (T^E, CL)
18: *Posterior simulation and forecasting:*
19: P=construct_prior_prob_distr (M, Sc^M)
20: y_{n+1}=MCMC_posterior_simulation (P, y)

The input for the algorithm is the discrete time series for prediction y and the set of template probabilistic models M, that need to be defined a priori. Within this research, we focus on supporting forecasting of short time series assuming that $n_{min} = 10, n_{max} = 20$.

The algorithm starts from the definition of imprecise concepts that describe the trends and linguistic summaries (Line #1). Secondly, the preprocessing of the time series data (Line #3) is performed to ensure that they are normalized and without missing values. Next, the supervised learning of the probabilistic models (Line #6) is executed. Its goal is to build the training database and to discover rules enabling the classification of the probabilistic models based on the sets of linguistic summaries describing the evolution of time series.

Then, the mining for the human-consistent prior information (Line #12) is performed. Its goal is to discover and validate with experts the linguistic summaries about the expected evolution of the predicted time series. Next, the prior probability distributions are calculated (Line #19). Finally, Markov Chain Monte Carlo Posterior Simulation is run (Line #20) to simulate the posterior probability distributions for the vector of interest and calculate the forecast y_{n+1}.

V. NUMERICAL RESULTS

The experimental study aims at showing the forecasting accuracy of the proposed *Forecasting with Linguistic Summaries* (F-LS) approach among other forecasting methods. The results are presented for the real-life benchmark time series data.

We use the subset of the 10 first yearly time series (N1-N10) that have length 20 from the M3-Competition Datasets Repository [6]. The performance of the proposed F-LS approach is compared to best 13 benchmark methods studied in [6]. These methods are briefly presented in Table I. Methods marked with * are commercially available in forecasting packages. The forecast accuracy is measured by Symmetric Mean Absolute Percentage Error (sMAPE).

Table II shows the medal classification based on sMAPE. Is is observed that the proposed F-LS forecast is number one for two series and has never performed worst. Only the Robust-Trend forecast has also been number one for two series, and number two for another two time series. Nonetheless, it has also been the worst for one series.

TABLE II. MEDAL CLASSIFICATION. TOP-3 AND THE WORST FORECASTING METHOD FOR N1-N10 TIME SERIES FROM M3-C DATASET.

	TOP-3				WORST
Method	**I**	**II**	**III**	**...**	
ForecastX	1	0	1		0
F-LS	**2**	**0**	**1**		**0**
Comb S-H-D	0	0	0		0
Robust-Trend	2	2	0		1
Theta	0	1	3		0
RBF	0	1	1		0
Auto-ANN	2	0	0		1
ForecastPro	0	0	1		1
B-J Auto	0	0	0		1
Naive2	1	1	1		0
Single	0	1	1		0
SmartFcs	1	1	0		2
ARARMA	0	1	1		1
Flores /Pearce2	0	1	0		2

Details about sMAPE for the 1 step horizon are gathered in Table III.

As demonstrated by the results in Table III, none of the benchmark methods outperforms or dominates the proposed F-LS method for all 10 time series. The best average sMAPE result of 6.1 is achieved by ForecastX method. At the same time, it is observed that for 4 (N1, N6, N7 and N9) out of all 10 time series the proposed F-LS approach delivers more accurate forecast than the ForecastX.

F-LS provides forecasts which are similarly accurate to the ones provided by Comb S-H-D, Robust-Trend, Theta and RBF methods. The average sMAPE amounts to 6.7 for F-LS, and 6.7, 6.8, 6.8, 6.9 for the other methods, respectively.

We conclude that the proposed approach delivers very competitive results in terms of the forecasting accuracy.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have discussed the human-consistent input to support the analysis and forecasting of time series. We have proposed a new approach consisting of the human-computer interaction for the retrieval of the natural language summaries and their application for the Bayesian forecasting.

One of the main advantages of the proposed solution is its interpretability, which is of special importance for experts

TABLE I. SELECTED BENCHMARK METHODS FROM THE M3-COMPETITION.

Method	Author	Description
Naive2	M. Hibon	Deseasonalized Naive (Random Walk)
Robust-Trend	N. Meade	Trend model - Non-parametric version of Holt's linear model with median based estimate of trend
Flores /Pearce2	B.Flores, S. Pearce	Expert system that chooses among four methods based on the characteristics of the data
SmartFcs*	C. Smart	Expert System - conducts a forecasting tournament among four exponential smoothing and two moving average methods
Theta	V. Assimakopoulos	Decomposition technique - projection and combination of the individual components
Comb S-H-D	M. Hibon	Trend model - combining three methods: Single / Holt/ Dampen
ARARMA	N. Meade	ARIMA models - Automated Parzen's methodology with Auto regressive filter
Single	M. Hibon	Single Exponential Smoothing
ForecastX*	J. Galt	Expert System - selects from among several methods
RBF	M. Adya, S. Armstrong, F. Collopy, M. Kennedy	Rule-based forecasting: using random walk, linear regression and Holt's to estimate level and trend, involving corrections, simplification, automatic feature identification and re-calibration
ForecastPro*	R. Goodrich, E. Stellwagen	Expert System - Expert System - selects from among several methods
Auto-ANN	K. Ord, S. Balkin	Automated Artificial Neural Networks
B-J Auto	M. Hibon	ARIMA models - Box-Jenkins methodology of 'Business Forecast System'

TABLE III. SMAPE FORECASTING ACCURACY FOR N1-N10 TIME SERIES FROM THE M3-COMPETITION. F-LS IS THE PROPOSED METHOD, OTHER ARE BENCHMARK.

Method	TS-N 1	TS-N 2	TS-N 3	TS-N 4	TS-N 5	TS-N 6	TS-N 7	TS-N 8	TS-N 9	TS-N 10	Avg sMAPE
ForecastX	1.8	10.6	15.7	4.1	4.0	1.7	5.0	0.7	16.9	0.7	6.1
F-LS	**0.2**	**10.9**	**18.6**	**7.5**	**6.3**	**0.7**	**1.6**	**10.4**	**9.2**	**1.0**	**6.7**
Comb S-H-D	2.0	12.8	14.8	3.0	3.2	3.5	1.6	7.4	17.3	1.5	6.7
Robust-Trend	3.3	6.1	19.9	8.7	8.4	0.2	5.3	4.2	11.5	0.1	6.8
Theta	0.6	7.1	21.4	4.5	2.5	4.6	0.7	13.3	12.8	0.6	6.8
RBF	3.1	12.0	17.2	8.5	2.3	0.8	3.0	8.5	12.2	1.3	6.9
Auto-ANN	1.4	8.7	5.1	11.9	5.3	9.7	0.3	6.1	20.3	3.3	7.2
ForecastPro	2.0	12.6	13.9	0.5	4.0	1.7	4.5	14.3	20.1	0.8	7.4
B-J Auto	2.0	12.6	18.2	0.5	5.0	2.1	0.7	6.1	22.3	5.1	7.4
Naive2	8.6	12.5	13.8	0.5	4.4	5.8	0.7	6.1	20.1	5.0	7.7
Single	8.6	12.5	13.8	0.5	4.4	5.8	0.7	6.1	20.1	5.0	7.7
SmartFcs	2.3	1.3	24.4	9.3	5.0	1.4	5.0	1.1	30.8	4.0	8.5
ARARMA	3.2	11.1	14.7	4.7	4.1	0.4	24.3	3.5	17.9	3.1	8.7
Flores /Pearce2	10.3	11.1	13.8	20.5	3.1	5.8	1.2	9.7	16.0	1.3	9.3

involved in the forecasting process. Instead of providing definitions of prior probability distributions, users are asked to evaluate linguistic summaries that are intuitive and easy for interpretation.

The performance of the proposed approach is illustrated with the experimental study for benchmark datasets. The numerical results of the forecast accuracy show that the proposed approach of combining human input about linguistic summaries and Box-Jenkins models through the Bayesian averaging may lead to the increase of the accuracy compared to the competitive methods. Although the human input is highly subjective, it helps to eliminate the need to express the assumptions as prior probability distributions, which may be difficult to understand for non-mathematician decision makers.

Further experiments on other benchmark datasets are planned to analyze all advantages and disadvantages of the proposed approach. Future research also assumes the analysis of other forms of imprecise information like fuzzy classification rules and frequent temporal patterns, and the modeling of multiple imprecise labels interpretations.

ACKNOWLEDGEMENTS

The authors thank Janusz Kacprzyk and Anna Wilbik for their support and advice concerning the Trend Analysis. Katarzyna Kaczmarek is supported by the Foundation for Polish Science under International PhD Projects in Intelligent Computing financed from the European Union within the Innovative Economy Operational Programme 2007-2013 and European Regional Development Fund.

REFERENCES

[1] J. D. Gooijer and R. J. Hyndman, "25 years of time series forecasting," International Journal of Forecasting, vol. 22, 2006, pp. 443–473.

[2] G. Box, G. Jenkins, and G. Reinsel, Time Series Analysis: Forecasting and Control, 4th Edition. Wiley, 2008.

[3] P. D'Urso, D. Lallo, and E. Maharaj, "Autoregressive model-based fuzzy clustering and its application for detecting information redundancy in air pollution monitoring networks," Soft Computing, 2013, pp. 83–131.

[4] O. Hryniewicz and K. Kaczmarek, "Bayesian analysis of time series using granular computing approach," Applied Soft Computing, 2014.

[5] J. Geweke, "Contemporary bayesian econometrics and statistics," Wiley series in probability and statistics, 2005.

[6] S. Makridakis and M. Hibon, "The m3-competition: results, conclusions and implications," International Journal of Forecasting, 2000, pp. 451–476.

[7] G. Box and G. Jenkins, Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, 1970.

[8] R. Clemen and R. Winkler, "Combining probability distributions from experts in risk analysis," Risk Analysis, vol. 19(2), 1999, pp. 187–203.

[9] E. Ley and M. Steel, "On the effect of prior assumptions in bayesian model averaging with applications to growth regression." Journal of Applied Econometrics, vol. 24, 2009, pp. 651–674.

[10] D. Nauck and R. Kruse, "Obtaining interpretable fuzzy classification rules from medical data," Artificial Intelligence in Medicine, vol. 16(2), 2014, pp. 149–169.

[11] S. Kempe, J. Hipp, C. Lanquillon, and R. Kruse, "Mining frequent temporal patterns in interval sequences," Fuzziness and Knowledge-Based Systems in International Journal of Uncertainty, vol. 16 (5), 2008, pp. 645–661.

[12] J. Yao, A. Vasilakos, and W. Pedrycz, "Granular computing: Perspectives and challenges," IEEE Transactions on Cybernetics, vol. 43(6), 2013, pp. 1977–1989.

[13] J. Kacprzyk, "Linguistic summarization of time series using a fuzzy

quantifier driven aggregation," Fuzzy Sets Syst, vol. 159 (12), 2008, pp. 1485–1499.

[14] C. Moewes and R. Kruse, "Zuordnen von linguistischen ausdrücken zu motiven in zeitreihen (matching of labeled terms to time series motifs)," Automatisierungstechnik, 2009, pp. 146–154.

[15] K. Kaczmarek and O. Hryniewicz, "Linguistic knowledge about temporal data in bayesian linear regression model to support forecasting of time series," in Proc. of Federated Conference on Computer Science and Information Systems, 2013, pp. 655 – 658.

[16] R. Yager, "A new approach to the summarization of data," Information Science, vol. 28 (1), 1982, pp. 69–86.

[17] L. Zadeh, "Fuzzy sets," Information and Control, 1965, pp. 338–353.

[18] M. Gil and O. Hryniewicz, "Statistics with imprecise data," Encyclopedia of Complexity and Systems Science, 2009, pp. 8679–8690.

[19] R. Kruse, C. Borgelt, F. Klawonn, C. Moewes, M. Steinbrecher, and P. Held, Computational Intelligence. Texts in Computer Science. Springer London, 2013, ch. Fuzzy Sets and Fuzzy Logic.

[20] L. A. Zadeh, "A computational approach to fuzzy quantifiers in natural languages," Computers and Maths with Applications, 1983, pp. 149–184.

[21] J. Kacprzyk, A. Wilbik, A. Partyka, and A. Ziółkowski, Trend Analysis System. Systems Research Institute, Polish Academy of Sciences, Warsaw, 2011.

Modelling Volo, an Augmentative and Alternative Communication application

Antonina Dattolo
SASWEB Lab, Dept. of Mathematics and Computer Science
Università di Udine, Gorizia, Italy
email: antonina.dattolo@uniud.it

Flaminia L. Luccio
Dept. of Environmental Sciences, Informatics and Statistics
Università Ca' Foscari Venezia, Venezia, Italy
email: luccio@unive.it

***Abstract*—In this paper, we present a formal representation of an Augmentative and Alternative Communication (AAC) application, called Volo, for users with Autism Spectrum Disorders (ASD). We discuss existing AAC applications and present a formal visualisation model of Volo based on zz-structures, hyper-orthogonal, non-hierarchical structures for storing, linking and manipulating data. To the best of our knowledge, this is the first work that tries to give a formal model to AAC techniques. We finally present the prototype of this new Volo application.**

Keywords–Zz-structures; mobile app; Augmentative and Alternative Communication; Autism Spectrum Disorders.

I. Introduction

Language delays and impairments are one of the common characteristics of children with Autism Spectrum Disorders (ASD) [1]. However, these children often show good visual abilities [2], thus in the last years standard speech therapies have been combined with Augmentative and Alternative Communication (AAC) techniques: they are powerful methods that combine different visual components in order to create syntactically and semantically correct sentences.

AAC techniques can be of different nature: they can either be based on unaided systems, i.e., rely on the use of body actions such, e.g., gestures, sign language, etc., or on aided systems which are based either on lite technology such as Picture Exchange Communication System (PECS) books, letter boards, or on high technology such as dedicated devices containing images and sounds [3] (see Figure 1).

A. History of AAC systems

The first examples of Augmentative and Alternative Communication (AAC) systems go back only to 1960, when letter and picture boards, and typewriters where first introduced as a way of communicating by very few nonspeaking individuals [4]. Issues that deeply impacted the development of AAC devices were: in 1975, the approval by the Congress of the Public Law 94-142 (Education of All Handicapped Children Act), which assured a free appropriate public education to all children with disabilities; the development and diffusion of microcomputer technology, given that prior to that these devices were handmade electrical systems, or non-portable computer systems [5].

The first example of a commercially available dedicated AAC device is the Canon Communicator from Canon Inc., a small device with a keyboard and a strip printer. After that researchers concentrated on portable voice synthesiser leading to the production, in 1978, of the first commercial AAC device with speech synthesis, Handivoice, from the Federal Screw Works and Phonic Ear. Nowadays, given the big advances in microcomputer technology, AAC devices have become more and more sophisticated and adaptable for all the wide variety of users [6].

Figure 1. An example of sign language (left) and of a PECS book (right).

In this paper, we are interested in formal models and visualising techniques for high technology AAC devices, so we will concentrate on them. In particular, we will consider dedicated applications for mobile systems, like smartphones and tablets.

The study of current literature highlighted that, independently from the type of app which is chosen, there are some basic principles followed by all of them: the main feature is a picture-based approach, given that it is proved to be effective on children that demonstrate the comprehension of cause and effect actions. A training approach that promotes symbolic communication and is based on lite technology is the PECS, developed in 1985 by Lori Frost and Andy Bondy. This approach works well for those children that demonstrate emerging intentional communication as it promotes interactions among users [7]. An example of a PECS book, which is a velcro communication board, is shown in Figure 1-right. Each child with ASD is different and has different abilities, thus the use of this book, and in general any learning target, may differ.

There are different training levels which are introduced to the child by a specialised therapist. In Phase 1, the simplest one, the child initiates a request by removing a picture from the board (the unique one there), and by handing it to an adult in exchange for a desired item or action. When the child shows a good ability at this level, it moves to Phase 2, in which it has first to select the picture, and then physically bring it to a partner which is not nearby. In Phase 3, the child has to discriminate one picture among many other pictures (as in in Figure 1-right); in Phase 4, it has to build a sentence using a sequence of pictures, e.g., a picture representing the concept

 ISBN: 978-1-61208-382-7

"I want", and another picture representing the desired object; in Phase 5, it has to reply to questions such as "What do you want?", with a sequence of pictures. Finally, in Phase 6, it has to respond to questions and comment on items, e.g., using pictures which define concepts such as "I see", "I hear", "I feel", etc. [8].

The AAC apps that we will consider in the following are all based on this picture-based approach.

B. Features of AAC devices and scanning methodologies

High technology AAC systems have many different features: they can range from dedicated devices to specialised apps from different platforms; their size and weight may hardly affect their portability and accessibility, which are very important features for children with ASD which are typically ambulatory. The methods for storing and retrieving sounds and messages may differ, e.g., offering real or synthesised words as a spoken output. Other features are the access methods which determine how the user will select symbols on the AAC device, such as pointing with fingers, using eye gaze, etc.. Finally, the methods we are interested to analyse and consider in this paper are related to the visualisation, scanning and choice of images.

The way users may scan and select images differ from a device to another:

- *Automatic scanning* is the basic standard technique. The device presents different symbol choices while the user waits and, when the desired picture is reached, it is chosen by the user by a pressure. This can be done using 1-step scanning such as *linear scanning* in which pictures are scanned linearly, and *circular scanning* in which are scanned by following a circular pattern. However, both types of scanning work well when the number of presented pictures is limited [2]. The process can be speed-up using the *row-column scanning* (or sometimes the *column-row scanning*) which consists of 2-steps scanning where a row of pictures is presented at once, it is scanned and when the desired picture is reached the user has to select it.
- *Group-item scanning* derives from the row-column scanning and assumes the refinement in subcategories of pictures. Groups, i.e., classes of symbols are presented at once; the user chooses a group and then the pictures inside this group are presented. These pictures can be subgroups or final items. This scanning technique assumes a three- to four-steps scanning process and is thus the most difficult for users to manage, but at the same time it allows their the use of a wider range of images.
- *Directed scanning* has an opposite approach: the user keeps the switch pressed and s/he releases it when the item is reached. This allows the user to move his/her hand on the device without accidentally activating symbols; on the other hand, it requires the user to focus on the pictures for a long time.
- *Step scanning* assumes that the user presses the switch one step at a time to move the to the next symbol, and when the desired symbol is reached, s/he stops.

Comparing these selection methods, it turns out that automatic scanning requires less actions by the user (only a click when the item is reached), but the user has to patiently wait for the picture. The other methods require more physical ability of the user but increase his/her cognitive process and provide a direct feedback for an action. For all methods, note that the communication rate is still very low, at most 43 words per minute [9]. For this reason, many existing AAC systems use picture-based accelerating techniques. For example, the sentence "Hi, how are you", could be coded as a picture of a waving hand + a finger pointing, i.e., a sequence of symbols that code in a compact way a more complicated sentence [2].

The paper is organised as follows: in Section II, we present the characteristic of many existing apps for ASD children, and in particular we concentrate on the visualisation techniques. In Section III, we describe the Volo formal model, based on the the zz-structures, while in Section IV we present the prototype. Note that, to the best of our knowledge, this is the first work that tries to give a formal model to AAC techniques. We conclude in Section V.

II. RELATED WORK

In this section, we describe some of the commercial applications based on AAC techniques, which are used by children with ASD, and we will compare them from a visualisation point of view.

The commercially available apps are many and range from free to very expensive tools. There are applications for iOS (as [10]-[11]), for Android (as [10][12][11]), or for special dedicated devices (as [13][14]). We concentrate here on tools that can be used by children with strong communication delays and allow their the selection of images. The differences among these apps are first obviously in the graphics: some apps use sketched images, other real pictures, most of them provide the user with an initial set of pictures and allow the import of new images from a personal computer, a camera, etc. (as [10][15][12]-[11]).

Another feature is the possibility of adding sounds, which can be synthetic or natural (as in [10][16][17][12][11]), or can also be recorded (as in [16][15][12]). Most of the applications are also multilingual. The Niki-talk application [15] has also social features, i.e., it allows the user to type a message and eventually to tweet it.

Some apps allow the creation of calendars: the daily routine might be organised in sequence of actions which describe the activities of the day in a fixed temporal order as in Figure 2-right where the user visualises a sequence of actions to follow, in order to get dressed (as in [11][18]).

Although all these apps might have some different features, the visualisation characteristics seem to be common: in particular, all the tools we have analysed assume dynamic displays, i.e., the set of pictures which the user is able to visualise may dynamically change after a selection. Moreover, pictures are displayed in a grid form, i.e., in rows and columns, whose size can be, in most cases, customised.

In Niki-talk [15], the user may select between two categories of actions ("I want", "I am"), a keyboard icon that links a page where the user may type and hear words, and a paint brush icon that opens another page where the user can draw and save his/her drawings.These two icons (keyboard and paint brush) appear in each internal view together with items of the

 ISBN: 978-1-61208-382-7

Figure 2. An example of the Niki-talk app [15] on the left, and of the Pics Aloud calendar [18] on the right.

same categories - see Figure 2-left where the user selected the action “I want”.

In Proloquo2Go - Symbol-based AAC [17], images of different categories may appear together (as subject, verbs, etc.), and are differentiated by the border color. In simple views, only images of the same category are displayed inside the grid. The app displays in the top the sequence of selected images. Images may be rearranged by a user customisation, for example, inserting images in a given position, etc..

TTalk AAC [19] shows categories and simple images together, while Let’s talk [16] displays many categories at the beginning and it allows for two distinct uses: the user can be alone, or helped by an external person that suggests images. Pics Aloud [18] has a slightly different approach: it displays many real pictures of different categories, and, when the user selects the image, the system repeats loudly a whole sentence. For example, if the image is a swimming pool, when the user selects it the app says something like ”I want to go to the swimming pool”, i.e., it produces a sentence for the user.

Voice4u [11] may display a set of predefined images in an alphabetical order, otherwise by clicking on a special button, categories appear and the user may browse inside them.

Alexicon AAC [10] and TalpToTalk [12] are Web applications; they displays many categories and objects in a grid form. Alexicon allows sequential and row/column scanning, and also auditory preview scanning. TalpToTalk allows the extension of an image word to a whole sentence. GoTalk [13] and QuickTalker [14] are two examples of apps that run on a dedicated device. They both display images in fixed 2×2, or 9×9, etc. grid, containing objects, categories and objects.

III. A FORMAL MODEL FOR PECS-BASED APPS

In this section, we introduce the basic definition of a formal and general model for describing generic AAC-based applications, that can be used to formally describe all the applications mentioned in the previous Section II. In this paper we concentrate on the formalization of our model, the extension to all the other applications is an ongoing work.

The model is based on zz-structures which are a graph-centric system of conventions for data and computing, able to simply connect different elements through contextual, semantic connections [20][21]. Readers interested in a deep discussion on zz-structures, its formal description and practical applitudes can refer to [22]-[23].

Using these zz-structures is possible to simply describe knowledge representation, visualisation and the scanning process.

Definition: An *AAC-based app* is a zz-structure, $AAC\text{-}A = (MG, SL, sl)$ where

- $MG = (V, E, f)$ is a multigraph (a graph where pairs of vertices might have multiple edges connecting them) composed of:
  - a set of *vertices* V;
  - a set of *edges* E;
  - a surjective function $f : E \to \{\{u, v\} \mid u, v \in V, u \neq v\}$;
- SL is a set of semantic labels;
- $sl : E \to SL$ is an assignment of semantic labels to edges of the multigraph.

$AAC\text{-}A$ is subject to the following constraint: $\forall x \in V,\ \forall k = 1, 2, ..., |SL|,\ \ deg_k(x) = 0, 1, 2$ where $deg_k(x)$ indicates the *degree* of the vertex x (that is, the number of edges incident to it) labeled by the semantic label k.

In other words, an AAC-based app is an edge-labeled (equivalently edge-coloured) multigraph where the vertices are either singletons, or are connected by the labels in linear paths, or cycles.

Each vertex of a zz-structure is called a *zz-cell* and each edge is a *zz-link*. We can extract from the zz-structure a sub-graph, composed by all the connected components (linear paths or cycles) of a particular semantic label. Each connected component is called *rank*, while the sub-graph of ranks of the same label is called *dimension*. Each label is associated to a specific semantic context; for this reason, ranks and dimensions provide a semantic interpretation of the zz-structure.

An example of $AAC\text{-}A$ is shown in Figure 3, where normal, thick, dashed, dotted and double lines represent five different dimensions. (Note that, the images used here in Figure 3, and later for the presentation of the zz-structure and the prototype, have been downloaded from many different sites of the Web.) The first row (the dimension called *category* -

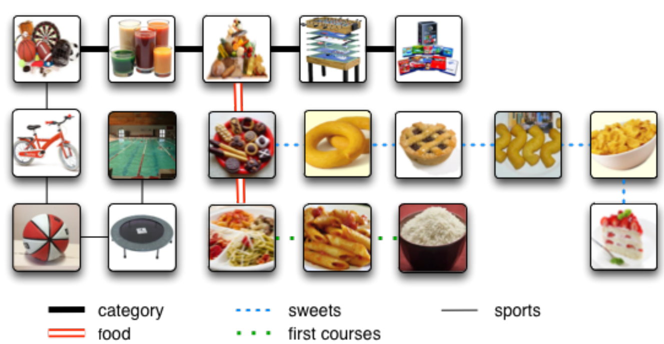


Figure 3. An example of a simple zz-structure.

thick line), collects the main categories of items (like food, sports, games, dvds), while the first zz-cells (sports and food) of the two vertical dimensions (identified respectively by normal and double lines) generalise the set of images contained in that dimension. They are called *maincells*, and their role in the dimension is different from the role of the other cells:

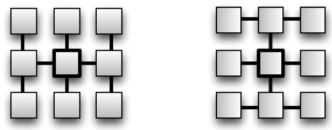


Figure 4. A schematic representation of a $3x3$ H-view and I-view.

Figure 5. An example of $5x5$H-view.

they symbolise the dimension, and represent their neighbouring cells. Analogously, also the second and third cells of the food dimension (sweets and first courses) are maincells: they identify two sub-categories of food.

An AAC-A may be visualised in different ways: in this paper, we use the H-view [24], a rectangular view (based on two dimensions, called primary and secondary) where the focus zz-cell is placed at the center of the view on a Cartesian plane (the cell with bold border), while the primary dimension occupies the horizontal central path, and, for each zz-cell, the zz-cells, connected to it,and belonging to the secondary dimension, occupy the vertical paths. A schematic example of a $3x3$ H-view is shown in Figure 4-right; the name H-view comes from the fact that the columns remind the vertical bars in the capital letter H. In Figure 4-left is schematised a $3x3$ I-view. Obviously these views contain exactly $3x3$ (in general nxm) zz-cells, only if each cell has the necessary neighbours in both primary and secondary dimensions. The formal definition of these views is provided in [22].

Figure 5 shows an example of $5x5$ H-view, where the focus is the zz-cell representing a basketball, highlighted by a bold border; the horizontal dimension is *sports*, and the vertical is *category*. Potentially, in an H-view the same cell could appear in different positions as it may represent the intersection of different semantic contexts (for example, a basketball may be also in the category *games*).

IV. The Mobile Prototype

In this section, we describe the prototype of a new app, called Volo we are willing to develop (Volo comes from latin, and its translation is "I want"). The app proposes a customisable and adaptive (the description of the adaptive aspects of the Volo app lies outside the goals of this paper and is not proposed here) user interface: a user that has very limited capacity to search through images will have an initial configuration containing only pictures of different, frequently used items. An example of an initial view is shown in Figure 6.

The interface shows five items, that slowly scroll up horizontally. Each item looks bigger when it comes to the center, while the others are smaller and opacified. When the user touches an item, it comes to occupy the central position (if it was not already there), the border assume a brilliant color, and the item is added to the history of the user and visualised at the top of the view.
The interface is very simple: in addition to the items, it contains only the home button (top-left corner), and the sequence of selected images. The selection of an item may be associated with a reading voice (customizable in different languages). There is also the option of setting the app so that the reading voice pronounces only the first syllable. This is useful when the child has started to repeat words but is not able to produce it by its own, so the app provides a small hint. To improve user performance the required pictures have to be efficiently accessed, thus objects are viewed following a statistic of most frequently searched items.

Figure 6. Volo app: a simple, initial view: the user selected a tart.

The first screenshot (Figure 6) goes back to Phases 1 to 3 of PECS. More capable users will have a more complex initial configurations. They will start by choosing a subject, then a verb, then an object (Phase 4 of PECS), and so on. Thus, we assume that the app is customizable depending on the different users, and depending on the different learning pathways they are able to follow (some decisions are taken by therapist/parents which have access to the app). Images are also customisable with imported pictures; the app however already provides the user with real daily life pictures.

We also assume that pictures for users in Phases at least 4 will be categorised: for example, to select a particular food the user will first choose the image of food, and then will visualise a list of different types of food. Depending on the user this can be refined by defined subcategories, for example, food-sweets-tart (see the example in Figure 7). While defining categories we are using images which include inside sub-elements of the category itself, for example, to represent food we use an image that contains vegetables, fruit, bread, etc. (see Figure 7 left). This choice derives from the fact that many people with ASD disorders represent concepts through sets of related images; for example, to represent the concept of the color green they may visualise different objects of that color, like grass, leaves, etc. [25].

No matter in which view the user is working, we assume that our app is based on the zz-structure approach. The user is able to visualise few objects in one dimension, i.e., in a row or

 ISBN: 978-1-61208-382-7

Figure 7. An example of division in subcategories: food → sweets → tart.

Figure 8. Volo app: the category food is proposed to the user, and he/she may select it.

in a column at the center of the screen. The user moves from a picture to another one using step scanning, i.e., it presses the switch to move to the next symbol, and stops when it reaches the desired picture, or by swiping the fingers to scroll through images and then pointing the picture.

At this point a new dimension (of the AAC-A) appears in the other direction still at the center of the screen, visualising an H-view (if the previous dimension was horizontal now it appears vertically, or vice versa), and with a subset of items. If the searched item or category is not viewed, the user may move using a circular scanning.

We now illustrate an example for a user that is able to work up to Phase 4 of PECS. The final request is of a tart. The app visualises a page with a row containing the images of sports, drinks, food, games, dvds (see Figure 8), which are general categories.

The user presses the switch or swipes to scroll, and stops when the image of the food is at the center. At this point a new vertical dimension appears and shows the images of sub-categories: sweets and first courses (Figure 9).

The user presses the switch and stops when sweets appears at the center, and a new horizontal dimension appears with images of specific sweets (Figure 10).

Finally, the user presses the switch and stops when the image of a tart appears. In the meantime, in the top the user visualises the sequence of images (zz-cells) that have been selected food-sweets-tart.

What is really new in our model and app is the use of the single zz-structure dimension. All the existing apps work using a grid view, i.e., the screen is filled with rows and columns of items. Using only two dimensions (an horizontal and a vertical

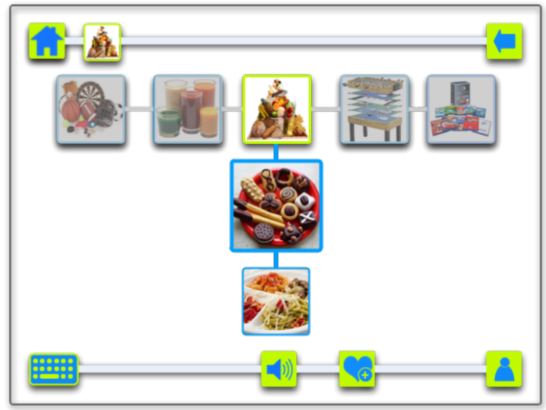


Figure 9. Volo app: the user selected the category "food", and the app proposes its sub-categories (sweets, main courses).

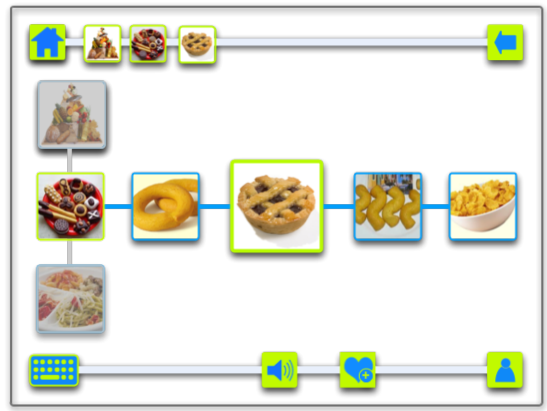


Figure 10. Volo app: the user selected the sub-category "sweets" and the app proposes a set of sweets.

line), limits the number of visible objects thus simplifying the understanding of the user. Moreover, we assume that while the user is navigating on an image (i.e., the image is at the center), it may visualise the other dimension (as a preview). For example, in the previous example depicted in Figure 9 the user navigates on the view, presses and arrives at the category "drink" (thus visualising a dimension with images of drinks), and while realising it is not correct it moves to the food image. This is very important when the child is still not able to categorise because it slowly starts to associate the category image to the sub-items.

We assume that if the user chooses a wrong image there are different solutions: (a) for not very experienced users, we assume the possibility of pressing a complete reset button (i.e., go back to the initial view), represented by the home button (on the top-left corner of Figures 6, 8, 9, 10); (b) more experienced users may also select a back button (represented by the arrow button on the top-right corner in Figures 8, 9, 10), that goes back to the previous view, both for the horizontal and vertical lines, and for the set of selected items in the top.

For children in school age that are starting to read, the app may also provide writings (in the chosen language) in capital letters, so to simplify the reading (if the desired language is not present, the writing is omitted). An example is shown in

 ISBN: 978-1-61208-382-7

Figure 8 where, below the image, the user can read the related English word.

Another feature for expert users is a "like button" (represented by the heart button in Figures 8, 9, 10) with which a preferred item is selected by the user itself. We also assume that some options are customisable by clicking on a special icon (represented by the user button, on the bottom-right corner in Figures 8, 9, 10): the colours, the size of the grid, the initial view, the border of the images, the possibility of adding or removing a writing, adding the sound of a complete word, of a syllable, etc.

In this paper, we chose to illustrate the setting for users with very strong language disabilities. However, our prototype allows to create more complicated views in which the sequence of images that can be chosen are, for example, subject → verb → adjective → object, where the subject can be chosen from a set of images containing the user, parents, brothers, sisters, etc. (assuming the app has been customised with imported images, otherwise the user finds sketched images), a set of images for verbs, adjective and objects.

An additional feature of our app is a reward button, a hand with a thumb up, which is not shown in the figures, but is possible to add it from the setting page. It is well known that positive rewards help the user fixing a concept, as a matter of fact positive reinforcement is the most widely applied principle of behaviour analysis (a commonly used therapy for ASD disorders). Thus, we add the possibility, for the person that is interfacing with the user (or by the user itself when it increases its ability to use the app), to reward the user not only by answering his/her request, but also by pressing this reward button. The app then executes something funny, for example, the app says "well done!" while opening a page where there are stars falling down.

V. Conclusion and future work

In this paper, we presented a formal model and the prototype of the new application Volo we are willing to develop in the coming future.

Volo has some nice features which, to the best of our knowledge, seem to be lacking in other applications. In fact, Volo has been developed having in mind users with very strong language disabilities, and thus its navigation structure is simple and usable, although the underlying model is general and scalable.

As a future work, from a formal point of view, we are preparing a formal and more detailed description for proving that our formal model includes, and can completely describe all the different existing AAC-based applications. From an implementation point of view, we are willing to engineer the app, starting from the actual prototype. Moreover, we are planning to include an expert therapist in the team in order to evaluate both the proposed model and, what is more crucial, to experimentally evaluate the impact on a group of users with ASD. Finally, it would be interesting to propose the real app to teachers at school, and to compare it with other existing approaches also from a usability and accessibility point of view.

Acknowledgment

This work has been partially supported by the PRIN 2010 Project *Security Horizons*.

References

[1] American Psychiatric Association, Ed., The Diagnostic and Statistical Manual of Mental Disorders: DSM 5. BookpointUS, 2013.

[2] S. Dodd, Understanding Autism. Elsevier Australia, 2005.

[3] C. Weitz, M. Dexter, and J. Moore, AAC and children with developmental disabilities., S. Glennen and D. C. DeCoste, Eds. San Diego: Singular Publishing Group, 1997, vol. Chapter 11.

[4] H. R. Goldberg and J. Fenton, Aphonic communication for those with cerebral palsy: Guide for the development and use of a communication board. New York: United Cerebral Palsy of New York State, 1960.

[5] G. C. Vanderheiden and K. Grilley, Nonuocal communication techniques and aids for the severely physically handicapped. Baltimore, MD: University Park Press, 1976.

[6] G. C. Vanderheiden, "A journey through early augmentative communication and computer access," Journal of Rehab R&D, vol. 39, no. 6, May 2009, pp. 39–53.

[7] A. Bondy and L. Frost, "Mands across the water: A report on the application of the picture-exchange communication system in Peru," The Behavior Analyst, vol. 16, 1993, pp. 123–128.

[8] "Pyramid educational consultants, inc." 2014, http://www.pecs-canada.com/ [accessed: 2015-01-08].

[9] A. Szeto, E. Allen, and L. M., "Comparison of speed and accuracy for selected electronic communication devices and input methods." Augmentative and Alternative Communication, vol. 9, 1993, pp. 229–242.

[10] "Alexicom AAC," 2014, http://www.alexicomtech.com/ [accessed: 2015-01-08].

[11] "Voice4u. Helping your communication," 2014, http://voice4uaac.com/ [accessed: 2015-01-08].

[12] "TalpToTalk," 2014, http://www.taptotalk.com/ [accessed: 2015-01-08].

[13] "GoTalk," 2014, http://www.attainmentcompany.com/ [accessed: 2015-01-08].

[14] "Quicktalker family of communication devices," 2014, http://www.ablenetinc.com/Assistive-Technology/Communication/QuickTalker [accessed: 2015-01-08].

[15] "Niki-Talk," 2014, https://play.google.com/store/apps/details?id=it.alessandrolarocca.nikitalk [accessed: 2015-01-08].

[16] "Let's talk by Toriken Media Information Lab," 2014, http://ne-kite.com/index_english.html [accessed: 2015-01-08].

[17] "Proloquo2go," 2014, https://itunes.apple.com/au/app/proloquo2go/id308368164?mt=8\&ign-mpt=uo\%3D4 [accessed: 2015-01-08].

[18] A. Burrington, "Pics aloud," 2014, https://itunes.apple.com/us/app/pics-aloud/id402333607?mt=8 [accessed: 2015-01-08].

[19] "TTalk AAC by T-Box srl," 2014, https://itunes.apple.com/it/app/ttalk-aac/id537958879?mt=8 [accessed: 2015-01-08].

[20] T. H. Nelson, "Welcome to ZigZag," 1999, http://xanadu.com/zigzag/tutorial/ZZwelcome.html [accessed: 2015-01-08].

[21] ——, "A cosmology for a different computer universe: data model mechanism, virtual machine and visualization infrastructure," Journal of Digital Information: Special Issue on Future Visions of Common-Use Hypertext, vol. 5, no. 1, 2004, p. 298.

[22] A. Dattolo and F. L. Luccio, "A formal description of zigzag structures," in Workshop on New Forms of Xanalogical Storage and Function - HT 2009. Turin, Italy, June 29-July 1 2009, pp. 7–11.

[23] A. Moore, J. Goulding, T. Brailsford, and H. Ashman, "Practical applitudes: Case studies of applications," in Proceedings of the 15th ACM Conference on Hypertext and Hypermedia (HT'04). Santa Cruz, California, USA, 9-13 August 2004, pp. 143–152.

[24] M. J. McGuffin, "A graph-theoretic introduction to Ted Nelson's Zzstructures," January 2004, http://http://www.dgp.utoronto.ca/~mjmcguff/research/zigzag/.

[25] T. Grandin, "How does visual thinking work in the mind of a person with autism? A personal account," Philosophical Transactions of the Royal Society, vol. 364, no. 1522, May 2009, pp. 1437–1442.

Experiments and Applications of Support System for Caregivers with Optical Fiber Sensor and Cleaning Robot

Junko Ichikawa and Norihiko Shinomiya

Graduate School of Engineering
Soka University
Tokyo, Japan
Email: shinomi@ieee.org

Tetsuya Kon

Core System Japan Co., Ltd.,
Tokyo, Japan
Email: con@core-system.jp

***Abstract*—This research aims to propose a supporting system with sensor network technology and cleaning robots to alleviate workload of caregivers in welfare facilities. Our focus is to reduce the labor load of nurses that has increased with elderly population growth. On the basis of the system requirements clarified from a site survey in a nursing facility, this paper proposes an integration method of sensors that keeps monitoring situations in a facility and cleaning robots that can approach an incident location detected by the sensing system to confirm safety of facility residents.**

Keywords–monitoring system; sensor network; cleaning robot; nursing.

I. Introduction

Global society has encountered the serious problem of aging nowadays. Especially, the number of over 65 years old Japanese citizens has been monotonously rising since 1950 and accounted for more than 25% of its population in 2013 [1]. The demand for nursing home care has also been increasing in accordance with the growth of senior citizens.

In Japan, the number of facilities that support elderly people is over 25,000 sites, and *group homes* account for about 47% [2]. The group home is a facility where caregivers stay overnight and look after the elderly with senile dementia. In recent situation, each caregiver needs to take care of approximately 10 elderly people during nighttime because of the increase in facility residents. However, this situation is likely to be difficult for caregivers to give enough attention to accidents in a whole facility.

In order to solve the lack of nurses, video monitor systems in group homes have been developed. Caregivers can monitor elderly people behavior through cameras installed in living rooms, bathrooms and entrances 24 hours a day, every day of the year. However, such a monitoring system directly displays and records private behavior with image and sound. Thus, this invasive surveillance would make elderly people stressed [3].

On the other hand, the use of sensor network is conceived as a solution that can prevent such invasive surveillance. Sensor networks can indirectly retrieve environmental information by setting many sensor nodes. Nonetheless, sensor systems can faultily recognize a normal situation as an emergency because they only gather indirect information such as pressure. For instance, something falling on the floor can be regarded as a fall of an elderly person, and this incorrect report confuse caregivers. Thus, another technique is needed to confirm a reported situation by sensor networks.

Use of robots in the care field has attracted much attention. The Ministry of the Economy, Trade and Industry in Japan began *Robotic Care Equipment Development and Introduction Project* that aims at alleviating burden of caregivers in 2013. This project supports development of robotic care equipments for monitoring systems in nursing facilities. According to the Ministry, these systems need a robot which possesses data communication and sensing functions to notify an accident to the care takers [4]. Such a robot is expected to collect more detailed information, which cannot be detected only with wired sensor networks because it can move and conduct interactive safety confirmation process with facility residents.

Hence, this research proposes a Support System for Caregivers (SSC) with optical fiber sensors and a cleaning robot, which can monitor behavior of elderly people and notify caregivers of an emergency without violating their privacy. The goal of our system is confirmation of safety of an elderly person using optical fiber sensors and a consumer-electronics robot instead of image sensors such as a video camera. In addition, the system identifies the place where a facility resident falls down, and informs caregivers of the situation of the resident.

In Section II , some works of daily life monitoring system for the elderly are described. In Section III, some requirements are analyzed based on our hearing researches to caregivers. In addition, the overview of our support system for caregivers and system components are presented in Section IV. Furthermore, in Section V, the prototype system is tested, and we explain experiment results.

II. Related Work

In order to assist the elderly to safely spend their daily life, some monitoring systems have been proposed. For example, a homecare monitoring system is proposed by Bourennane et. al. [5]. The research suggests that the utilization of multi-sensor networks realizes behavior observance of an elderly person, who lives alone at home. Due to the observance, the proposed system provides the alert function in case of dangerous accidents. However, this research does not suppose the use of such a system in welfare facilities where many elderly people inhabit.

 ISBN: 978-1-61208-382-7

TABLE I. Data of visited nursing facility

Residents	100
Occupancy rate	100 %
Caretakers (day)	10
Caretakers (night)	5
Living quarters	two-story
Private rooms (quad)	14
Private rooms (double)	22

Another instance is a monitoring system in group homes [3]. The paper suggests a method to monitor facility residents with videos and install such equipments into the group homes. Nevertheless, the use of cameras is hesitated in terms of privacy problems. Moreover, such a camera system cannot be sited in bathrooms where accidents such as tumble frequently occur.

Both the proposed systems partly solves the problems in care of the elderly; however, a support system that considers the use in facilities where many incidents simultaneously happen and privacy issues is required. Thus, this paper proposed the support system, which is to find an abnormal behavior of residents in welfare facilities and prevent invading privacy of the elderly.

III. Support System for Caregiver

This section summarizes some requirements for SSC that gained from our hearing researches to caregivers. This section also discusses necessary elements for the systems based on the requirements.

A. Assumed Situation

In this research, welfare facilities which fulfills the following conditions is assumed to be the location where our system works. One of conditions is that the facility takes care of senior citizens all day and night. Additionally, a capacity of sickbed is comparatively large.

B. Factual caregivers' burden from site surveys

We interviewed caregivers who have been working at a nursing facility, and table I shows the overview information about the facility. Additionally, this facility has a sensor network system, which involves a pressure sensor and an infrared radiation sensor in the private room to detect the elderly falling down or moving away from the bed.

The survey indicates that welfare facilities using sensor networks face the following problems related to the increase of health care burden of caregivers. One of problems is false detection of sensors, and the other problem is that caregivers have a huge amount of work beside nursing. Whenever a false detection by sensors occurs, caregivers would feel both physical and mental burden because caregivers should confirm safety of residents by visiting a private room each time. Furthermore, we have identified that caregivers have a lot of work other than nursing such as cleaning, managing health information and planning the care. Furthermore, work during night is harder than one during daytime because a smaller number of caregivers are in the facility senior at night. The interviewees stated that they did not have enough time to take after each resident because of their workload.

C. System Requirement

From the above interview, this section clarifies some requirements for our system. First, the function to confirm false detection is needed because the confirmation process forces nurses to spend much extra time and confuses them. Another requirement is a function to reduce workload besides nursing. The interviewees listed cleaning, recording health information and scheduling care plans as their major extra work for them. In those three kinds of works, healthcare information systems developed by Fujitsu [6] and ND software [7] provide solutions to record health information and care planning. Thus, this research focuses on reduction of burden caused by cleaning.

IV. Prototype of SSC

This section presents details of system components, such as sensor device, network management, SSC manager and robot in our prototype system.

A. System Overview

Our system presupposes a situation that an elderly person suddenly would fall down on the floor, and caregivers might not become aware of the accident because they may work at a separated place distant from the point of the incident. Moreover, it is expected to reduce the burden of care such that caregivers have to keep their eye on elderly people all day long and get ready for rushing over to them whenever they need care.

This support system involves sensors embedded in a floor to detect the fall which can be identified by more than two sensors continuously retrieving pressure over certain time interval. Also, a robot with a touch sensor is required to confirm the safety of an elderly person. Figure 1 shows the system overview of support system for caregivers. An agent of Simple Network Management Protocol (SNMP) regularly observes the state of sensors and sends a trap ID to a manager in case of detecting the elderly falling down. The manager constantly retrieves the information and monitors a behavior of an elderly person. When the manager recognizes the fall, it specifies a route to the incident location and sends the commands for a robot to act on the routing information. After the robot moves toward the location and arrives at the goal, it generates a beep to check if the elderly person retains consciousness. The robot sends the state information to the manager according to the touch sensor pushed by the elderly in response to the beep. Then, the manager notifies caregivers about the abnormal situation.

B. Sensor Device

The system uses sensor technologies to detect the accident. Generally, wireless sensor networks are set to observe the motions such as falling [8]. A wireless sensor gathers a lot of information to distribute the sensor nodes among a wide area. However, wireless sensor networks need to supply an electric power to all sensor nodes. Therefore, sensors which do not need to be supplied power are suitable for our system.

One of the prospective sensors is the *hetero-core spliced optical fiber sensors*, and a structure of the sensor is shown

 ISBN: 978-1-61208-382-7

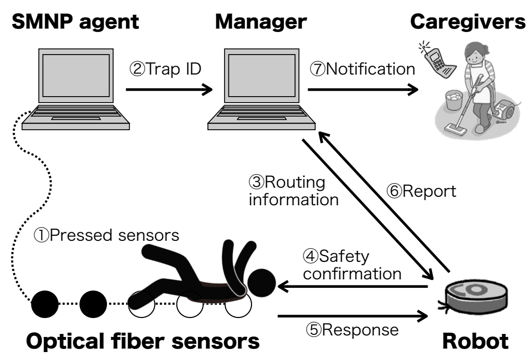


Figure 1. System component diagram.

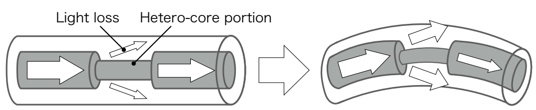


Figure 2. Structure of hetero-core spliced optical fiber sensor.

in Figure 2. It is composed of two single mode transmission fibers. One fiber is cut in two, and the other which has a smaller diameter is inserted between cut fibers. The spliced potion is called *hetero-core portion*. When the hetero-core portion is bent from outside by pressures, light waves leak into the cladding region. The hetero-core portion of an optical fiber sensor can work as a sensor by measuring the light leakage. Furthermore, plural hetero-core sensors are distinguished by changing light leakage of each of the sensors. We carried out an experiment in order to check whether the sensors were correctly distinguished. In this experiment, two sensors are pressed at a minute intervals, and a combination of sensor states is detected. The states were recognized by the difference in loss shown in Figure 3.

C. SSC Manager

An SSC manager fulfills two types of roles: detecting the state of sensors and routing for robot.

1) Detecting the State of Sensors by SNMP: One of the roles of SSC manager is to detect the state of sensors, which is realized by using Optical Sensory Nerve Network (OSN) with hetero-core spliced optical fiber sensors. OSN is a network that realized communication and sensing simultaneously, and it uses SNMP to manage sensor network equipments and distinguishes the sensor states by trap [9][10].

Figure 4 shows the component of sensor network management system. First, the media converter (A) generates light which went through sensors, and the media converter (B) receives the light. Then, an SNMP agent measures the optical loss and issues trap to an SNMP manager. The manager can detect the sensor pressed location to check the trap because the trap previously was set to a threshold voltage per each sensor. The SSC manager regularly analyzes the sensor information provided by an SNMP agent that observes state of each sensor and detects an emergency. If the SSC manager finds the situation, it calculates the route and sends the routing information to the robot.

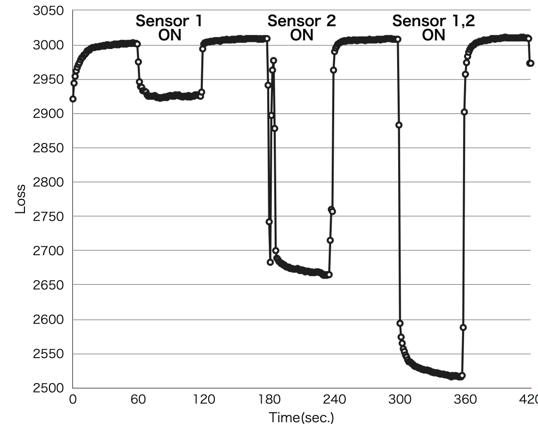


Figure 3. Identification of sensor states.

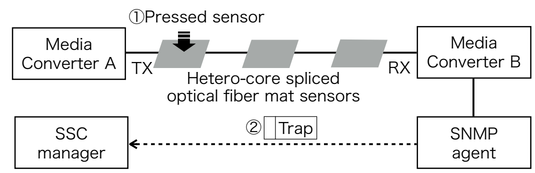


Figure 4. Sensor states detection by SNMP.

2) Routing of Robot: The SSC manager also has a routing function to give a route to a robot. In order to calculate the path, the manager generates a grid graph by Java Universal Network/Graph Framework (JUNG) as shown in Figure 6. This framework is an open source library for Java in order to analyze and visualize the structure of graphs. Furthermore, a shortest path algorithm that has been studied in the field of graph theory is adopted. The algorithm is a solution to obtain a path with the minimum weights from a given source vertex to a destination vertex in a graph consisting of vertices and edges. After the calculation, the manager creates and sends a set of command to robot in order to move it toward the destination.

D. Robot

The system uses a robot to confirm safety for the elderly in order to reduce misinformation of the sensor and to alleviate the labor load of caregivers. The following is the requirements for a robot that is suitable for our system.

First, an appearance of the robot does not offend a user, and any cameras are not attached on the robot. Furthermore, the robot can autonomously move or can be controlled remotely. In order to confirm safety, the robot has beep function and a touch sensor. Finally, it needs to be equipped with vacuum cleaner function.

As a robot which fulfills these requirements, a vacuum cleaning robot, such as Roomba [11] is likely to be selected. Recently, Roomba has penetrated in general households; thus, it is easily configured in our system. Roomba designed simply that does not threaten the elderly. Additionally, it can beep sounds as its alert, and a bumper is installed on ahead of the body, which is used as a button in our system. Besides, though it autonomously moves around rooms in it's cleaning mode, its

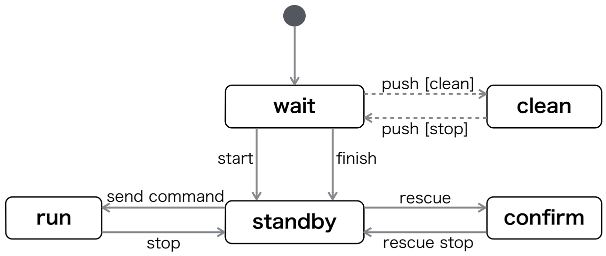


Figure 5. State diagram of Roomba.

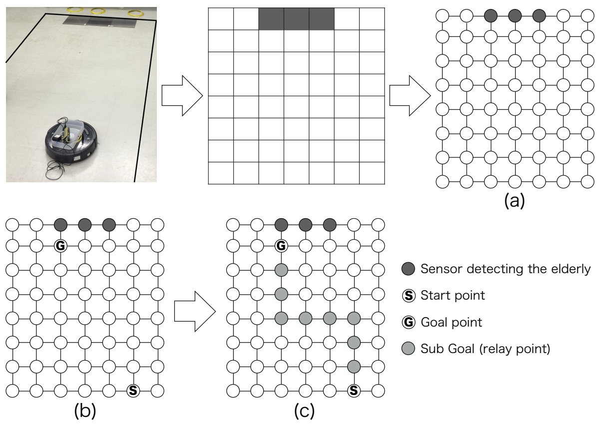


Figure 6. Generated logical map by manager.

movements are able to be controlled by programs from remote locations. Figure 5 shows the state diagram of Roomba in our supporting system.

V. Experiment Results

This section explains results of SSC trial experiment, examining the sensor management and robot control function. At first, points which the robot stops and changes direction on are shown as white node in Figure 6. In the experiment of robot control function, we use 3 mats that sensors are embedded in. When more than two adjacent sensors would respond, the manager could judge the exact place as the accident location which is shown as gray nodes in Figure 6(a). Then, the manager sets both the current place of the robot as a start point and one of the pressed sensors as a goal point as shown in Figure 6(b). Additionally, the manager calculates a path between the start and the end, and Figure 6(c) shows the relay points. Finally, the SSC manager sends the robot a set of commands on a path to a relay point, and the robot moves along the commands as shown in Figure 7.

VI. Conclusion and Future Work

This paper proposes a support system for caregivers, which uses an optical fiber sensor and a cleaning robot. This system aims to monitor the elderly without the use of video recording in order to prevent privacy invasion. The hetero-core fiber sensors and a vacuum cleaning robot enable our system to detect fall of an elderly person accurately. However, the experiment is not considered to lessen the gap between a calculated route and an actual moving route of a robot, which is a remaining issue of this study.

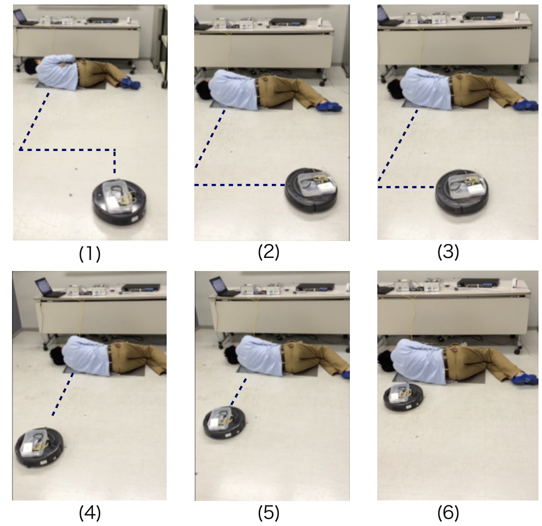


Figure 7. Routing control of Roomba.

References

[1] Cabinet Office, government of Japan. Statistics and analysis of population aging in japan in 2013. [retrieved: August, 2014]. http://www8.cao.go.jp/kourei/whitepaper/w-2014/zenbun/26pdf_index.html

[2] Ministry of Health, Labour and Welfare, government of Japan. Survey of number of elderly supporting facility in 2012. [retrieved: August, 2014]. http://www.mhlw.go.jp/toukei/list/24-22-2.html

[3] T. Sugihara, T. Fujinami, and R. Takatsuka, "An analysis of problems on development and installation of mimamori-care support camera system for persons with dementia in group home," Sociotechnica, vol. 7, 2010, pp. 54–65.

[4] Ministry of the Economy, Trade and Industry of Japan. Revision of the four priority areas to which robot technology is to be introduced in nursing care of the elderly. [retrieved: September, 2014]. http://www.meti.go.jp/english/press/2014/0203_02.html

[5] W. Bourennane, Y. Charlon, F. Bettahar, E. Campo, and D. Esteve, "Homecare monitoring system: A technical proposal for the safety of the elderly experimented in an alzheimer's care unit," IRBM, vol. 34, no. 2, 2013, pp. 92–100.

[6] FUJITSU. Hope wincare-v2. [retrieved: September, 2014]. http://jp.fujitsu.com/solutions/medical/products/archive/wincarev2/

[7] NDSoftware. Nursing care supporting software honobono next. [retrieved: September, 2014]. http://www.ndsoft.jp/next/kinou_kyotaku.html

[8] A. Stefano, A. Marco, C. Paolo, L. Janet, and V. Alessio, Monitoring of Human Movements for Fall Detection and Activities Recognition in Elderly Care Using Wireless Sensor Network: a Survey, 2010.

[9] L. S. Goh, K. Watanabe, and N. Shinomiya, "Feasibility evaluation of multi-point sensing for hetero-core spliced optical fiber sensor using internet-based protocol." IEEE ICST, Dec 2013, pp. 569–572.

[10] N. Abe, Y. Hanada, N. Shinomiya, and Y. Teshigawara, "A demonstration of sensing and data communication in optical sensory nerve network with hetero-core fiber optics sensors," IEICE Technical Report. USN, vol. 107, no. 294, Oct 2007, pp. 67–71. http://ci.nii.ac.jp/naid/110006453275/

[11] iRobot. Roomba. [retrieved: June, 2014]. http://www.irobot-jp.com

 ISBN: 978-1-61208-382-7

Evaluation of a Vibrotactile Device For Outdoor and Public Transport Pedestrian Navigation Using Virtual Reality

Olivier Hugues and Philippe Fuchs
MINES ParisTech, PSL Research University
Center For Robotics
Paris, France
firstname.lastname@mines-paristech.fr

Lucie Brunet and Christine Megard
CEA-LIST
Sensory and Ambient Interfaces Laboratory
Fontenay-aux-Roses, France
firstname.name@cea.fr

***Abstract*—It can be difficult to find your way in public transport, especially when the journey combines indoor and outdoor transportation. We designed an innovative vibrotactile device dedicated to guide a pedestrian in public transport. This multi-modal interface can be used to guide a pedestrian in unknown public transport. The device can be used by visually impaired person. The device has been tested during two main phases. The first step was to test the device using virtual reality while the second step test was to test the device in a real environment. This paper presents the first part of the evaluation of the device. We have developed a virtual reality scenario to assess the objective and subjective utility of the device. The results showed that the device could properly guide users. We also evaluated the usefulness of a warning vibration preceding a message. It was found that the vibration seems to introduce confusion to the pattern recognition by the user.**

Keywords–Vibrotactile device evaluation;multi-modal interface; tourism mobile device; public transport; virtual reality.

I. INTRODUCTION

In a large city, the range of public transport servicies is wide –buses, subways, trams, bicycles, electric cars– making the transportation network more and more complex. To address this complexity and facilitate the movement of users, information systems are available in many guises such as signs, information boards or mobile device applications. However, existing systems to assist pedestrians mostly offer visual cues, sometimes combined with sound. The need to inform the pedestrian can lead to saturation of these sensory modalities, making it difficult to grasp information during the journey. The problem is even more serious if the user is not familiar with the transport system, or if it's his first time navigating. These observations led us to consider the development of a new way to interact with the navigation aid systems. There are two objectives in this paper. The first one is to evaluate, using virtual reality, the potential interest of a preliminary alert to draw the user's attention. The second one is to evaluate a vibrating wristband to help pedestrians to navigate the city and public transport.

In Section II, we present related work on devices used to guide pedestrians using multi-modality. In Section III, we describe the device and give some details about the design of the hardware and the software. Then, in Section IV, we present the user study which was conducted using virtual reality. Finally, we present the results and discussions in Section V.

II. RELATED WORK

Pedestrian navigation aid systems mainly employ visual and auditory modality [1][2][3][4]. Indeed, cognitive resources for visual and auditory modality are already heavily used in mobile environments [5]. The solicitation of visual modalities when interacting with navigation aid systems strongly degrades the performance of pedestrian mobility [6]. In this case, the pedestrian must perform several tasks simultaneously. He must look at the screen of his mobile interface and pay attention to the environment at the same time. The auditory modality, slightly less used, is not the best communication channel in public transport. In a noisy acoustic environment saturated with visual information, the haptic modality seems to be more promising and worth exploring to provide information to travelers. The haptic perception is defined as a perceptual system composed of two subsystems: tactile –cutaneous channel– and proprioceptive –kinesthetic channel [7][8]. We believe that the use of the haptic channel is useful for navigational interfaces because it does not overlap with the main channels (auditory and visual) used during the journey. This is based on the theory of Wickens [9], which highlights the existence of different resource reservoirs, each associated with a particular processing channel. The processing channels are independent. In other words, when two different tasks are performed at the same time by different processing channels, the model predicts no performance degradation. In addition,Lee et al. [10] also shows that a vibration can automatically draw the driver's attention to information delivered by the system. Another experiment confirmed this hypothesis, indicating that a vibratory stimulation serves to focus the driver's attention on the driving situation in order to pay attention to potential risks. For example, a study was conducted consisting of sending vibration signals to the conductor's waist [11] or through the seat [12]. The haptic modality was shown relevant to pay attention to information about the environment without

 ISBN: 978-1-61208-382-7

disrupting the visual and auditory channels. This validates the choice to create devices for pedestrian navigation employing the haptic modality. It is important to keep in mind that the user is continually confronted with multiple sensory stimuli. Some of them are useful while others are irrelevant. The users' perception has to filter information in order to limit the number of informations to be processed [13]. This system may be faulty when the user faces an unexpected situation. This can be caused, for example, by performing recreational activities not related to the principal activity of transport such as reading/writing an SMS. In this case, there is a shift in focus from relevant information to the secondary activity [14]. However, we know that the appearance of an event is similar to a distraction, induces an automatic attention redirection to this event. It was in 1980 that Posner [15] described this loss of concentration as "exogenous focus of attention." Haptic stimulation, used as a distraction, could therefore help to focus the user's attention back to commute.

Issues regarding the displacement in virtual worlds have been widely studied [16] and especially for large virtual environments [17][18] to avoid the cyber-sickness [19]. Slater [20] showed that the sense of immersion is impacted by the metaphor of displacement and a system allowing the user to walk under degraded conditions is better than a classical interface with buttons and joysticks. This is why we choose a particular metaphor to help the user in his displacement and given below are some details about the design of this user experience.

III. Design Details

The activities extracted from the study conducted by Brunet et al. [21] allowed us to select 8 functions to assist the user. According to this study, there are two main functions: one for guidance and one for warning the user. These two functions are represented by two different devices shown on Figure 1. Guiding is provided by D_1, a hand-held device composed of a body (in white) and a small movable part (in black). This part can be tilted in 8 directions (cardinal and diagonal). It is used to indicate the direction by putting a finger on the tilting part. For reassurance and warning functions, the second part, D_2, is worn around the wrist and is composed of 8 vibrating units. This setup allows creating specific vibration patterns for each information and alert. Changing the following settings creates different patterns of vibration:

- Vibration time
- Vibrator sequence
- Delay between each pulse

A. *Wayfinding in virtual environment*

The commute consists of following a path through nodes and segments. The participant must move from one node to another. This is done to simplify the interaction needed for the displacement. When he reaches a node, the movement automatically stops. Then, he can move his head to choose an orientation he wants to take. Once a direction has been selected, the participant must step forward (beyond a mark on the ground). His foot must remain in front of the mark until he reaches the next node. Between two nodes, the participant can stop moving forward by repositioning his foot behind the mark on the ground. When he decides to move again, he simply puts his foot forward and movement resumes.

The general concept of this device is based on the differences between each haptic pattern. The interaction mainly consists of seven vibrating messages delivered to the user through the bracelet. Each vibration pattern is associated with an image, which is an activity related to the user's commute (see Table I). Please find details about the signal used for each pattern in the study of Brunet et al. [22].

Figure 1. Illustration of the device. Right: Hand-held device (D_1) used to indicate the direction by tilting the small black part. Left: Worn around the wrist (D_2) composed by 8 vibrators.

TABLE I. VIBRATIONAL MESSAGES ASSOCIATED WITH THEIR IMAGES AND RELATED CONCEPTS.

Name	**Sign**	**Concept**
Knock-knock		Vehicle Alert: you need to get in/out from the vehicle.
Siren		Incident Alert: incident on your route.
Bug		Unavailability Alert: technical issue (escalator...).
Half-lap		Wrong road alert: you're going the wrong way.
Heart		Point of interest alert: you are next to one of your POI.
Waltz		Information Alert: you're next to an information center.
It's Ok		Reassurance Alert: you're on your way, no problem.

IV. User Study

The use of Virtual Reality (VR) to realize the evaluation of our prototype will help us to shape the next user study that will take place in a real environment. VR allows us to perform device evaluation faster than in a real environment and with better control of many constraints [23]. Furthermore, VR helps us to make few design choices and minimize the cost of producing devices that would neither be used nor accepted by users [24]. In addition, VR allows the control of parameters that we could not manage in the real world such as the noise level of the environment to assess the impact of this parameter on the user's activity.

A. Hypothesis

The first hypothesis (H1) we wish to verify is that the presence of a preliminary vibrational message before the vibration itself will improve pattern detection during the commute in the virtual environment. This hypothesis is focused on D_2. It postulates that the presence of this message should reduce the number of misinterpretations. The second hypothesis (H2) consists of testing if D_1 is well designed for guiding participants. This will be confirmed by obtaining a minimum number of misorientations.

The system is designed for a wide population range so we want to verify if the devices D_1 and D_2 are easy to use for a large portion of the population. We also want to compare results among different age groups. Finally, the third hypothesis (H3) consists in verifying, by questionnaires, if the device improves the user experience. In this study, participants experience immersion in a virtual environment representing a metro station and its external surroundings. The user's dominant hand holds the device D_1. Around his dominant wrist, the user wears the device D_2 to receive vibration alerts and reassurance. The study was conducted in four steps. The first one (E1) is dedicated to learn how to employ devices D_1 and D_2. The second one (E2) is needed to learn how to move through the virtual environment with D_1 and D_2. The third (E3) is the user study itself and the fourth (E4) is assigned to filling out a questionnaire. We will describe each step in the following sections.

B. Tasks

Step E1 is dedicated to familiarizing the participant with devices D_1 and D_2. The user starts by learning 5 vibration patterns: The experimenter says the name of a pattern, sends it to the device (worn by the participant) and then says the name again. This procedure is repeated for each of the five different vibration patterns randomly. The procedure is repeated a second time without saying the name. We then move to the stage of verbalization by the participant itself. Each vibrational pattern is sent to the device and the participant must identify the pattern by its name. This phase is repeated as long as the participant makes errors. The participant learns how to move in the virtual environment in step E2. In this preliminary step, the device D_2 is worn by the participant for practical reasons, but is inactive. Moreover, the participant does not have access to the device D_1 during the first few minutes. The participant moves along a predetermined path (the same for all subjects) and is guided by the experimenter who tells him the directions to take. When the subject comes to a particular node (same for all participants), the experimenter allows the participant to use device D_1. At this step, the participant can move without guidance from the experimenter, but helped by D_1. During the movement, vibration patterns are sent to the participant through D_2. For each pattern, the participant must tell the experimenter the name of the pattern recognized. The experimenter valids the name of the pattern and repeats the vibration in case of error. After several nodes, the participant reaches the end of the learning path.

The next step (E3) is the user study itself. A scenario is proposed to the participant before he starts. The scenario begins on the platform. He must go downtown, to a street next to the subway entrance in order to go shopping. He must then join a friend next to a subway entrance to visit a museum. To do this, the participant and his "virtual friend" should take the metro. During the scenario, the participant is only guided by D_1 and D_2. Prior to joining his friend, the participant is asked to send an SMS to agree on the meeting place. The message to be sent is the same for all participants. The experimenter gives the participant a mobile phone at that time. While writing the SMS, D_2 vibrates. This operation is used to assess the ability of the device D_2 to be understood, even if the user is doing many different tasks at the same time –commuting, writing SMS–. During the commute to the subway, an incident on the line is announced (by D_2) and it is recommended to take another line by making a U-turn. Once back on the platform, the experiment ends. During the scenario, cultural and commercial points of interest or alerts are presented to the participant through D_2. During the experiment, the software logs steering errors. An error is recognized whenever the subject wishes to leave a node towards another node in a direction which is inconsistent with that indicated by D_1. The experimenter also records misunderstood vibration alerts.

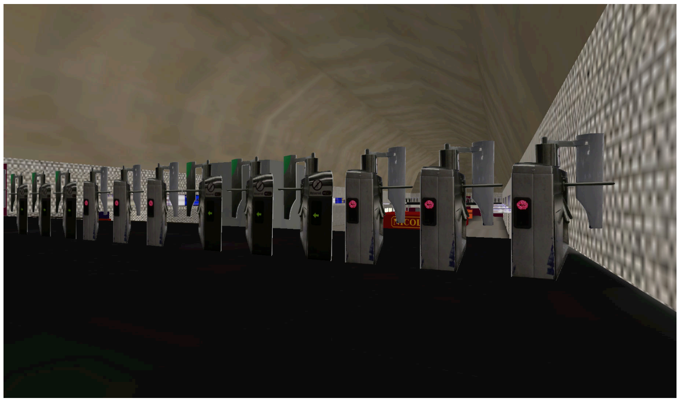
C. Implementation

A 3D model of the metro station was implemented in Unity3D (Figure 2). The station had to be large enough so that the path can be complex so we modeled one of the largest Paris metro station. The virtual environment was rendered on a back-projected wall $(3.1m \times 1.7m)$ with monoscopic rendering. We use of ART cameras for motion tracking: both the head and the dominant foot of the user were tracked in real time thanks to passive markers. The two devices were connected to a computer using Bluetooth. Keyboard control was available for the experimenter to record errors.

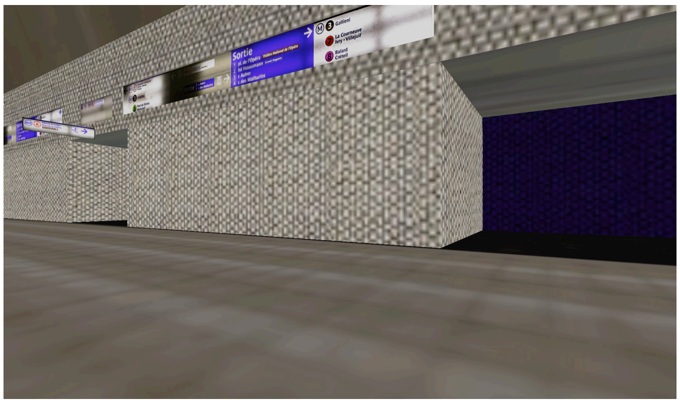
V. Results and Discussions

A total of 21 subjects participated in this study. The duration of the experiment for each subject was about 1 hour. We chose to separate subjects randomly into two groups: during the experiment, the first group (G1) received a vibrational warning before the vibration itself, whereas the second group (G2) received the vibrational messages without this vibration

 ISBN: 978-1-61208-382-7



(a) A view of the metro station 3D model.



(b) The signs are modeled in the 3D model.

Figure 2. An example of the virtual metro station.

warning. The age distribution of our population was 37.5 years on average with a median of 32.5 years. The gender distribution was balanced with 11 men and 10 women. We will discuss the results related to data acquisition in the next section and we will deal with the subjective experience in the following one.

A. User Study Measures

For both groups G1 and G2 (the entire population), we observed a tolerable error rate of 5.39% for D_1, while the error rate was relatively large for D_2 with 28.95%. We performed 1-way ANOVAs to detect any significant effects of the vibration alert. There was no significant influence on the user's errors on D_1 and whether or not they received the preliminary message via D_2 (F (20,1) = 0.29, p = 0.59). D_1 allowed to properly guide participants in the virtual environment, and was not affected by the different modes of D_2, which validates H2.

We found a significant influence of the preliminary message on the vibration pattern recognition (F (20,1)=3.22, p=0.09). The G1 group experienced more difficulties recognizing the vibration pattern with an error rate of 40.47% than the G2 group with an error rate of 15.39%. Many participants made a confusion between the preliminary alert and the message itself and few users tried to recognize a pattern when the preliminary alert was started and not when the message itself started. These error rates were quite large, so we have suggestions for possible improvements. We could increase the amount of time to learn how to use D_2 or simply reduce the number of messages required (5 in this study). In addition, another clue about this high error rate was the device itself. During the experiments, the vibrating motors of the prototype occasionally lost contact with the skin of the participant due to his movements.

Analysis relating to age has shown interesting results. We decided to divide our population into two groups. The separation was the average age (37.5 years old). A significant difference (F(20.1)=7.41, p=0.01) was observed for D_1 error rates. On average with D_1, participants under 37.5 years old (12 people) made 1 error, whereas participants over 37.5 years old (8 people) made 3.4 errors. The older age group experienced more difficulties in recognizing the vibrational patterns of D_2 (F (20,1)=3.35, p=0.08). Age seems to be a significant factor for the perception of the vibration of D_2.

We will now look at the particular case of the vibration recognition when the participant was asked to write an SMS during his commute. We saw a significant impact of the preliminary vibration when writing an SMS with an error rate of 18% for G2 and an error rate of 50% for G1 (F (20,1)=5.35, p=0.03). This result confirms the previous results that showed a greater error rate for G1 than for G2.

B. Subjective user experience

To evaluate the subjective user experience it was necessary to collect subjective data reflecting the experience felt by each participant during interaction with the prototype. The subjective aspect of the user experience takes into account the emotion felt during the session. According to the work in this field, we use some classic usability tests to evaluate the user experience starting with an evaluation of presence and immersion. All participants were asked to answer the questionnaire "presence and immersion" from [25]. Participants could give a score between 1 and 7 (the higher the better) on the control of events, system responsiveness, the naturalness of interaction, visual appearance, consistency of movement and involvement. Results are shown in Figure 3.

All responses were above average. We have not noticed any particular problems during the experiment, such as simulator sickness potentially caused by the commute except for one person prone to vertigo at the top of the virtual stairs.

We can observe a significant difference (Student's t-test) between the two groups for the first question regarding the level of control (Figure 4). We see that the older group felt less control over the system than the younger group.

C. Overall user experience

Mood can be experienced directly or at a reflexive level. Mood can be organized into two main dimensions: pleasant/unpleasant and calm/excitement. To extract the mood experienced, all participants were asked to fill out the Brief Mood Introspection Scale (BMIS) [26] and the SAM scale [27] at the

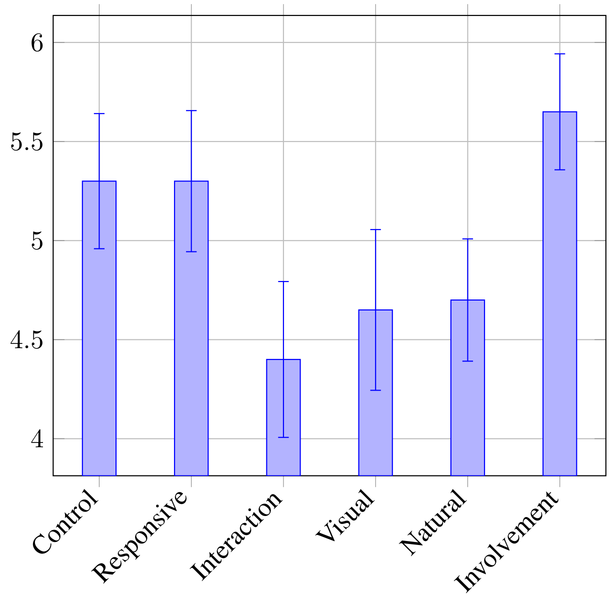


Figure 3. Response to the questionnaire on presence and immersion for the entire population.

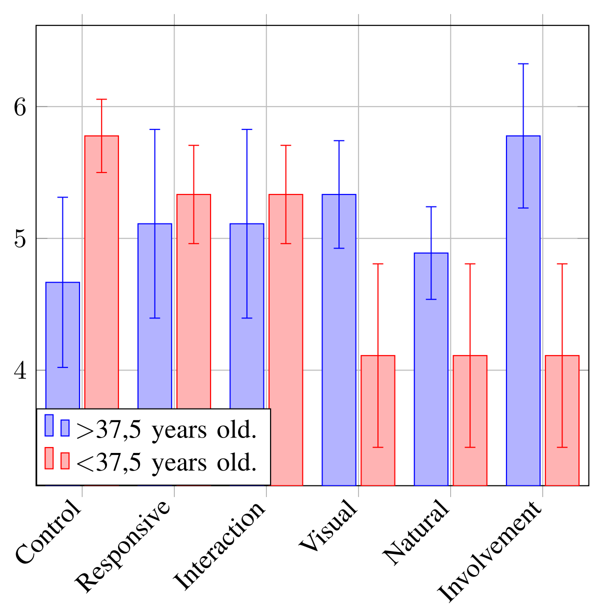


Figure 4. The answer averages to the question of presence and immersion for both age groups.

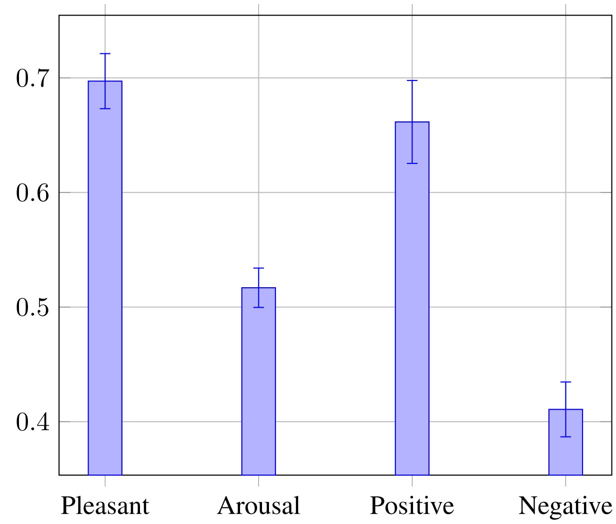


Figure 5. The BMIS mood scale in average for the entire population.

end of the user study. The well-known BMIS scale consists of 16 adjectives. Two adjectives are selected for each of the eight emotional states. The adjectives are: Lively, drowsy, happy, grouchy, sad, peppy, tired, nervous, caring, calm, content, loving, gloomy, fed up, jittery and active. Participants had indicate how well each adjective describe their mood by choosing among the different sentences:

- definitely do not feel;
- do not feel;
- slightly feel;
- definitely feel.

The BMIS questionnaire allowed us to assess the emotion felt during the interaction. Positive mood was observed for all participants with a pleasant and positive experience with a lower level of excitement for negative sentiment (Figure 5). Figure 6 shows the result averages of the BMIS questionnaire for both age group.

Emotions can be described in terms of three independent dimensions: pleasure/displeasure, degree of arousal, domination/submission (PAD model) [28]. These three elements are independent and may occur without impacting each other. The SAM questionnaire [27] is an instrument to measure emotional states based on pictures to achieve a self-evaluation of an object or event based on the three main emotional dimensions. SAM provides a list of pictures for each dimension of the PAD model associated to a scale from 1 to 9. It has the advantage of being filled out very quickly, hence there are no mistranslation issues and both children and adults can fill it out. The SAM scale allows us to extract a general emotional state of the participants. Figure 7 shows the score averages for all the participants. It can be seen that participants had fun during the experiment. Younger participants felt more pleasure than the older group, and they also felt a stronger sense of dominance compared to the older participants (Figure 8). The feeling of excitement is relatively low for both groups, which confirms

 ISBN: 978-1-61208-382-7

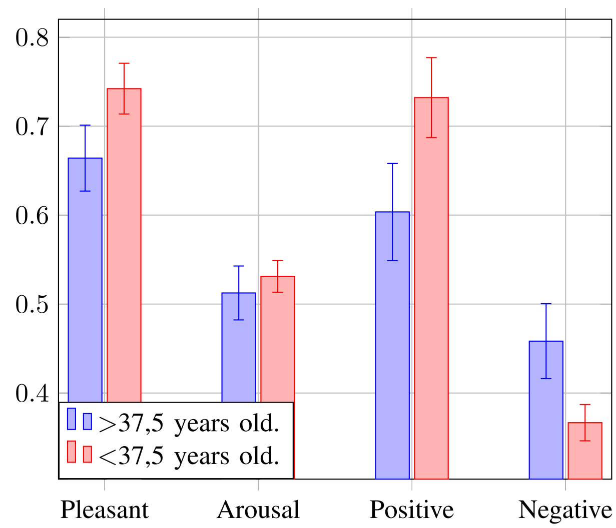


Figure 6. The score averages of the BMIS questionnaire for each group.

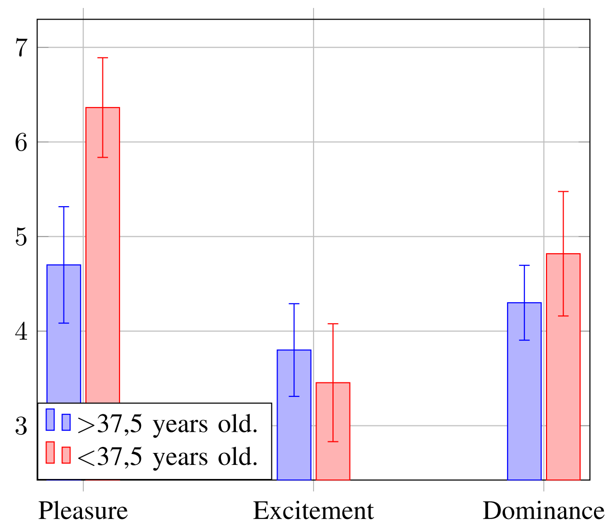


Figure 8. The score averages of the SAM scale for each age group.

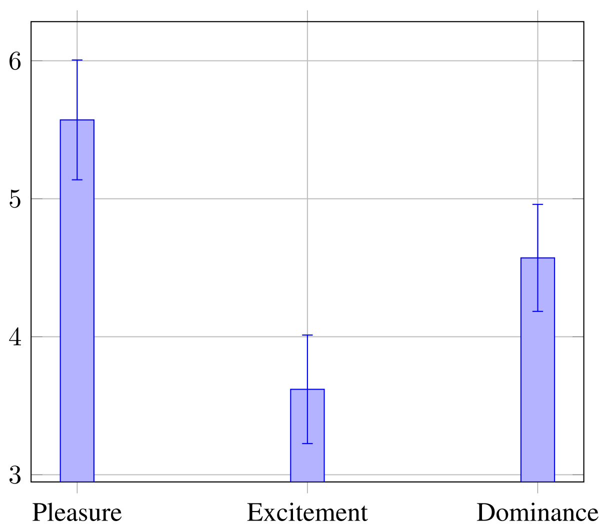


Figure 7. The SAM scale score averages for the entire population.

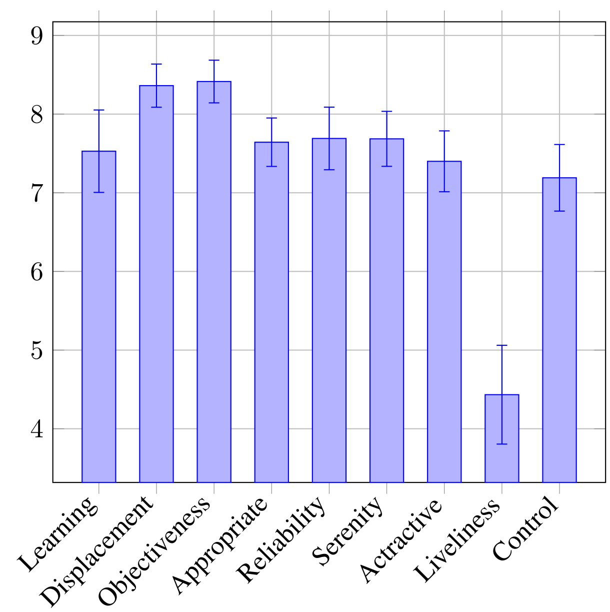


Figure 9. The score averages for all the participants regarding their feelings about interaction with D_1.

that the participants were relaxed during interactions with the device.

We will now discuss the user experience for both devices D_1 and D_2 separately. Using a Likert scale, we will be able to evaluate the usability of each device.

D. User experience with D_1

Concerning D_1, it can be noted that the score is never below 7 out of 10 except for the liveliness (Figure 9). The subjects perceived a slow interaction with D_1.

Table II summarizes the responses to open-ended questions about D_1. Four participants thought that D_1 was simple and an intuitive guide during the commute and five participants felt that the system properly indicated the direction. However, five respondents indicated that they preferred to find their way using a map or a GPS-based application.

E. User experience with D_2

Use of D_2 is overall lower than the score of D_1 as shown by Figure 10 especially for the training phase. Subjects were able to adapt to the device and seemed not to have trouble memorizing the patterns. Table III summarizes the responses to open-ended questions about D_2. Three participants thought that device D_2 provided good vibration recognition and two

 ISBN: 978-1-61208-382-7

TABLE II. SUMMARY OF COMMENTS REGARDING THE USE OF D_1.

Positive points - D_1	**Number of people**
Simple and intuitive guide	4
Properly indicates the direction	5
Usefulness in open space without signs	2
Useful if I'm in a hurry	1
Useful for a blind or visually impaired person	4
No opinion	5
Negative points - D_1	**Number of people**
It would be easier if the button were still ahead when we have to go forward	1
Cumbersome	3
I prefer to find my way using a map/GPS	5
Lack of independence	1
Smartphone is sufficient	1
Using signs is easier	1
Requires concentration	2
Lack of autonomy	1
Lack of direction update	3
No opinion	5

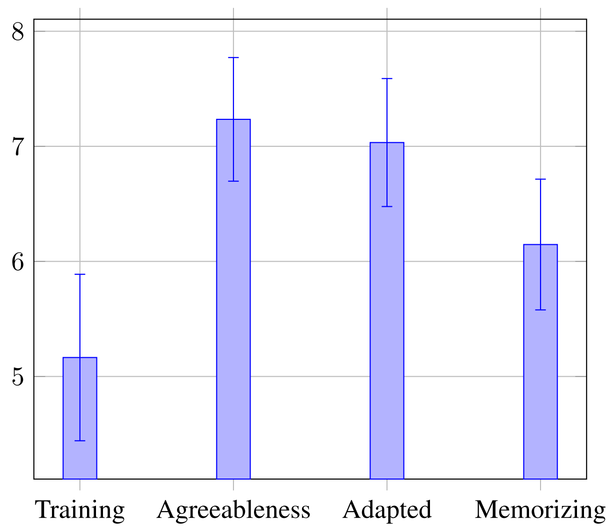


Figure 10. The score in average for all the participants regarding their feelings about the interaction with D_2.

TABLE III. SUMMARY OF COMMENTS REGARDING THE USE OF D_2.

Positive points - D_2	**G1 (10)**	**G2 (10)**	**Number (20)**
Good vibration recognition	2	1	3
Provides vital information	1		1
Not encumbered hands	1		1
Fun aspect	1	1	2
Discrete interaction	1	1	2
No opinion	12		12
Negative points - D_2	**G1 (10)**	**G2 (10)**	**Number (20)**
Cognitive load	3	3	6
Lack of discrimination	2	4	6
Smartphone is sufficient	1		1
Prefer to use eyes and ears	1		1
Memorization effort	1	4	5
Lack of control	1		1
Vibrations disturbance		2	2
No opinion	3	2	5

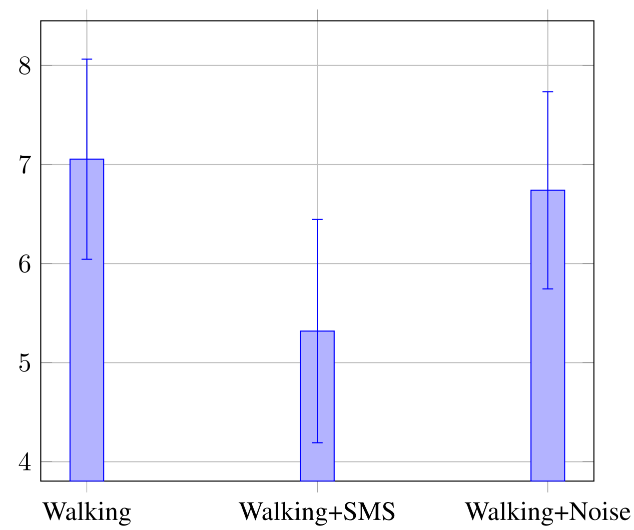


Figure 11. Feeling for G1 on the impact of the preliminary alert during different commuting conditions.

participants have highlighted the fun aspect of D_2. However, six participants have noted difficulties to distinguish patterns (lack of discrimination) and five participants have found important memorization effort.

F. Impact of the alert

Based on answers to the question (does the jingle allow to anticipate the identification of the vibration?), participants said yes up to 6.9/10. Depending on the different conditions during the commute (walking, walking and writing an SMS, walking in a noisy environment), some participants seemed to be distracted by the SMS (Figure 11) and indicated that they had experienced difficulties related to the recognition of vibration. This is confirmed by the observed error rate of 50% for the pattern recognition.

We can note that the alert seems not to be useful while writing a text message for G1, the group which received the alert (Figure 11). This has been confirmed by observations with a significant recognition error rate during this phase (more than 60%). We can observe (Figure 12) that the alert was well received for the younger group during all three phases (walking, walking + SMS, walking + noise) unlike the other group where the alert was rather negative while writing an SMS. Table IV summarizes the responses to open-ended questions about the alert. Five participants thought that the alert helped them to listen to the message but two participants said that the presence of the alert was annoying and caused

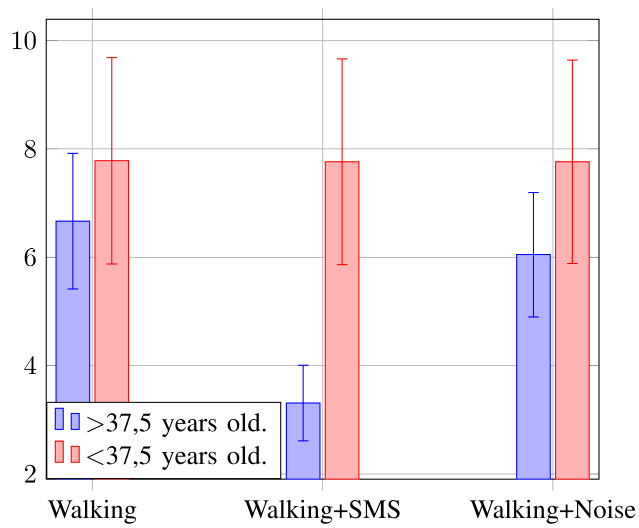


Figure 12. Feeling averages for each age group regarding the vibration pattern recognition.

confusion with the message itself.

TABLE IV. SUMMARY OF COMMENTS FORM PARTICIPANTS IN GROUP G1 REGARDING THE ALERT.

Positive points - Alert	**G1 (10)**
Helps prepare to "listening" the message - promotes wakefulness - caution - concentration	5
No opinion	4
Negative points - Alert	**G1 (10)**
Not essential	2
Presence of alert is annoying and causes confusion with the message itself	2
No opinion	3

VI. Limitations and Future Work

As we said earlier, the used device was a prototype and still has many possibilities of evolution. Indeed, the bracelet could rarely not have been in perfect contact with the skin of the user, which could have lead to a loss of information.

Following this study using virtual reality, we conducted a study in a real environment. We designed the test protocol for the real environment following the results of the study presented here, and we have, for example, specifically emphasized the appearance of the jingle. Results of the study conducted in real environment will be presented in another paper.

Today, we plan to target a population with visual impairments to confirm the interest of this device for this population.

VII. Conclusion

This paper focused on the evaluation of a vibrotactile device for outdoor and public transport pedestrian navigation using virtual reality. A user case was implemented in which a pedestrian had to commute in a large virtual metro station. In this study, we evaluated each of the four proposed hypotheses. We note that our first hypothesis is not validated because the preliminary alert seems to bring confusion to pattern recognition with lower performance for participants compared with those who did not receive the preliminary message. This result, however, can be addressed due to the fact that the device used is only a prototype. The second hypothesis was to assess the guide performance of the device and the data shows that the system is useful to guide users through the station. We noticed a significant difference concerning the age of users. People in the younger age group generally report that it is easier for them to recognize a vibration. Finally, the analysis of questionnaires allows us to conclude that the user experience is quite positive. However, it is difficult to decouple the impact of the experience in the virtual environment with the use of the device itself, which is why we will conduct another user study in the real environment of the metro station modeled in this paper.

VIII. Acknowledgments

We would like to thank all the subjects that participated in the user study. We also want to thank the National Research Association who is founding this work and the RATP Group (Régie Autonome des Transports Parisiens).

References

[1] Y. Miyazaki and T. Kamiya, "Pedestrian navigation system for mobile phones using panoramic landscape images," in Applications and the Internet, 2006. SAINT 2006. International Symposium on, 2006, pp. 7 pp.–108.

[2] S. Kaiser, M. Khider, and P. Robertson, "A pedestrian navigation system using a map-based angular motion model for indoor and outdoor environments," Journal of Location Based Services, vol. 7, no. 1, 2013, pp. 44–63.

[3] H. Furukawa and Y. Nakamura, "A pedestrian navigation method for user, save and easy wayfinding," in Human-Computer Interaction. Users and Contexts of Use, ser. Lecture Notes in Computer Science, M. Kurosu, Ed. Springer Berlin Heidelberg, 2013, vol. 8006, pp. 156–165.

[4] M. Kluge and H. Asche, "Validating a smartphone-based pedestrian navigation system prototype: an informal eye-tracking pilot test," in Proceedings of the 12th international conference on Computational Science and Its Applications - Volume Part II, ser. ICCSA'12. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 386–396.

[5] A. Oulasvirta, S. Tamminen, V. Roto, and J. Kuorelahti, "Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile hci," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ser. CHI '05. New York, NY, USA: ACM, 2005, pp. 919–928. [Online]. Available: http://doi.acm.org/10.1145/1054972.1055101

[6] M. Pielot and S. Boll, "Tactile wayfinder: Comparison of tactile waypoint navigation with commercial pedestrian navigation systems," in Pervasive Computing, ser. Lecture Notes in Computer Science, P. Floréen, A. Krüger, and M. Spasojevic, Eds. Springer Berlin Heidelberg, 2010, vol. 6030, pp. 76–93. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-12654-3-5

[7] K. E. Maclean, "Haptics in the Wild: Interaction Design for Everyday Interfaces," In Carswell, M. (Eds.), Review of Human Factors and Ergonomics (HFES). Santa Monica: United States, 2008. [Online]. Available: http://www.cs.ubc.ca/labs/spin/publications/spin/maclean2008hfes.pdf

[8] S. Lederman and R. Klatzky, "Haptic perception: A tutorial," Attention, Perception, and Psychophysics, vol. 71, no. 7, 2009, pp. 1439–1459. [Online]. Available: http://dx.doi.org/10.3758/APP.71.7.1439

[9] C. D. Wickens, "Processing resources in attention, dual task performance, and workload assessment." DTIC Document, Tech. Rep., 1981.

[10] J. D. Lee, J. D. Hoffman, and E. Hayes, "Collision warning design to mitigate driver distraction," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ser. CHI '04. New York, NY, USA: ACM, 2004, pp. 65–72. [Online]. Available: http://doi.acm.org/10.1145/985692.985701

[11] C. Ho, N. Reed, and C. Spence, "Assessing the effectiveness of "intuitive" vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator," Accident Analysis & Prevention, vol. 38, no. 5, 2006, pp. 988–996.

[12] J. B. V. Erp, H. A. V. Veen, C. Jansen, and T. Dobbins, "Waypoint navigation with a vibrotactile waist belt," ACM Transactions on Applied Perception (TAP), vol. 2, no. 2, 2005, pp. 106–117.

[13] D. E. Broadbent, "The role of auditory localization in attention and memory span." Journal of experimental psychology, vol. 47, no. 3, 1954, p. 191.

[14] C. Lemercier and J.-M. Cellier, "Les défauts de l'attention en conduite automobile: inattention, distraction et interférence. [attention deficits in car driving: Inattention, distraction and interference.]," Le travail humain, vol. 71, no. 3, 2008, pp. 271–296.

[15] M. I. Posner, C. R. Snyder, and B. J. Davidson, "Attention and the detection of signals." Journal of experimental psychology: General, vol. 109, no. 2, 1980, p. 160.

[16] K. M. Stanney, R. R. Mourant, and R. S. Kennedy, "Human factors issues in virtual environments: A review of the literature," Presence: Teleoper. Virtual Environ., vol. 7, no. 4, Aug. 1998, pp. 327–351. [Online]. Available: http://dx.doi.org/10.1162/105474698565767

[17] R. Darken, T. Allard, and L. Achille, "Spatial orientation and wayfinding in large-scale virtual spaces: Guest editors' introduction," Presence: Teleoperators and Virtual Environments, vol. 8, no. 6, 1999, pp. 3–6.

[18] J. L. Chen and K. M. Stanney, "A theoretical model of wayfinding in virtual environments: Proposed strategies for navigational aiding," Presence: Teleoper. Virtual Environ., vol. 8, December 1999, pp. 671–685.

[19] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks Jr, "Walking¿ walking-in-place¿ flying, in virtual environments," in Siggraph, vol. 99, 1999, pp. 359–364.

[20] M. Slater, M. Usoh, and A. Steed, "Taking steps: The influence of a walking technique on presence in virtual reality," ACM Trans. Comput.-Hum. Interact., vol. 2, no. 3, Sep. 1995, pp. 201–219. [Online]. Available: http://doi.acm.org/10.1145/210079.210084

[21] L. Brunet, C. Megard, S. Paneels, G. Changeon, J. Lozada, M. P. Daniel, and F. Darses, """invitation to the voyage": The design of tactile metaphors to fulfill occasional travelers' needs in transportation networks," in World Haptics Conference (WHC), 2013. IEEE, 2013, pp. 259–264.

[22] L. Brunet, C. Megard, S. Paneels, G. Changeon, J. Lozada, M. Daniel, and F. Darses, "Invitation to the voyage: The design of tactile metaphors to fulfill occasional travelers' needs in transportation networks," in World Haptics Conference (WHC), 2013, April 2013, pp. 259–264.

[23] P. Fuchs and G. Moreau, "Le Traité de la Réalité Virtuelle," Presse de l'Ecole des Mines de Paris, Troisième Edition. Mars 2006.

[24] J.-M. Burkhardt, Réalité virtuelle et ergonomie : quelques apports réciproques, PUF, Le travail humain ed., ser. 1, DOI : 10.3917/th.661.0065, 2003, vol. 66.

[25] G. Robillard, S. Bouchard, T. Fournier, and P. Renaud, "Anxiety and presence during vr immersion: A comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games," Cyberpsy., Behavior, and Soc. Networking, vol. 6, no. 5, 2003, pp. 467–476.

[26] J. D. Mayer and Y. N. Gaschke, "The experience and meta-experience of mood," Journal of Personality and Social Psychology, no. 55, 1988, pp. 102–111.

[27] P. J. Lang, "Behavioral treatment and bio-behavioral assessment: computer applications," in Technology in Mental Health Care Delivery Systems, J. B. Sidowski, J. H. Johnson, and T. H. Williams, Eds. Norwood, NJ: Ablex, 1980, pp. 119–137.

[28] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," Current Psychology, vol. 14, no. 4, 1996, pp. 261–292. [Online]. Available: http://dx.doi.org/10.1007/BF02686918

One Hand or Two Hands? 2D Selection Tasks With the Leap Motion Device

Manuel César Bessa Seixas
FEUP,
University of Porto,
Porto, Portugal
ei09011@fe.up.pt

Jorge C. S. Cardoso
CITAR/School of Arts,
Portuguese Catholic University,
Porto, Portugal
jorgecardoso@ieee.org

Maria Teresa Galvão Dias
INESC TEC/FEUP,
University of Porto,
Porto, Portugal
tgalvao@fe.up.pt

***Abstract*—In this paper, we present the results from an experiment designed to compare two selection gestures (hand grab and screen tap) for the Leap Motion controller in 2D pointing tasks. We used the ISO 9241-9 multi-directional tapping test for comparing the devices, and we analyze the results using standard throughput and error rate measures as well as additional accuracy measures. We also present the results from the ISO 9241-9 assessment of comfort questionnaire. To complement this analysis, the computer mouse was also evaluated in order to serve as a comparison. Results indicate that the hand grab gesture performs better than the screen tap.**

Keywords-Interaction Device; Leap Motion; HCI; Pointing devices; Performance evaluation; Selection tasks.

I. INTRODUCTION

The Leap Motion (LM) controller is a 3D sensing device for hand gesture interaction. The LM is a small device that plugs to the computer via USB (it is also sold embedded in the HP ENVY Leap Motion Notebook PC and in the HP Leap Motion keyboard) and is operated by positioning the hands over the device. Through stereoscopic computer vision techniques, it is able to determine the position and orientation of the fingers of the hands, as well as the palm orientation and curvature. The controller can be used to point to a computer screen with a finger or with a tool (a pen or pencil, for example), or perform various hand gestures.

Although not meant to be a replacement of the mouse, many of the interactions with the LM involve pointing at a computer screen. There are situations where users would want or need to perform typical Windows, Icons, Menus, Pointer (WIMP) tasks with the LM, such as selecting buttons, navigating through menus and options or dragging graphical objects. Many applications in the Leap App Store are meant to give users various degrees of control over the computer, from selection and launching predefined applications and settings to scrolling content on webpages. Some applications even emulate the mouse, allowing cursor control and mouse actions [1]–[3]. Most applications that take advantage of the LM device still require users to perform typical WIMP tasks at some point (in many cases giving users the option of using the mouse or the LM device). For example, in many games users still need to select options and activate buttons; some software for surgery rooms also provides cursor control for specific functions [4].

If we assume that the LM device gains commercial traction and becomes embedded in additional laptop computers and desktop keyboards, we must also assume that it will become an additional alternative to typical WIMP tasks. In a situation where the user is operating the LM device in a specific LM task it may be faster to perform a WIMP task also with the LM, instead of moving the hand to operate the mouse.

Previous work [5] has evaluated the LM device for 2D selection tasks using a single hand for both the pointing and target selection actions. The LM performed poorly in that situation. One of the reasons that may justify the poor performance of the LM in that study is the fact that only hand was being used, forcing the user to move the pointer and select with the same hand. This may originate errors and delays in the "clicking" part of the gesture. We hypothesize that using a different hand for performing the selection gesture may improve the task.

In this work, we compare two gestures for selection tasks with the LM. One gesture uses the same hand for pointing and selecting; another uses one hand for pointing and the other hand for selecting. We also compare the selection task made with a traditional computer mouse. We used the standard ISO 9241-9 multi-directional tapping test [6] for pointing devices and calculated various accuracy measures [7] for the various selection gestures and devices. We have also used the ISO 9241-9 assessment of comfort questionnaire to get a subjective device preference, with additional questions regarding the selection gesture preference.

The contributions of this paper are: comparison of two gestures for selecting targets in 2d graphical interfaces with the LM device and a computer mouse; an analysis of the differences between pointing paths for the two LM gestures and the computer mouse; an assessment of the subjective preferences and comfort of the LM device versus the computer mouse; an assessment of the subjective user preference of the selection gesture.

The rest of the paper is organized in the following way. In Section 2, we present work that has used the LM device either to evaluate the device itself, or to evaluate new interaction techniques implemented with the LM. In Section 3, we describe the LM device in more detail. In Section 4,

 ISBN: 978-1-61208-382-7

we describe the experimental setup. In Section 5, we present and discuss the results from the experiment; In Section 6, we conclude.

II. RELATED WORK

Previous work has addressed the LM device from different perspectives.

Weichert et al. [8] analyzed the accuracy and robustness of the leap motion controller. They performed an experiment where a robotic arm would hold a pen in its hand and was programmed to place the tip in several real world known positions. These positions would then be compared to the ones acquired by the LM controller, being the difference between each other the precision. These measures were repeated several times in order to find repeatability, for two cases: static and dynamic (with a moving pen). They found the accuracy of the LM to be less than 0.2mm for the static case and less than 1mm for the dynamic case. Weichert et al. focused on the accuracy of device itself; in this paper, we focus on the accuracy of the user performing a task with the device.

Vikram et al. [9] present a new type of user input for writing, using the LM. Using the finger position data from the LM they are able to identify characters and words written "in the air". They propose an algorithm that is capable of recognizing gestures without pen down/pen up gestures to mark the beginning and end of a gesture. Although their interaction technique relies on users performing finger gestures, their analysis is concerned with the gesture recognition algorithm. In this paper, we address the issue of the performance of doing the gestures (for simple pointing tasks).

Nabiyouni et al. [10] performed a usability testing in order to find which of the implemented 3D travel techniques was the most efficient in bare-hand interaction. Five techniques were tested in a set of 3 tasks and the interaction was performed through the use of the LM controller. The techniques developed were based on a "Camera-in-hand" metaphor, where the Leap Motion workspace was directly mapped to the virtual world, and an "Airplane" metaphor, that, similar to driving a vehicle, had the camera always moving straightforward being the user responsible for controlling its velocity and orientation (the orientation was the same as the hand). A 3D virtual scenario, modeled as a city, was used to perform the tests. This is an example of a task that is out of the scope of our evaluation since it uses LM-specific features that are outside of the WIMP paradigm.

III. THE LEAP MOTION DEVICE

The LM is a small input device (7.6 x 3 x 1.3 cm) developed by Leap Motion Inc., which detects and recognizes users' hands posture and gestures (Figure 1).

Programmers can use the Leap Motion SDK (available for C++, Java, Objective-C, C#, Python, Javascript, and other programming languages) to develop applications that take advantage of the device's capabilities.

Currently, the SDK provides high-level functions such as:

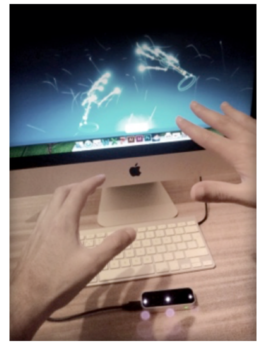


Figure 1. The Leap Motion device.

- Detection of the hands, and their 3D position in space, within the range of the LM.
- Orientation and curvature of the hand's palm.
- Overall scale, rotation, and translation motions calculated from the movement of the hands.
- 3D orientation and position of individual fingers and normalized 2D pointing position on the screen.

Applications developed for the LM can be distributed via the Airspace store [11], an online store from which users may download applications to use with their device. Several applications are currently available, from games to productivity applications.

The LM controller can be used as a traditional pointing device, but this functionality is not included directly in the driver software. To do this, an application must be used. Touchless [2] is an example of such applications, developed by Leap Motion Inc., with versions for Mac and Windows computers. Touchless provides several ways to interact with the OS:

- By pointing with a finger, users can control the position of the mouse cursor on the screen.
- By making a screen tap gesture (i.e., moving the finger towards the screen quickly), users can perform a mouse click.
- By swiping multiple fingers in the air, users can scroll horizontally or vertically.
- By pinching the fingers, users can zoom in and out.

IV. EXPERIMENT

A. LM Gestures

We compared two selection gestures for the LM device: screen tap, and hand grab (Figure 2). The screen tap gesture consists in moving the pointing finger towards the screen and returning the original position, quickly. This gesture is supported directly by the LM SDK that provides functions to configure the gesture's speed and motion amplitude and is an often-used gesture by applications on the Airspace store. The hand grab gesture requires two hands to point and select: the dominant hand is used for controlling the position of the pointer on the screen; the auxiliary hand is used to perform the selection by closing and opening the hand (i.e., making a fist).

To select these gestures for the experiment, we ran a preliminary session where we asked participants to try out

 ISBN: 978-1-61208-382-7

different gestures in the ISO 9241-9 multi-directional tapping test, and then collected their subjective preference regarding the gestures. In this preliminary session, each of the six participants was exposed to the following gestures: screen tap, hand grab, key tap, touch zone entered, and touch zone exited. The key tap gesture is performed with the auxiliary hand by flicking the index finger as if playing a piano key. The touch zone entered gesture uses a virtual vertical plane as a threshold: if the index finger crosses that threshold in the direction of the screen, a touch zone entered gesture is performed. The touch zone exited works in the opposite way to the touch zone entered: if the index finger crosses the threshold in the direction of the screen and than crosses it again in the opposite way, a touch zone exited gesture is performed. Participants experimented with all these five gestures and were then asked to rate them. The preferred gesture was the hand grab gesture, and the least preferred gesture was the screen tap. We thus decided to evaluate the performance of the two gestures that were rated best and worst and compare them to the computer mouse.

B. *Setup*

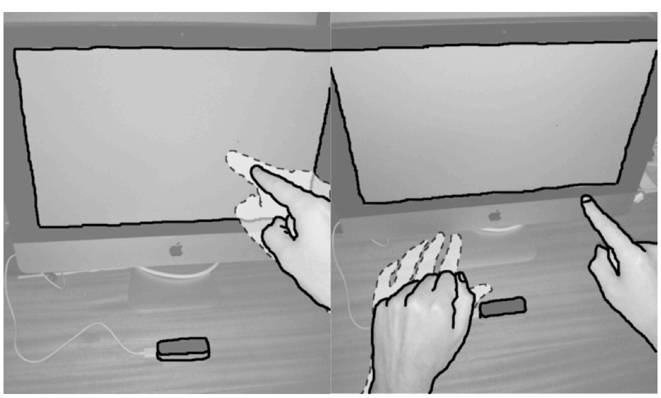
The experiment was a 3 × 5 × 7 within-subjects factorial design:

- Device {Mouse, LMScreenTap, LMHandGrab}
- Sequence {1,2,3,4,5}
- Block {1,2,3,4,5,6,7}

We configured the multi-directional tapping test with 16 circular targets, each with 13mm, in a circular layout with diameter of 180mm. The nominal index of difficulty used was 3.8 bits. The experiment was structured in “sequences” and “blocks.” A sequence corresponded to 15 target selections. A block had 5 sequences. Each participant was tested with all devices/gestures. The order of device/gesture differed for each participant according to a balanced Latin square.

We developed an application for collecting the pointer data for all devices/gestures, at 40 samples per second.

At the beginning of the experiment the participants were explained the purpose of the experiment, the task to be performed, and the devices to be used. Participants were also asked to fill in a questionnaire to determine their computer literacy and experience with the devices. Age and gender were also asked.



Screen tap Hand grab

Figure 2. Evaluated LM gestures.

Participants were asked to perform the selection task as fast as possible without exceeding one error per sequence. Participants were allowed to perform practice trials until they felt ready to start the experiment and could use their preferred hand to operate the devices. Breaks were allowed between sequences.

At the end of each device’s trials we asked participants to fill in the 12 item ISO 9241-9 comfort and effort questionnaire. At the end, we asked participants which device they preferred and several questions about the LM gestures. The experiment lasted about 1 hour and 30 minutes.

C. *Participants*

Nine non-paid participants (4 male, 5 female) were recruited. Their ages ranged from 10 to 35 years old. All participants were daily computer and computer mouse users (except one that stated to use the computer/mouse often). No participant had used the LM before.

D. *Apparatus*

We used the following hardware and software for the experiment:

- Apple Mac Mini (2.5GHz Intel Core i5, with 4GB RAM), running Mac OS X 10.8.3;
- HP L1706 LCD Display, with resolution set to 1280 x 1024;
- Genius Xscroll USB mouse, with the tracking speed set to third tick mark;
- Leap Motion device (commercial version), with tracking priority set to "Balanced";
- The Touchless software [2] for the screen tap gesture.

V. RESULTS AND DISCUSSION

Raw data from the experiment and R [12] analysis scripts are available at [13].

A. *Movement time, Throughput and Error rate*

Figure 3 shows the movement time (in seconds) as a function of block.

To estimate the learning effect, we ran pairwise t-tests for average throughput per block (considering all devices) with a significance level of 5%. The results indicate a clear learning effect in blocks 1 to 3, so these blocks are discarded in subsequent results.

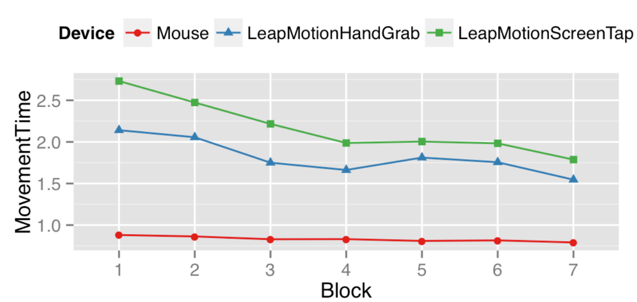


Figure 3. Movement time as a function of block.

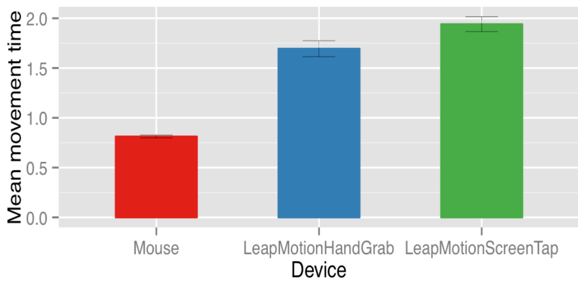


Figure 4. Mean movement time for each device/gesture.

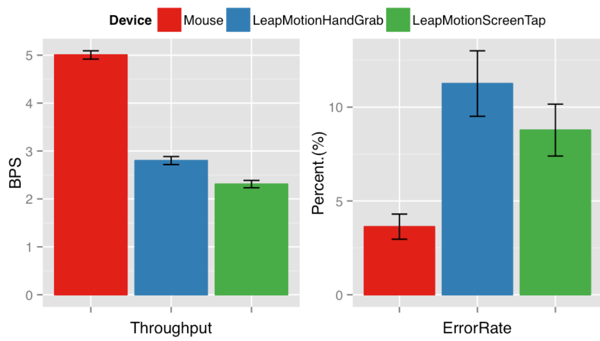


Figure 5. Throughput and error rate.

It is obvious that the mouse outperforms the LM in either gesture. The average movement time is 812 ms for the mouse, 1694 ms for the LM HandGrab gesture, and 1940 ms for the LM Screen Tap gesture (Figure 4). The average movement time for the mouse is less than half than for the LM device. Accordingly, the throughput of the mouse (Figure 5) is much higher than either LM gesture. However, it is also interesting to note that the hand grab gesture results in a faster overall movement time. A paired t-test comparing the two LM gestures shows significant differences in movement time (*t(179)=-6.0954, p-value = 6.539e-09*). The LM hand grab gesture represents a reduction in movement time of over 12% relative to the LM screen tap gesture.

The error rate for the mouse was also lower than the error rate for both the LM gestures. The differences between the two LM gestures were not statistically significant so, although the hand grab gesture seems to decrease the time needed to select an object, it does not appear to contribute to a less error-prone selection.

B. MacKenzie's accuracy measures

The Mackenzie's accuracy measures (see [7] for a description of the measures) allow us to see the differences between the devices/gestures in greater detail. Figure 6 and Table 1 shows the means, standard deviations, and F statistic for all accuracy measures. It also shows the *t* statistic comparing both LM gestures. Analysis of variance indicates that there are significant differences between devices for all measures except Movement Offset (MO). Student's t test comparing both LM gestures indicates significant differences in Task Axis Crossing (TAC), Movement Direction Change (MDC), Orthogonal Direction Change (ODC), Movement Variability (MV), and Movement Error (ME) measures.

As expected, the mouse outperforms the LM device in various measures (TRE, TAC, MDC, and ODC).

We can observe that, based in MV, ME and MO, the movement of the pointer when being controlled by the LM is quite similar to the movement of the pointer when controlled by the computer mouse. When comparing only the LM gestures, however, a few observations stand out as unexpected. The target re-entry measure (TRE), which measures the number of times the pointer re-entries the target before the selection is made, is equivalent in both LM gestures. We expected that the hand grab gesture would result in a lower TRE since the selection gesture is made with the auxiliary hand so selecting the target would not

TABLE I. MEANS AND STANDARD DEVIATIONS OF ACCURACY MEASURES FOR EACH DEVICE/GESTURE.

	Mouse		*LMHandGrab*		*LMScreenTap*			
Accuracy measure	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*F*	*t(179)*
Target re-entry (TRE)	0.10	0.08	0.37	0.25	0.37	0.30	78.4*	-0.0317
Task axis crossing (TAC)	1.61	0.34	1.92	0.67	2.24	0.67	53.1*	-5.338*
Movement direction change (MDC)	4.26	0.85	7.33	2.48	8.37	2.76	170*	-4.507*
Orthogonal direction change (ODC)	1.17	0.53	3.61	1.78	4.20	2.27	161*	-3.177*
Movement variability (MV)	20.62	7.08	26.40	13.37	21.87	7.01	11.4*	4.222*
Movement Error (ME)	20.09	5.56	21.11	10.79	17.73	5.48	9.16*	3.828*
Movement Offset (MO)	-2.46	6.63	-1.57	8.06	-1.81	5.18	0.85	0.335

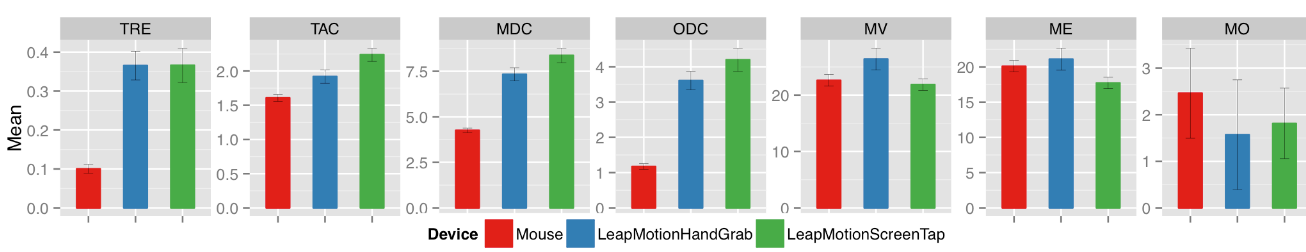


Figure 6. Accuracy measures for the three devices.

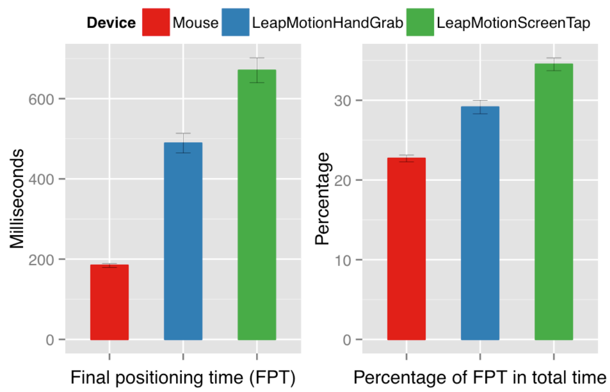


Figure 7. Final positioning time.

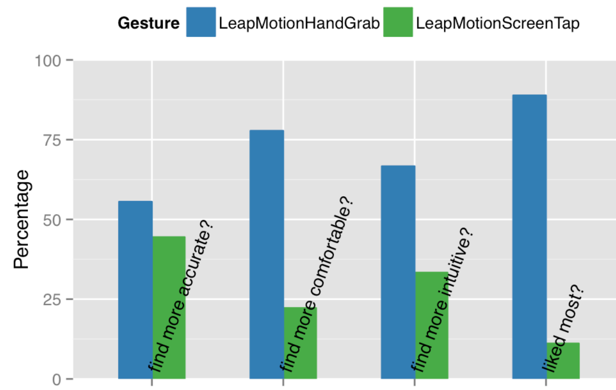


Figure 9. Gesture preference.

influence the pointer movement (as we expected to happen with the screen tap gesture). However, the results show no difference in TRE, indicating that maybe the selection gesture has little influence on the target re-entry measure, and that the higher TRE compared to the mouse is due to the pointer movement itself and not to the final selection gesture.

We also expected lower movement variability (MV) (and movement error – both are highly correlated) with the hand grab gesture than with the screen tap gesture. Again, we reasoned that because participants had separate control over the movement of the pointer and the selection of the target with the hand grab gesture it would result in more control over the pointer movement. However, the opposite seems to be true, having both hands over the LM seems to make it more difficult to control the pointer movement, resulting in a higher movement variability.

C. *Final positioning time*

To better understand the differences in the movement between the various devices/gestures, we analyzed the final positioning time (FPT) [14] for the selection task. The FPT measures the time it takes since the moment the cursor enters the target to the moment the user selects the target - we consider only the last target (re-)entry. Figure 7 shows the average FPT for the various devices/gestures in milliseconds, and the percentage that the FPT represents in the overall movement time. Again, it is clear that the mouse outperforms the LM, but more interesting to the current study it the fact that the hand grab gesture clearly reduces the FPT of the selection task ($t(179)= -10.73$, $p\text{-value}<2.2e\text{-}16$). However, it is also clear that the FPT for the LM device is still higher than that of the mouse. The difference in FPT of the two LM gestures explains most of the different in the overall movement time for the two gestures.

D. *Effort and comfort*

We also collected subjective device preferences and comfort through the ISO 9241-9 assessment of comfort questionnaire. Figure 8 shows the average scores for each question. As expected, for the evaluated task, the mouse was, in general, rated higher by participants.

E. *Users' opinion on the LM gestures*

At the end of the experiment participants were asked to indicate which device they liked best for performing this type of tasks. All the participants answer the computer mouse. We then asked participants to indicate which LM gesture they preferred by answering the following questions:

- Of the used gestures which one did you find more accurate?
- Of the used gestures which one did you find more comfortable?
- Of the used gestures which one did you find more intuitive?
- Of the used gestures which one did you like more?

Results are show in Figure 9 as the percentage of participants that preferred each gesture for each question. The results indicate a clear preference for the hand grab gesture, with only one participant saying he liked the screen tap the most.

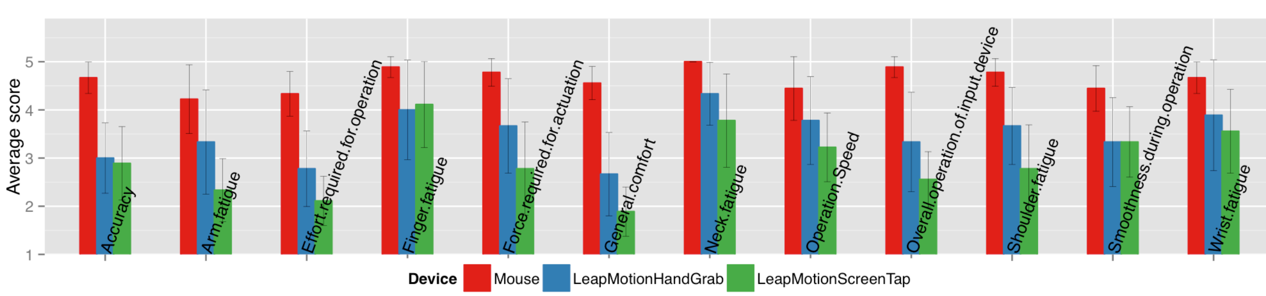


Figure 8. Average scores for the various comfort questions.

 ISBN: 978-1-61208-382-7

In general, participants also seem to find the hand grab gesture more comfortable and intuitive, but were more divided regarding whether any of the gestures was more accurate than the other.

VI. CONCLUSIONS

We have presented the results from an experiment designed to compare two selection gestures for 2D interfaces for the Leap Motion device. We compared the screen tap gesture to the hand grab gesture, in addition to the computer mouse.

Results indicate that the hand grab gesture that uses two hands improves the performance of the selection task when compared with the screen tap gesture. Movement time using the hand grab gesture is roughly 12% faster than using the screen tap gesture. This different is mostly accounted for by the lower final positioning time achieved when using the hand grab.

These results can be used when designing the interaction for the LM device, providing additional design options: one hand vs two hands, slower vs faster selection.

It is important to note that the comparison between the mouse and the LM is not completely fair. The mouse uses a non-linear mapping between device displacement and cursor displacement: faster movements translate to greater cursor displacement. This does not currently occur with the LM, but it would be interesting to try to implement a similar technique for the LM. It would also be interesting to evaluate and compare further selection gestures.

We should stress out that this study must be interpreted with care. We performed an evaluation of a very specific graphical interaction 2d task, for which the LM was not specifically designed. We believe that the LM may be used for these tasks, and hence it is important to know how it performs, but it is more suited for general gestural interactions, which was not the focus of the current experiment.

ACKNOWLEDGMENT

This paper was financially supported by the Foundation for Science and Technology — FCT — in the scope of project PEst-OE/EAT/UI0622/2014. We would like to thank all the participants in this experiment for their time and collaboration.

REFERENCES

[1] P. Lab, “Pointable,” 2014. [Online]. Available: https://apps.leapmotion.com/apps/pointable/windows. [retrieved: December, 2014].

[2] Leap Motion Inc., “Touchless for Mac.” [Online]. Available: https://airspace.leapmotion.com/apps/touchless-for-mac/osx. [retrieved: December, 2014].

[3] Nu-Tech, “Mudra Mouse,” 2014. [Online]. Available: https://apps.leapmotion.com/apps/mudra-mouse/osx. [retrieved: December, 2014].

[4] A. Manolova, “System for touchless interaction with medical images in surgery using Leap Motion,” in Proceedings og the 9th INTERNATIONAL CONFERENCE on Communications, Electromagnetics and Medical Applications, October, 2014, pp: 2-6.

[5] M. Seixas, J. Cardoso, and M. T. G. Dias, “The Leap Motion movement for 2D pointing tasks: Characterisation and comparison to other devices,” In Proceedings of the 5th International Conference on Pervasive and Embedded Computing and Communication Systems. Angers, France. 2015. (in press).

[6] International Organization for Standardization, “Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Requirements for non-keyboard input devices,” ISO, 2000.

[7] I. S. MacKenzie, T. Kauppinen, and M. Silfverberg, “Accuracy measures for evaluating computer pointing devices,” in Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’01, 2001, pp. 9–16.

[8] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, “Analysis of the accuracy and robustness of the leap motion controller.,” Sensors (Basel)., vol. 13, no. 5, Jan. 2013. pp. 6380–93.

[9] S. Vikram, L. Li, and S. Russell, “Handwriting and Gestures in the Air, Recognizing on the Fly,” CHI 2013 Ext. Abstr., 2013.

[10] M. Nabiyouni, B. Laha, and D. A. Bowman, “Poster: Designing Effective Travel Techniques with Bare-hand Interaction,” in EEE Symposium on 3D User Interfaces (3DUI), 2014. pp. 139-140.

[11] Leap Motion Inc., “Airspace store,” 2014. [Online]. Available: https://airspace.leapmotion.com/. [retrieved: December, 2014]

[12] R Core Team, “R: A Language and Environment for Statistical Computing,” Vienna, Austria, 2014.

[13] J. C. S. Cardoso and M. Seixas, “Leap Motion experiment raw data and analysis scripts - FEUP,” 2014. [Online]. Available: http://dx.doi.org/10.6084/m9.figshare.1263636. [retrieved: December, 2014]

[14] M. Akamatsu, I. S. MacKenzie, and T. Hasbroucq, “A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device.,” Ergonomics, vol. 38, 1995. pp. 816–827.

Developing Evaluation Matrix of Digital Library Interface by Analyzing Bloopers of Korean National Digital Library Sites

Miah Kam, Jee Yeon Lee
Department of Library and Information Science
Yonsei University
Seoul, South Korea
makiyma@hanmail.net, jlee01@yonsei.ac.kr

***Abstract*— The importance of digital libraries is increasing with the advancement and proliferation of networked online services. This work in progress focuses on developing an evaluation model for analyzing the real-world cases. Firstly, the web bloopers of the Korean digital libraries were identified, then heuristic evaluations were applied to categorize the bloopers into five types, which occur in three main components of the digital libraries. The resulting evaluation matrix consists of one axis for the web blooper types and the other for the digital library components. Each matrix cell has different weighting derived from the heuristic evaluation of the digital libraries in service. Digital library developers, managers, and subject matter experts should be able to consult the evaluation matrix to improve the usability and accessibility of their libraries. Our digital library evaluation matrix, based on the heuristic evaluation model, should raise the efficiency of digital library user interface evaluation.**

Keywords-Digital Library User Interface; Library Service Components; Web Bloopers; Heuristic Evaluation; Evaluation Matrix.

I. INTRODUCTION

The advancement of information technology enabled ordinary people to browse and access online library resources with ease. This new mode of access caused fundamental changes in library user behavior.

Users often visited physical libraries to get the information they needed, but today, more users with information technology exposures, initially access digital libraries before visiting physical libraries [1][2][3]. In addition to many obvious advantages, the ubiquitous access to digital libraries reinforced its use and importance as a *de facto* source of information.

Previous studies found usability, especially interface usability, to be one of most important factors in understanding user satisfaction with digital libraries. Xie [4] found that 'interface usability' was the most important factor in assessing digital library user satisfaction. Hernon and Calvert [5] also claimed that 'ease of use' was one of the most significant factors in measuring e-service quality.

Indeed, many digital library-related studies focused their research on the user interface aspect of libraries, such as Liew [6]. Hariri and Norouzi [7] classified digital library-related topics into 3 groups: (1) user interface and digital libraries, (2) digital libraries and usability, and (3) other studies related to user interface. However, these studies did not provide concrete guideline on how to develop or improve the digital library interfaces. To fill this gap, our aim is to develop a digital library blooper matrix that can be applied easily by practitioners to improve the interfaces of their digital libraries.

The term *blooper*, which refers to a silly mistake, was introduced by Jeff Johnson [8] in 2000 to describe his finding of problematic user interfaces. He conceptualized the Graphical User Interface (GUI) and web bloopers as mistakes, that are committed frequently in designing the interface and consequently influence usability. Web bloopers are often used as a checklist, which guides what not to do in detail and helps managers to improve interface efficiently and effectively. Web bloopers are closely related to the heuristics evaluation method, which was introduced by Nielsen and Molich [9]. Both methods use a checklist to identify usability problems. The heuristics evaluation method is "a method of reviewing the usability of software to find potential problems. Reviewers go through the software systematically with a list of UI design guidelines in hand, noting places where the software's UI violates the guidelines" [8]. Web bloopers show real-world examples of what not to do in the interfaces and, thus, it is possible to simulate the heuristics evaluation by counting how many problematic features exist.

The concept of web bloopers has not yet been fully examined in the academic community, although this concept has great research potential. Only some studies mention web bloopers [10][11], because this concept is firstly written for the practitioner's community. It is difficult to find digital library related studies that use web blooper related ideas to either evaluate or implement the interfaces. Thus, our work in progress attempts to determine whether the use of web bloopers can effectively improve digital library user interfaces.

The Korean digital libraries became more accessible and interactive for users with the advancement of the digital information technology. Although digital libraries place greater concern on the searching and full-text viewing related problems than other web services, libraries share many common usability requirements with other services. In this study, the practical notion of web bloopers was combined with the examination of the digital library specific usability issues to evaluate the Korean digital libraries' interfaces.

The aims and method of this research are as follows. Firstly, we inductively generate web blooper types and digital library components by analyzing the operational Korean national digital libraries such as 'Dibrary of the National Library of Korea', 'The National Assembly Digital Library' and 'The National Library for Children and Young Adults'. Secondly, we develop a digital library blooper matrix with the web blooper types and digital library components. Finally, we develop a digital library user interface evaluation matrix based on the heuristics evaluation.

In Section II, five cases of the actual web bloopers were shown. In Section III, we discovered bloopers and components which are divided into two axes. In Section IV, we developed an evaluation matrix by using the evaluation model and assigning weightings. Finally, in Section V, we concluded this work to show contributions, applicable area, limitations and potential future works.

II. CASE ANALYSIS

Three Korean national digital library sites were analyzed and over 260 web bloopers were found. The web bloopers were categorized into five groups according to their characteristics considering simplicity of errors, the amount of information and convenience of use. This categorization were used to evaluate usability regarding user interface design.

A. *Case 1 - User Support and Purpose of Operation*

The blooper shown in Figure 1 was from the Online Archiving & Searching Internet Sources (OASIS) [12] site of the National Library of Korea. This site behaves differently for different web browsers. The 'Back' button does not work when viewing the site via Internet Explorer (IE): however, it works when viewing via Google Chrome. This kind of inconsistency is an example of what can go wrong in the 'general website' component of the digital libraries.

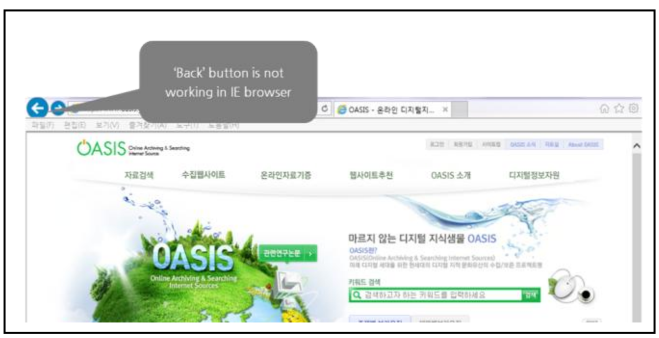


Figure 1. A web blooper of 'User Support and Purpose of Operation'.

B. *Case 2 – System Menu and Navigation*

The blooper shown in Figure 2 was found in the Government General Gazette of the Chosun [13] site of the National Library of Korea. Selection of one of the search results did not always produce the expected outcome. Unselected results sometimes showed up or nothing showed up at all. This can be regarded as a navigation problem and an example of what can go wrong in the full-text viewing component of digital libraries.

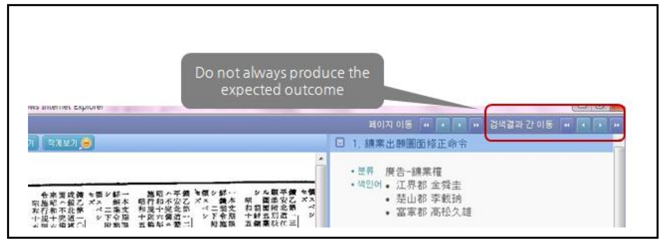


Figure 2. A web blooper of 'System Menu and Navigation'.

C. *Case 3 - Motion and Interaction*

This web blooper of Figure 3 occurred in the Dibrary [14], which is the digital library for the National Library of Korea. There was a problem with the checkboxes, which limited the search scope to a specific resource type. The checkboxes were under the main search menu. It was not possible to uncheck the boxes unless the user selected another checkbox. This probably occurred as the checkboxes were implemented as radio buttons. This problem occurred in the 'general website' component.

Figure 3. A web blooper of 'Motion and Interaction'.

D. *Case 4 - Information Provision*

The case in Figure 4 was also gathered from the Dibrary [14] site. In the federated search, which targeted resources on other sites, only the top five results were shown for each site. To see the lower ranked results, users had to go to the corresponding external site. This problem was categorized to occur in the 'searching and search results' component.

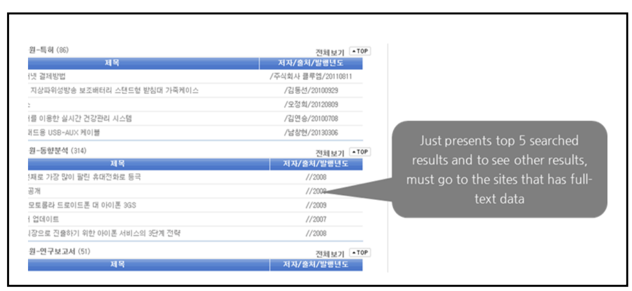


Figure 4. A web blooper of 'Information Provision'.

E. Case 5 - Visual design

The web blooper shown in Figure 5 was yet another case from the Dibrary [14] site. For the subject category search, it was difficult to tell which one was selected, as none of the icons or the font colors of descriptions' changed, even when the user's mouse pointer was on a specific subject description. This was a visual design problem, and it occurred in 'searching and search results' component.

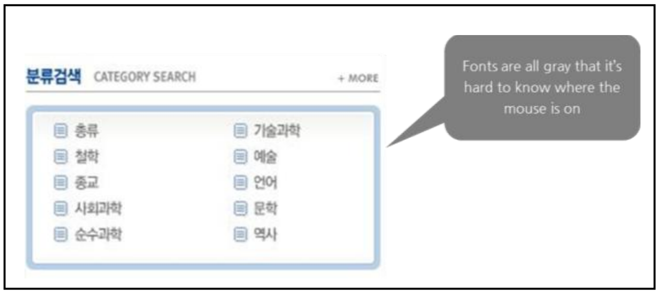


Figure 5. A web blooper of 'Visual design'.

There are also various other cases of web bloopers on Korean national digital library sites. By categorizing these bloopers and identifying corresponding digital library components where the bloopers were found, an evaluation matrix with two axes was developed.

III. DISCOVERING BLOOPERS AND COMPONENTS

A. Axis 1 – Bloopers Types

Based on the analysis of the discovered bloopers, the bloopers were categorized into five types: 1) User Support and Purpose of Operation, 2) System Menu and Navigation, 3) Motion and Interaction, 4) Information Provision, and 5) Visual Design. These types are similar to the ones used by Jeff Johnson [15][16]. Each type in turn was further divided to reflect the finer understanding of the bloopers. A brief description of five blooper types and the further divided sub types are as follows:

1) User Support and Purpose of Operation: supporting users, language, customization and browser;

2) System Menu and Navigation: clearance of navigation, structure and location path;

3) Motion and Interaction: matter of overlapped link, link motion, form, loading speed, system feedback;

4) Information Provision: related to relative link, consistency, relevance, recency, and understandability; and

5) Visual Design: icon, color, image, font and layout.

B. Axis 2 – Digital Library Components

By conducting a literature review [15][16] and analyzing ten prominent websites of libraries and information centers, three digital library components were identified: 1) general website, 2) full-text viewer, and 3) searching and search results.

IV. DEVELOPING EVALUATION MATRIX

A. Evaluation Model

The resulting matrix of Table 1 consists of five blooper types (22 subtypes) and three library components. The second and third stages of Jakob Nielsen's Heuristic Evaluation [17] were used to develop a weighted evaluation table. The five phases of Heuristic Evaluation were: 1) pre-evaluation training: give evaluators needed domain knowledge and information on the scenario, 2) evaluation: individuals evaluate user interface and make a list of problems, 3) severity rating: determine how severe each problem is, 4) aggregation: group meets and aggregates rating, and 5) debriefing: discuss the outcome with the design team.

TABLE I. WEIGHTED BLOOPER EVALUATION MATRIX OF DIGITAL LIBRARY INTERFACE.

		Digital Library Components(DLC)		
Blooper Type	***Blooper Subtype***	***DLC 1)***[a]	***DLC 2)***[b]	***DLC 3)***[c]
Type 1) User Support and Purpose of Operation	1)-1 : Users			
	1)-2: Language			
	1)-3: Customizing			
	1)-4 : Browser			
Type 2) System Menu and Navigation	2)-1 : Navigation			
	2)-2 : Structure			
	2)-3 : Location path			
Type 3) Motion and Interaction	3)-1 : Overlapped link			
	3)-2 : Link motion			
	3)-3 : Form			
	3)-4 : Loading speed			
	3)-5 : System feedback			
Type 4) Information Provision	4)-1 : Relative link			
	4)-2 : Consistency			
	4)-3 : Relevance			
	4)-4 : Recency			
	4)-5: Understandability			
Type 5) Visual Design	5)-1 : Icon			
	5)-2 : Color			
	5)-3 : Image			
	5)-4 : Font			
	5)-5 : Layout			

a. Digital Library Component 1), General website

b. Digital Library Component 2), Full-text viewer

c. Digital Library Component 3), Searching and search results

 ISBN: 978-1-61208-382-7

B. Weights

Each cell is colored differently to show different weights. The weights were assigned not just by counting the frequency of specific bloopers but also to reflect the needs of digital library sites. This should allow evaluators to refer to the matrix easily and efficiently when evaluating the interfaces. There are three levels in weighting. The most frequently occurring blooper types in a component with high importance are dark colored. The problem, which occurs on this level, should be fixed as soon as possible. Light gray-colored cells refer to less serious problems but with higher frequency of occurrence than the white-colored cells. The white-colored cells refer to general bloopers with lesser impact on the interface's usability than the light gray ones. In summary, these weightings was assigned by considering the seriousness of each observed example. The resulting matrix is as shown in Table 1.

V. CONCLUSTIONS AND FURTHER WORK

A. Contribution

The evaluation matrix should be used for assessing the usability of the digital library interfaces. As the matrix was developed by analyzing the sites in service, it should be also applicable to the real-world cases by the developers and managers of the digital libraries.

Most of the existing web bloopers were about general web sites and no special attention was paid to the particularities of the digital libraries. Thus, our digital library-specific evaluation matrix based on the heuristic evaluation model should raise the efficiency of digital library user interface evaluation.

B. Applicable area

The evaluation matrix can also be used as an evaluation tool of various general websites. Although the development started with the national digital library sites, this matrix should be applicable in various areas, because it is composed with combinations of essential elements of websites and critical web bloopers. This can be used in evaluating web sites such as search engines, and educational websites which need to be checked continuously to ensure the usability.

C. Limitations & future work

Although it was possible to find 260 web bloopers, the resulting evaluation matrix was only based on three Korean national digital libraries. Thus, it is not possible for us to claim strong reliability of the research outcome. Thus, additional digital libraries, especially in countries other than Korea, will be analyzed to augment the current evaluation matrix.

The resulting evaluation matrix with one axis of five web bloopers types and the other of three digital library components makes 15 cells and 66 cells when we further categorized the blooper types. In the future, each digital library component will be re-examined to check the benefit of further dividing each component. In addition, examples and explanations will be added to each cell to further assist the users of the evaluation matrix.

REFERENCES

[1] S. Joo, J. Lee, "Measuring the usability of academic digital libraries: Instrument development and validation," The Electronic Library, vol. 29 issue 4, 2010, pp. 523-537.

[2] I. Xie, "Users' evaluation of digital libraries (DLs): Their uses, their criteria, and their assessment," Information Processing and Mangement, vol. 44, 2008, pp. 1346-1373.

[3] J. Hwang and E. Lee, "Development of service quality measurement model and index for digital libraries", Journal of Korean Library and Information Science Society, vol. 41 issue 1, March. 2010, pp. 121-147.

[4] I. Xie, "Evaluation of digital libraries", Library & Information Science Research, vol. 28, 2006, pp. 433-452.

[5] P. Hernon and T. Calvert, "E-service quality in libraries: Exploring its features and dimensions", Library & Information Science Research, vol. 27 issue 3, 2005, pp. 377–404.

[6] C. L. Liew, "Cross-cultural design and usability of a digital library supporting access to Maori cultural heritage resources," Victoria University of Wellington: New Zealand, 2008.

[7] N. Hariri and Y. Norouzi, "Determining evaluation criteria for digital libraries' user interface: a review", The Electronic Library, vol. 29 issue 5, 2011, pp. 698-722.

[8] J. Johnson, "GUI Bloopers: Don'ts and do's for software developers and web designers," San Francisco, CA: Morgan Kaufmann Publishers, 2000.

[9] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," In Proceedings of ACM CHI'90 Conference on Human Factors in Computing Systems, 1990, pp. 249-256.

[10] J. Cappel James and Z. Huang, "A usability analysis of company websites," Journal of Computer Information Systems, vol. 48 issue 1, September. 2007, pp.117-123.

[11] M. Shelstad, "Content matters: analysis of a website redesign", OCLC Systems & Services, vol. 21 issue 3, 2005, pp. 209-225.

[12] OASIS, <http://www.oasis.go.kr/ctrlu?cmd=main> 2014.12.28

[13] Government General Gazette of Chosun, <http://gb.nl.go.kr> 2014.12.28

[14] Dibrary, <http://www.dibrary.net> 2014.12.28

[15] J. Johnson, "Web bloopers: 60 common web design mistakes, and how to avoid them," Morgan Kaufmann; 1 edition, April. 2003.

[16] J. Johnson, "GUI Bloopers 2.0, second edition: common user interface design don'ts and dos," San Francisco, CA: Morgan Kaufmann Publishers; 2nd edition, September. 2007.

[17] J. Nielsen, "Heuristic evaluation," in Nielsen, J. and Mack, R. L, Eds. New York: John Wiley and Sons, 1994.

Implementing the Tactile Detection Task in a Real Road Experiment to Assess a Traffic Light Assistant

Michael Krause, Verena Knott, Klaus Benger
Institute of Ergonomics
Technische Universität München
Garching, Germany
emails:{krause, knott, bengler}@ lfe.mw.tum.de

***Abstract*— The tactile detection task (TDT), a vibrating detection response task (DRT), was used to assess the mental demand of an in-vehicle information system (IVIS), which recommends a driving speed to the driver on a smartphone. In the experiment, the TDT was recorded as a single task, with the driving task as a baseline as well as with additional IVIS task, and with a cognitive task as reference and control. Results show that IVIS use did not significantly prolong the TDT reaction times, which can be interpreted as no increase in mental workload caused by the IVIS. The control task increased the reaction times significantly. The data of the real road experiment are analyzed in terms of correlations showing that the TDT was a reliable cognitive workload measurement tool in the experiment. Sideline: Driven speed revealed no correlation with TDT reaction times when the vehicle was in motion.**

Keywords-IVIS, detection response task; smartphone; cognitive workload

I. INTRODUCTION

A traffic light assistant for rural roads was developed on a smartphone in the Bavarian KOLIBRI project. The project involved human factors engineering along the whole project. The development was started in a static driving simulator for safety reasons. Selection of a favored graphical interface was done, related specifically to gaze behavior and assessment of the driving behavior, while using the assistant in the simulator [1][2]. After the system was validated in the simulator, further experiments were carried out on the real road. The mental demand needed to use the system is reviewed in the paper presented here.

Pauwelussen et al. [3] assessed their green wave assistant system in a driving simulator with a visual peripheral detection task (PDT), which would now be called (head mounted) detection response task (HDRT). They found for their system a significant increase in the reaction times of participants, which is interpreted as increased mental demand. An interface similar to Pauwelussen et al. [3] was not liked by our participants and was opted out by subjective ratings in an early stage of the development process [4]. The work presented here assesses the mental demand of the developed traffic light assistant, but has a focus on tactile detection task method (TDT) used.

The TDT is a version of the detection response tasks (DRT), which is currently subject to standardization by ISO TC22 SC13 WG8.

Section 2 will describe the methods, used in this experiment. In section 3 the results are presented and discussed; first the subjective ratings from questionnaires, than the objective measurement from the TDT. The TDT results were correlated between different experimental conditions, to the subjective ratings and to the driven speeds.

II. METHODS

A. Test Environment - Test Track

A section of federal road 13 (B 13) in the North of Munich was used for the individual test runs of the study. The test section of B 13 has a length of almost 7 km and 7 light signals control the individual intersections. In addition the route has two lanes in each direction. There is a median barrier separating the opposing highway lanes. Under dry conditions the speed limits on the track are usually restricted to 100 km/h and reduced to 70 km/h near intersections. In the individual test runs the test track was driven either from the North (GPS N 48.303477, E 11.572568) to the South (GPS N 48.245692, E 11.602002) or in the opposite direction. During the experiment, two turning points at the end of the track were used for evaluation of the individual test run (questionnaires). The traffic lights were acting on a coordinated scheme (green wave). Rush hours, which have a highly directional traffic load, were avoided for the experiments. The traffic density was about 500 cars/h in each direction. The study was conducted in May/June 2012.

B. Test Environment - Test Vehicle

The test vehicle used for the study on the real test field was a BMW X5. For the test series of the experiment, the experimental setup was completed with the following systems: A smartphone of the type Samsung GalaxyAce S5830 (3.5-inch display, resolution of 320 x 480 pixels, Android version 2.3.3) was used during the experimental procedure as mobile driver information system. Comparable to the driving simulator studies already carried out, the nomadic device was mounted on the ventilation slots on the right of the steering wheel on the center console of the BMW X5 (Figure 1).

Figure 1 Experimental setup in the car

Also an instrument for measuring the reaction time (RT) was installed in the test vehicle, because the experiment applied the TDT method. For the test run, in which a cognitive task (CoTa) had to be performed in addition to driving and the detection task, a microphone was mounted behind the driver's seat and the speakers were fixed in the armrest in the middle area of the vehicle. Two cameras were mounted in the test vehicle to capture the road scene ahead of and behind the car. A self-developed program allowed the synchronous recording of objective data (TDT, camera pictures and GPS from smartphone).

C. *KOLIBRI Human Machine Interface (HMI)*

Initial studies in a static simulator resulted in findings regarding the representation of the human-machine interface (HMI) [1][2][4][[5]. Thus the state of the HMI-display of the traffic light assistant on the mobile device, which was used for the experiments in real traffic, includes the favored approach of the previous simulator-based studies. Figure 2 illustrates the display states of the HMI that are offered to assist the driver. The favored concept of the traffic light assistant on the mobile device includes a recommendation in the form of a speed carpet. The green area represents the speed recommendation that supports the driving behavior for a green wave. The shown car position corresponds to the current vehicle speed. The white vehicle indicates that at the present rate the driver is within this recommendation and his driving behavior is optimal for achieving a green wave (Figure 2-1). The black car indicates the contrary (Figure 2-2). The position of the pointer of the Heuer traffic light (on the right upper corner of the HMI) shows the current status of the next traffic light - green or red signal in form of a clock. When the speed limit is exceeded by more than 10 km/h, the speeding display will appear (Figure 2-3). If the vehicle is outside of the calculation criteria for a speed recommendation, "preparing to stop" ("Vorbereiten auf Halt") appears, which notifies the driver that the next traffic light will display a red signal on arrival (Figure 2-4). The fifth presentation - a combination of the Heuer traffic light and a countdown - appears at a speed less than 5 km/h at a red traffic signal. This function shows the waiting time until the traffic light changes back to green (Figure 2-5).

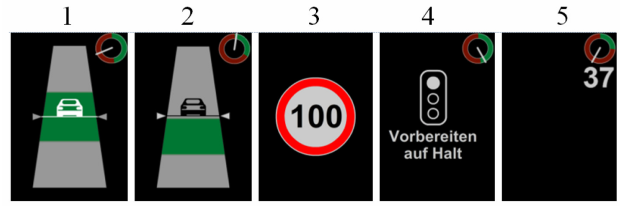


Figure 2 State of displays of the KOLIBRI-traffic light assistant

D. *Tactile Detection Task (TDT)*

A tactile detection task is a common method for assessing the mental demand of the driver. All DRTs show sensitivity to cognitive workload [6]. The attentional demands are measured in terms of reaction times and the hit rate [7]. The tactile stimuli are transmitted via a vibration motor. Merat and Jamson [8] propose application of the vibration motor on the left shoulder. The responses are typically given by pressing the button mounted on the left index finger against the steering wheel [7]. Figure 3 illustrates the mounting of the vibrating motor and the button on the test subject.

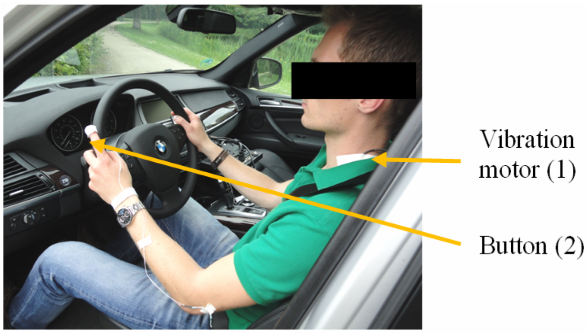


Figure 3 Attachment of the vibration motor (1) and the button (2) on the test subject

For the TDT realization in this study the test procedure was programmed into an Arduino Uno. The vibration stimulus was triggered every 3-5 seconds with duration of 1 second. The activation was interrupted when a reaction was carried out. This also provided the test subject a feedback to the response. The mental demand can be determined as a result of the reaction times and the hit rate, in which only a response time within 200-2000 milliseconds was counted as a hit [7]. Reactions faster than 200 ms were labeled as cheats. Responses slower than 2000 ms were logged as miss. The metric hit rate is the number of hits divided by the number of stimuli [7][9]. For quality reasons, the hit rate of a data segment should exceed 70% [9]. The TDT-method is usually - as in this study – simultaneously applied to the driving task and additionally other tasks, such as the operation of a vehicle internal system or the processing of a cognitive task (CoTa). Therefore, a combination of up to three tasks (e.g., driving + TDT + CoTa) is performed by the test subjects. The TDT works in these combinations as a measurement task to get the reaction times. The subsequent section deals with the method of the cognitive task. This threefold task condition is compared with a two-fold task execution, which does not include a secondary task. The difference can be interpreted as a measure of attentional demands; or different experimental conditions with the triple task setting can be compared.

E. Cognitive Task (CoTa)

The method of performing a cognitive task was used as a reference and control test. Simultaneous to the secondary task, the TDT is again carried out to determine the cognitive demand. Also terms, such as “Sternberg’s memory scanning task” or “Auditory Sternberg task (AUST)” are common in the literature for the here used ”Cognitive Task” [10]. The application Cognitive Task 1.0 from DaimlerChrysler AG (Stefan Mattes, 2005) was used in the experiment. The stimuli in the experiment were presented auditive. The test subject heard a three digit number (e.g., “3 5 2”). This was followed by a pause of 15 seconds for the cognitive processing. After this time interval, a further single digit (e.g., “3”) was announced, after which the subject had to decide whether the check digit was contained in the memorized sequence of digits as quickly as possible. Then, the participant gave a verbal response (“yes” or “no”).

F. Test Procedure

First, the test subjects explained their consent to study participation. In addition, a demographic questionnaire collected personal data. After that followed both the explanation of the test, which took up to 120 minutes time, and the description of the HMI on the nomadic device. The experiment began with a **pre-test** study in order to detect the reaction time on the TDT in the stationary vehicle, without any secondary or tertiary task. For this purpose, the vibration motor and the button were attached on the participant with suitable tape. Both vibration motor and button remained on the test person for the further course of the experiment. Afterwards a run on the described road section of the B13 for acclimatization followed while using the traffic light assistant. Four test conditions were then performed in randomized order: A baseline drive (**BL**), in which only the task of driving had to be fulfilled, and three conditions with execution of the TDT:

In addition to a **BL+TDT** run, there was a TDT-run in conjunction with the cognitive task (**CoTa+TDT**). In addition, one test run was combined with the KOLIBRI-traffic light assistant (**KOLIBRI +TDT**). In this test run, the smartphone was used to display speed recommendations. All conditions were driven only in one random direction, from North to South (NS), or in the opposite direction (SN). Except the condition **BL+TDT**, which was carried out in both directions (random first **BL+TDT NS** or **BL+TDT SN**) by each participant, to get more data from driving while operating the TDT and to check whether direction has an influence. To sum it up:

- **Pre-test** (TDT only, car standing)
- Driving and IVIS accommodation
- **BL** (driving only)
- **KOLIBRI +TDT** (driving+IVIS+TDT)
- **CoTa+TDT** (driving+Cognitive Task+TDT)
- **BL+TDT:**
    - **BL+TDT NS** (driving +TDT, direction NS)
    - **BL+TDT SN** (driving +TDT, direction SN)
- **Post-test** (TDT only, car standing)

After each of the individual test conditions a (raw) NASA-TLX questionnaire had to be completed by the participants in order to assess the subjective workload during the test run. For the two BL+TDT-conditions the NASA-TLX [11] was filled in after the first driven direction. After completing the test run with the nomadic device (KOLBRI IVIS), the System Usability Scale (SUS) questionnaire [12] was answered in order to assess the usability of the KOLIBRI-traffic light assistant and the AttrakDiff2 questionnaire was used to measure the quality and attractiveness of the system [13]. To determine correlations, as well as fatigue and learning effects, a study identical to the pre-test study was done at the end (**post-test**). The subjects were instructed to prioritize the tasks as follows: the driving task should be the highest priority, followed by the tactile detection task. The secondary tasks were in third position.

G. Participants

The study involved 23 test subjects. With an age range of 24 years for the youngest participant up to 58 years for the oldest subject, the arithmetic mean of the age of the subjects is 30.6 years with a standard deviation of 9.9 years. With 18 men and 5 women more than three-quarters of male subjects participated in the experiment. Two people had a self-reported red/green color weakness, while one person has an eye disease. All test subjects had a normal or corrected to normal visual acuity. 26% of the participants corrected the visual acuity by glasses. All persons were right handed. 48% of people knew the test track and 35% had already participated in one of the KOLIBRI driving simulator experiments. Except for one person, all participants had driven a car with automatic gear before.

III. RESULTS AND DISCUSSION

A. Questionnaires

1) NASA-RTLX (NASA-TLX raw)

The average subjective results of the NASA-RTLX shows a value of about 15 (SD: 10) for simple driving on the rural road. To work additionally on the TDT roughly doubles this value to 32 (SD: 13). In the triple task condition of driving, TDT and additionally engaging the CoTa the value is about three-fold: to 49 (SD: 16). The triple task with the KOLIBRI traffic light assistant (NASA-RTLX: 38 SD: 16) is rated in-between 32 and 49.

2) System Usability Scale - SUS

The System-Usability-Scale questionnaire used to assess the traffic light assistant in the KOLIBRI test condition resulted in a score of 79 (SD: 16). According to [14], this value can be related in between the adjectives good and excellent.

3) AttrakDiff2

The mean values of hedonic and pragmatic quality are in the “desired” sector, but the confidence rectangle touches “self-oriented”, “neutral” and “task-oriented”. Consequently, there is a tendency (not a full assignment) for “desired”.

B. Tactile Detection Task

The different experimental conditions resulted in reaction times as shown in Figure 4.

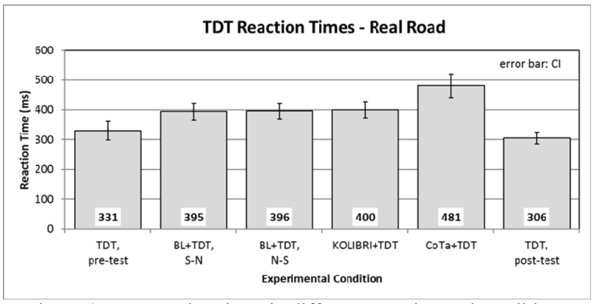


Figure 4 TDT reaction times in different experimental conditions

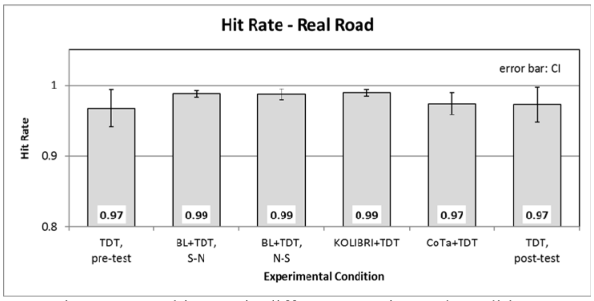


Figure 5 TDT hit rates in different experimental conditions

Associated hit rates for the experimental conditions are shown in Figure 5. T-tests (two sided, paired) detected no statistical difference in the reaction times between normal driving and driving with the KOLIBRI traffic light assistance:

BL-SN to KOLIBRI (t: -0.478; df: 22) p= 0.637

BL-NS to KOLIBRI (t: -0.615; df: 22) p = 0.545.

In contrast, the experimental condition CoTa prolonged the reaction times significantly.

BL-SN to CoTa (t: -6.15; df: 22) p<0.0001

BL-NS to CoTa (t: -7.43; df: 22) p< 0.0001

This indicates an increased mental workload and is in line with a slightly reduced hit rate for the condition CoTa. The hit rates for the TDT pre-test and post-test (only working on TDT) is also lower. The drop in the pre- and post-test is due to “Cheat” classification (reaction < 200ms); the drop in CoTa mainly due to Misses (reaction > 2000ms). Figure 6 holds the results of experimental condition baseline driving South-North (BL-SN) and shows a typical distribution.

During the experimental condition CoTa, the participant answered on average 14 Sternberg sequences (SD 1.7). The answers were to 98 percent correct, which is a clear indication of task engagement.

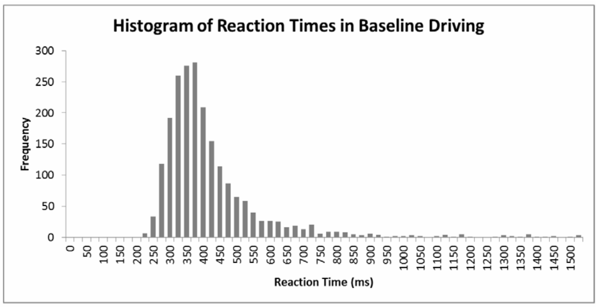


Figure 6 Histogram of reaction times in experimental condition baseline south-north (BL-SN)

For the condition KOLIBRI is no direct measurement of the task engagement available (eye tracking was not involved). From a former experiment in the driving simulator with TDT, the KOLIBRI system and an eye tracker, we know, that even with this combination the KOLIBRI system is very frequently observed [5]. Another study [15] in real traffic with the KOLIBRI system and an eye tracker, revealed an increase in the NASA-RTLX from about 14 to 20 between baseline driving and using the KOLIBRI system. This is similar to the shift from 32 to 38, in the here reported study (with TDT), which still does not indicate a high demanding task. These indications together with the observations from the examiner, led to the assumption that the KOLIBRI system was used by the subjects while driving.

C. Tactile Detection Task - Correlations and Repeatability

The experimental conditions pre-test and post-test, as well as baseline in direction North-South and South-North (BL-NS, BL-SN) can be seen as test-retest. Thus the correlations are interesting for these conditions. Table 1 holds the correlations of all conditions.

TABLE I. CORRELATION BETWEEN REACTION TIMES IN DIFFERENT EXPERIMENTAL CONDITIONS

	Pre	BL SN	BL NS	KOLIBRI	CoTa	Post
Pre	1.00					
BL SN	0.61	1.00				
BL NS	0.55	0.71	1.00			
KOLIBRI	0.66	0.70	0.86	1.00		
CoTa	0.66	0.68	0.81	0.80	1.00	
Post	0.59	0.65	0.64	0.50	0.75	1.00

The test-retest correlation of pre-test to post-test is r=0.59. The test-retest correlation between the baseline direction North-South to South-North is r=0.71. The time gap between pre-test and post-test is the entire experimental time of about typically two hours. Baseline driving in direction North-South and South-North is carried out directly after one another, but in random order. The pre-and post-test are identical (out of real traffic in a standing car), as in a laboratory. In contrast, the baseline driving can be influenced by uncontrollable circumstances (real traffic, overtaking, stops, etc.). On a larger scale, these microscopic events are centered out and are not visible (see condition BL+TDT,SN and BL+TDT,NS in Figure 4). The order (SN, NS) was random. In a further analysis the two baseline drivings were correlated, depending on the order of driving (first baseline to second baseline) independent of the direction; the result was again r=0.71.

From pre-test to post-test two-thirds (15 participants out of 23) get faster in the reaction times; on average 56ms (SD: 51ms). One-third (8 out of 23 people) gets slower in reacting; on average 32ms (SD: 22ms). The test-retest reliability of pre-test to post-test of r=0.59 is not very high, and perhaps the time gap of about two hours, with car driving and experimental conditions between test and retest, had an influence. For some persons in the form of a fatigue effect, for others in a training effect. The test-retest of driving in North-South and North-South direction of r=0.71 has a higher correlation. This result must be valued against the background of a real road experiment: The participants do not experience exactly the same situations each time while driving.

The other real road conditions also have a good correlation of 0.68 to 0.86. The internal consistency (split-half correlation) in pre-test is r=0.75 and in post-test r=0.83.

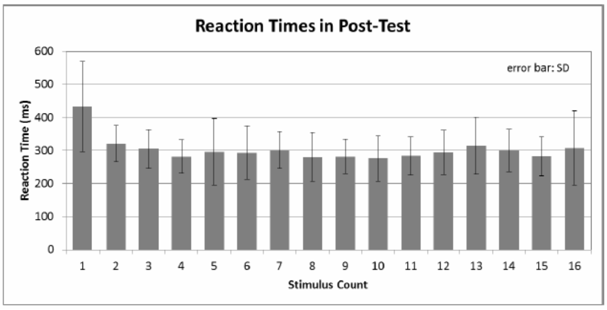


Figure 7 Average reaction times in post-test for every single stimulus

In pre-test and post-test, the participants were exposed to 16 stimuli. Figure 7 shows that participants need, even in the post-test (after extensive training on the TDT in different experimental conditions), at least one first "training" stimulus at the start of a condition in order to achieve nearly constant results.

NASA-RTLX scores were correlated to the reaction times. The correlation of individual reaction times in baseline driving South-North and North-South to individual NASA-RTLX scores is r=0.14 and r=0.09. In the condition KOLIBRI r=0.34 and in CoTa r=0.33. So, the congruence seems higher in triple task than in double task conditions.

D. *TDT Dependence on Speed and Acceleration*

For the following analysis each TDT reaction in the South- North-Driving-Baseline (BL SN) was normalized (divided) by the individual mean reaction in the TDT-only-pretest data. For example, if one test person had data while driving such as 238ms, 320ms and 281ms and a mean TDT-only-pretest (car standing) value of 274ms: the data were normalized by 274ms to 0.87, 1.17 and 1.03, to weaken intra-individuality. For the analysis, reactions were only accounted for when the vehicle was moving (faster than 0km/h) and the reaction result was a Hit (no Miss or Cheat). For the resulting N=1955 reaction time values the correlation to the driven speed while reacting revealed no correlation (r= -0.03). The same analysis for the opposite direction North to South revealed r= -0.04 (N=2022) and for the COTA-condition r= -0.04 (N=1975). Figure 8 is a (typical looking) plot of the speed versus the normalized reaction time. On the one hand, it is a positive aspect that the driven speed seems to not influence the reaction times for the TDT. Thus the procedure is more robust against nuisances. On the other hand, it is curious: If the TDT measures mental demand, why does a higher speed seem to not impose more mental demand on the driver? An analysis of absolute speed difference between stimuli (treating accelerations and decelerations the same way) revealed the same result. The correlation between absolute speed difference and reaction time for the South-North-Driving-Baseline (BL SN) is r=0.02. For the CoTa condition it is r=0.03. Therefore, even acceleration or deceleration seems not to influence the reaction time.

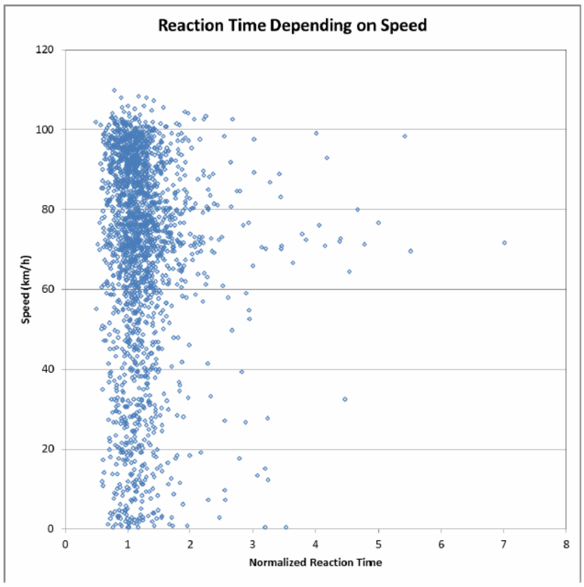


Figure 8 Normalized reaction times depending on speed in experimental condition BL SN

The duration of one experiment run was typically around seven minutes (about 100 stimulus onsets). For the experiment in South-North direction (BL SN) the correlation between normalized reaction time and the stimulus count up number (thus, related to experiment time) was r=0.04; for North-South run (BL NS) r=0.08 and for CoTa r= -0.02. It seems like the durations of about seven minutes of an experimental condition is not too long and fatigue effects do not affect the results.

 ISBN: 978-1-61208-382-7

With the GPS values, we were also able to map the reaction times to road segments of 250m length. The results revealed plausible results (like non-overlapping confidence intervals) for e.g., some long, straight sections with low demand and more demanding sectors.

IV. CONCLUSION

We could use the TDT in our experiment as an easy to administer and valuable tool. The TDT was able to detect the mental demand of a control task and showed no significant increase in reaction time, while using the carefully designed traffic light assistant.

ACKNOWLEDGMENT

This work was supported by Bayerische Forschungsstiftung (Bavarian Research Foundation). The authors would like to thank the project partners TRANSVER, BMW Group and the Center for Traffic Management of the Bavarian Road Administration.

REFERENCES

[1] M. Krause and K. Bengler, “Traffic Light Assistant - Driven in a Simulator”, in Proceedings of the 2012 International IEEE Intelligent Vehicles Symposium Workshops: IEEE Intelligent Vehicle Symposioum (IV-2012), Madrid, June 3-7, 2012, pp. 1-6.

[2] M. Krause and K. Bengler, “Traffic Light Assistant - Evaluation of Information Presentation”, in Advances in Human Factors and Ergonomics 2012 14 Volume Set: 6 Advances in human aspects of road and rail transportation: Proceedings of the 4th AHFE Conference 21-25 July 2012, G. Salvendy and W. Karwowski, Eds, Boca Raton, FL: CRC Press, Taylor & Francis Group, 2012, pp. 6786–6795

[3] J. Pauwelussen, M. Hoedemaeker, and K. Duivenvoorden, CVIS Cooperative Vehicle-Infrastructure Systems: Deliverable D.DEPN.4.1b Assess user acceptance by small scale driving simulator research, 2008.

[4] M. Krause and K. Bengler, “Subjective Ratings in an Ergonomic Engineering Process Using the Example of an In-Vehicle Information System”, in Human-Computer Interaction. Applications and Services: Springer, 2013, pp. 596–605.

[5] M. Krause, A. Rissel, and K. Bengler, “Traffic Light Assistant-What the Users Want”, in ACHI 2014, The Seventh International Conference on Advances in Computer-Human Interactions, 2014, pp. 235–241.

[6] A. S. Conti, C. Dlugosch, R. Vilimek, A. Keinath, and K. Bengler, “An assessment of cognitive workload using detection response tasks”, in Advances in Human Factors and Ergonomics Series, Advances in Human Aspects of Road and Rail Transportation: CRC Press, 2012, pp. 735-743.

[7] J. Engström, “The Tactile Detection Task as a Method for Assessing Drivers' Cognitive Load”, in Performance metrics for assessing driver distraction: The quest for improved road safety, G. L. Rupp, Ed, Warrendale, Pennsylvania: SAE International, 2010, pp. 90–103.

[8] N. Merat and A. H. Jamson, “The Effect of Stimulus Modality on Signal Detection: Implications for Assessing the Safety of In-Vehicle Technology”, Human Factors: The Journal of the Human Factors and Ergonomics Society, vol. 50, no. 1, 2008, pp. 145–158.

[9] N. Merat, E. Johansson, J. Engström, E. Chin; F. Nathan and T. Victor, “Specification of a secondary task to be used in Safety Assessment of IVIS”, AIDE Deliverable 2.2.3. http://www.aide-eu.org/pdf/sp2_deliv_new/aide_d2_2_3.pdf [retrieved: Jan., 2015].

[10] K. Bengler, S. Mattes, O. Hamm, and M. Hensel, “Lane Change Test: Preliminary Results of a Multi-Laboratory Calibration Study,” in Performance metrics for assessing driver distraction: The quest for improved road safety, G. L. Rupp, Ed, Warrendale, Pennsylvania: SAE International, 2010, pp. 243–253.

[11] S. G. Hart, NASA Task Load Index (TLX): Volume 1.0; Paper and Pencil Package. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000021488_2000015069.pdf [retrieved: Jan., 2015].

[12] J. Brooke, “SUS - A quick and dirty usability scale,” in Usability evaluation in industry, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds, London, Bristol, PA: Taylor & Francis, 1996, pp. 189–194.

[13] M. Hassenzahl, M. Burmester, and F. Koller, “AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität,” in vol. 57, Mensch & Computer 2003: Interaktion in Bewegung, G. Szwillus and J. Ziegler, Eds. 1st ed, Stuttgart: Teubner, 2003, pp. 187–196.

[14] A. Bangor, P. Kortum, and J. Miller, “Determining What Individual SUS Score Mean: Adding an Adjective Rating Scale,” JUS Journal of Usability Studies, vol. 2009 Volume 4, no. Issue 3, 2009. pp. 114–123.

[15] M. Krause, V. Knott and K. Bengler, “Traffic Light Assistant – Can Take My Eyes Off of You”, Human Factors and Ergonomic Society Europe Chapter Turin 2013.

Perspective and Use of Empathy in Design Thinking

Andrea Alessandro Gasparini
Department of Informatics
University of Oslo
Oslo, Norway
a.a.gasparini@ub.uio.no

***Abstract*—The paper takes a closer look into one of the main attributes of Design Thinking: Empathy. The motivation for doing so has its roots in the post Design Thinking period, which we are entering now, following a rich decade of the use of this approach to innovation. Approaching a Designerly Thinking perspective of what the designer does in practice, five different epistemological paths will give an understanding of the effects empathy has in the design process. Empathy is addressed by exploring two main aspects, the emotional and the cognitive. The theoretical perspective of Design Thinking, seen as a reflexive practice, or as a creator of meaning, or even as a problem solving activity, is used to understand how empathy can be used in a design context. This aspect is then further analyzed using the results of a large workshop where Design Thinking was used.**

Keywords-Empathy; Design Thinking; Designerly Thinking; Service Design.

I. INTRODUCTION

During the last decade, Design Thinking was considered by many to be one of the best ways to foster innovation and creativity in companies and organizations, to attempt to solve complex problems, also named wicked problems [1], and to innovate products and services. Nowadays the future of Design Thinking is unsure. From the management research field, some of the previously strongest supporters are confident the era of Design Thinking is over [2-5].

Others from the field of design research argue that more ownership of relevant parts of the method should be taken [6], and abandon those that do not work. This situation has many reasons, but risking oversimplification, we can say that the two fields, business management and design research, have been pulling design thinking in two very different directions.

On the one hand, the field of management adopted Design Thinking on their terms, as best explained by Nussbaum [4]: “Companies absorbed the process of Design Thinking all too well, turning it into a linear, gated, by-the-book methodology that delivered, at best, incremental change and innovation.” On the other hand, the design research developed a different approach. For the latter, thinking as a designer is not exactly a new savoir-faire, and therefore possibly, some of the relevant attributes of Design Thinking have been overlooked [6].

The strength of Design Thinking is the ability this approach has to combine the desirability a human can experience, and the economic viability and the technical feasibility of an innovative idea (see Fig.1).

As elaborative forces, Design Thinking uses rapid prototyping, abductive reasoning and empathy to enact innovative results. Although all three are of interest, the focus will be on the latter. Empathy will be addressed in this paper using two main aspects, the emotional and the cognitive.

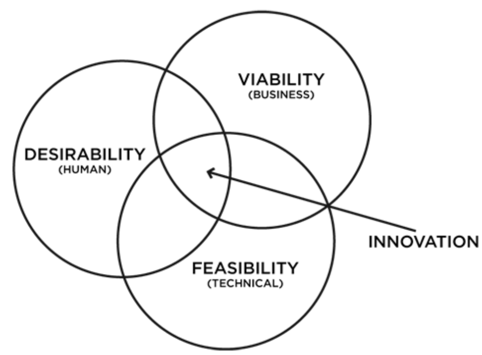


Figure 1. Overview of Design Thinking.

In order to emphasize the effect empathy may have in a design process we need to define different discourses explaining what designers really do in practice. Approaching a Designerly Thinking perspective on designer’s activity, five different epistemological discourses will be used [6] to explain the role the two aforementioned aspects of empathy havw in each discourse.

As a case study, to address the aforementioned intersection between Design Thinking and empathy, the paper will present results from a workshop organized by a Norwegian University Library addressing Open Access services. The methodologies used in this research are photo-ethnography and annotations in vivo. A longer discussion group meeting was done the day after, to collect additional data and perspectives.

The article is structured as follows: In Section 2, an introduction to Design Thinking is followed by Section 3 where the Designerly Thinking perspective is presented. Section 4 is an introduction to empathy. Section 5 will present the role empathy has in five different perspectives of Designerly Thinking. Section 6 highlights the results from the workshop, while Section 7 is the discussion, and Section 8 concludes the paper.

II. Design Thinking

As an approach, Design Thinking may help to innovate services and products. Using a combination of immersive understanding of users need, rapid prototyping, and abductive reasoning, this approach may result in the best possible solution [7][8]. The rapid prototyping act in the Design Thinking process, provokes small advancement based on partially known information [2], and the outcome is a high quantity of results one can further analyze. Sorting the outcome in the end, the result of a Design Thinking process, gives a handful viable ideas. A negative and interesting side effect is the problem with the "recreation" of all the steps of the rapid prototyping process [9]. This will leave out some learning issues for the participants where this approach is used, and obviously probing a repetition of the process with the same outcome. The empathy ensures a broad representation of knowledge in the process [10], while abductive reasoning explains why the result is adequate, given the context. Overall, the process of Design Thinking allows the result to be a technical viable solution, a desirable output for users and an economical feasible project [7] (see Fig. 1). Another strength Design Thinking has, is the possibility to tackle complex and ill-defined problems [8], ranging from business to societal context [11]. These types of problem, often coined "wicked" [1], have been debated in the design milieu for an extended period of time, as they can be an entry port for designers into new areas [9], bringing new understanding to complex issues.

III. Designerly Thniking

To emphasize the effect empathy may have in a design process, we need to define different discourses explaining what designers really do in practice. Designerly Thinking addresses how the practice of being a designer, and the theories trying to explain and understand the act of designing, coexist in the same sphere, and how we can understand the two from an academic perspective [6].

Approaching a Designerly Thinking perspective on designer's activity, five different epistemological discourses will be used [6] to explain the role the two aforementioned aspects of empathy has in each discourse:

- Design and Designerly Thinking as the creation of artefacts. The core concept is the science of the artificial [6].
- Design and Designerly Thinking as a reflexive practice. The core concept is the reflection after the creation process, helping the cyclic process of designing with added new competence and understanding.
- Design and Designerly Thinking as a problem-solving activity. The core concept is a methodology to solve complex and wicked problems [1].
- Design and Designerly Thinking as a way of reasoning and making sense of things, and is based on the practical activity done by designers.
- Design and Designerly Thinking as creation of meaning. The core concept is relevant as it links theory and practice.

Each of the five discourses has their own theoretical foundation and background [6], and also describes the concepts they address.

The relation between empathy and the aforementioned theoretical discourse of Design Thinking will be discussed in Section 7, where this pluralistic perspective will try to point out what empathy contributes to the "designerly" part of the process.

IV. Emapthy

It is possible to divide reflections around empathy into two main dimensions. The first may be seen as an emotional empathy, being an instinctive, affective, shared and mirrored experience [12]. More specifically, as a person, one feels what other people experience. The other dimension of empathy is cognitive, where one understands how others may experience the world from their point of view [12][13]. Since this is a state that is not actually experienced by a person, it may lead to misunderstandings and subjectivity. The lack of a degree (how much of empathy one experiences) may reinforce this problem [14].

In a design process, we can address and use empathy in different ways. Firstly as a tool to design with, requiring the transformation of this emotional feeling in an attribute [15]. Secondly designers can use empathy to acquire insight into users' needs and in doing so, inform the design process [7]. For instance, in a Design Thinking process all the participants in a design team need to be empathic with the users they are designing for in order to create relevant solutions. Using an approach toward cognitive empathy, designers apply different methods to build up that competence and insight, enabling them to prioritize the needs of the users and make the results of the process more desirable [7].

Designers may use a variety of approaches to gain a cognitive empathic insight. The following two examples illustrate how this can be done. First, designers can use "experience prototype". Using a medical wearable device [15][16], like a small remote heart monitoring device, over a period of time, would inform a design team of how a person wearing the device feels in everyday situations. This would be very difficult to understand otherwise. For example, driving to work, taking a bus or eating, are easy tasks that, for a person with special needs, may be extremely hard to perform. Then, the design team may get insight into how difficult it is to perform these simple tasks and can gain emapthy by understanding. The second example is related to how a group of interaction design students solved their project task. The task was to design a rescue boat. In order to gain an understanding of the experience and feeling of getting rescued at sea, they rescued each other in a swimming pool, and the empathic insight helped them to develop a very interesting and relevant prototype. Both examples show how to acquire a cognitive empathic understanding and insight, in this case, the designers did not need to *feel* what the real experience was.

Design Thinking invites participants in a design process to share their own empathic insights related to the task at hand. In fact, this is one of the strengths of the approach: all participants bring empathy into the process.

Cognitive empathy has also an effect on the way participants of a design team work together. It is observed in [17] that differences in competence and knowledge between members of a design team do not affect the overall team performance, since the empathy for others participants points of view, expressed as a cognitive based “social sensitivity”, functions as an equalizer [18].

In the design process, the participants contribute to the process through different roles: as themselves, as designers, librarians, managers, IT people and so on [8], bringing with them the cognitive empathy represented by the roles they hold. In addition, they may also have empathy for the role of a user they argue for or against in a given context of the project.

V. DESIGNERLY THINKING AND EMPATHY

Mapping the pluralistic perspective of the theoretical discourse of Design Thinking, also the Designelry Thinking aspect, and how we can use empathy in the design process, we can tentatively produce the following overview presented on Table 1. This will be further analyzed in Section 7.

TABLE I. OVERVIEW OF DESIGNERLY THINKING AND EMPATHY

Theoretical Perspective	***Core Concept***	***Empathy***
Design and Designerly Thinking as the Creation of Artefacts	The science of the artificial	Emotional
Design and Designerly Thinking as a Reflexive Practice	Reflection in action	Cognitive and Emotional
Design and Designerly Thinking as a Problem-Solving Activity	Wicked problems	Cognitive (Holistic)
Design and Designerly Thinking as a Practice-Based Activity and Way of Making Sense of Things	Designerly ways of knowing	Cognitive (Constrained)
Design and Designerly Thinking as a Creation of meaning	Creating meaning	Cognitive (Interpretation of context)

VI. THE CASE

The analysis of empathy in regards to Designerly Thinking in this paper is compared with findings from a day long workshop organized by the University of Oslo Library to address Open Access services provided by the University of Oslo.

The problem was a lack of cooperation and coordinated strategies by all the departments involved in helping researchers use Open Access as a channel for publications. To envision and map how the journey for a researcher would looks like when approaching all the different steps to, at the end, publish an Open Access article, the workshop used a specific method belonging to the Design Thinking sphere. “User journey” and “touch points” [19] are widely used to address services. User journey is the representation of all the steps a customer need to perform to achieve the final goal of the service. An easy example is a trip using a plane. The journey starts usually online where the ticket is bought and then printed. The next stage is arriving at the airport using train, bus or car, and then the trip goes further until the customer lands at the destination. Each situation where the user is in direct contact with the service provider is called a touch point. Using Service Design Cards [20][21] (see Fig. 2) one can make the aforementioned journey, where each card is a touch point.

Figure 2. Representation of a User Journey from one of the workshops groups.

For instance in Fig. 2, the user journey in the picture is from one of the groups in the Open Access workshop. The journey represents a researcher that has finished a research project, and then realizing that the contract done with the research project fund provider requires publications in an Open Access journal or make available the papers in an Open Access repository.

As represented in Fig. 3, the researcher then may use different touch points to achieve that goal. Those touch points are well suited to be redesigned in the spirit of Design Thinking method. One can, for instance, combine them or replace them with new technologies that improve the service experience.

The methodologies used in this study are photo-ethnography and annotations in vivo. Eighteen participants were invited. Two participants were interaction designers, four participants were from the library working with Open Access, and four participants from the library working directly with researchers. Finally, eight participants from different department of the University working as research consultants, in regard to different types of research projects. After an hour with an introduction to Design Thinking, all the participants were then formally divided in three groups. Each group had a unique exercise consisting of different user journey researchers may have had when publishing Open Access. A longer discussion group meeting the day after with three of the library staff that had organized the workshop, gave the possibility to collect additional data and perspectives. Analysis of pictures, annotations and results from the discussion group meeting, was performed in regard to empathy.

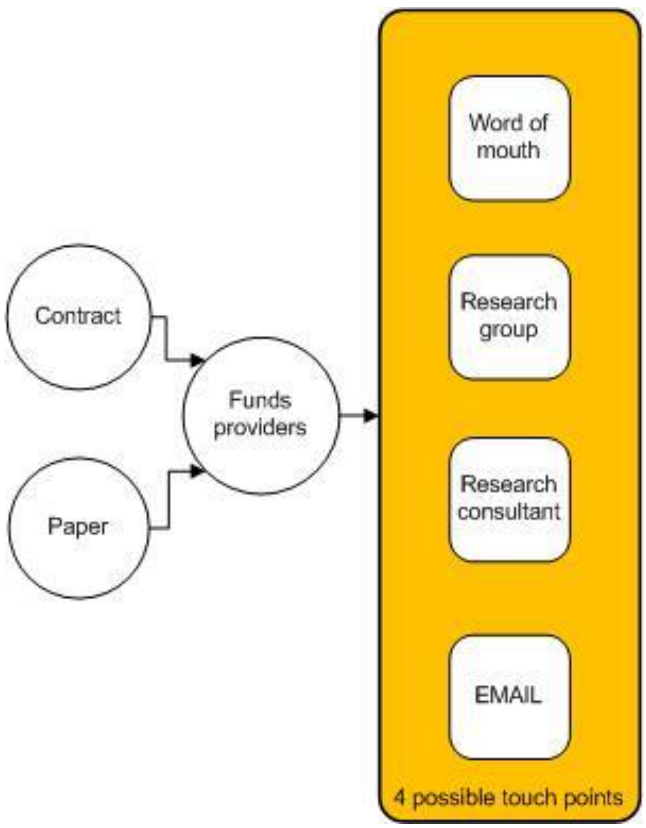


Figure 3. Excerpt from Figure 2, showing a very complex group of touch points.

Observations, among others, concluded that all three groups managed to develop good user journey. Discussions about issues researchers had in their work and research situations were contextualized and represented well by the Service Design cards. On the other hand the participants discussion about the issues the research faced were not unanimous. The researcher needs were not equally supported. Different department had different methods to approach and help the researchers, resulting in misunderstandings about the users perspective.

VII. DISCUSSION

Table 1 shows some interesting points to be discussed. In the first row, the Designerly Thinking perspective, invites understanding the making of an artifact as the core result. In the Design Thinking process, when creating artifacts, one of the generative drives is the making of various prototypes in a rapid way.

An instinctive and affective experience, as in emotional empathy, can be necessary to foster creativity and innovation [22] when a designer is prototyping in a rapid way. The use of tactile, visual and audio inputs in the creation process, can explain the necessity of the designer not using cognitive empathy.

Also the "quantity" and "quality" of empathy applied probably cannot be equal in all the prototypes. In this case a graduation of empathy can be used as an extra indicator to help designers choose the most relevant prototype.

The next phase of a design process can be the selection of the best prototypes. The type of empathy used in this situation seems not to be cognitive, therefore it can be more relevant to focus on emotional empathy. To sort out all the prototypes, an instinctive, emotional, affective experience can be a valuable first insight and can make the design process more effective. For instance, trying a new model of a bike, gives naturally a better insight then *imagining* how a user experiences the ride.

The second row has also some points worth to mention. The reflexive practice based on Schöns [23] approach, implies a reflection-upon-the-creation effort from the designer. As a result, the practical competence can have an incremental learning boost [23]. Cognitive empathy may explain partly how the designer embodied the improved new competence gained from the practice and their tacit knowledge. Using emotional empathy, on the other hand, we may explain what the effects of instinctive, affective and emotional new experiences, are in relation to their own abilities as designers, creativity and theirs learning processes. A possible use of this relation between reflection-upon-the-creation and empathy can be in the context of the educational curriculum to form design practitioners [24].

The third row is straightforward when it comes to empathy. Large complex problems, also known as wicked [1], can only be solved if the design process takes in accounts a holistic view of the user needs. Point eight in the definition of the properties of a wicked problem states the necessity to take in account that "solving a wicked problem is one shot operation with no room for trial and error."[1]. This definition requires from the designer a deep insight of the problem area and the user perspective. Cognitive empathy can, in this regard, be a valuable source of information.

Row four addresses partly the reflective tradition of [23] and the experience-centered design [25], nevertheless it has a scope more focused on how practitioners elaborate knowledge grounded in practical experience. Cognitive empathy may have a role since the result of making an artefact must be grounded in knowledge on prior usage. For instance, the designer of an Alessi coffee maker, must base

 ISBN: 978-1-61208-382-7

the new idea on prior knowledge of how such an artifact works. At the same time, a cognitive empathy is needed to understand how the new design can change the experience when making a coup of Italian coffee.

Finally row five advocates for a Designerly Thinking approach to the act of creating meaning. In this case the artifact is only a medium to articulate and transmit the result of the creation [6]. The Design Thinking process already from the first immersive stages of discovery and interpretation process [7][8], seems to gain substantial support from cognitive empathy, giving insight of user needs and the context.

The analysis of data from the workshop and the discussion group meeting was done in regards to how empathy works and gives additional support for the aforementioned understanding. The main findings suggest a need to better include the cognitive empathy of the participants, and also understand how the interpretation of different contexts may contribute on the process of including empathy. This is in line with the Designerly Thinking understanding in row 5 in Table1, the Design and Designerly Thinking as a creation of meaning.

Related to the aforementioned workshop, to help the outcome of a design process where the results are better services, the participants need to experience how the situation of a researcher is when publishing a paper. What we learned from the workshop, was the unique role the cards had. In fact they helped bringing out problems and misunderstanding, while the empathy was changing in the discussions from a cognitive perspective, where the experience gained over the years when in touch with researchers was relevant, to an emotional one, where first-hand experience in the process of participating with the researchers in the process of publishing Open Access was an emotional new experience.

VIII. Conclusion

The paper has presented an overview where the use of different types of empathy in the pluralistic perspective of the design process seems fruitful. Firstly, it gives an overview of this attribute in regards to the theoretical discourse of Designerly Thinking, secondly it address also the necessity to understand how different types of empathy work during different design effort. The table used in the paper, can be a valuable tool when addressing the correct type of empathy used in different design situations.

The use of emotional and cognitive empathy in the design process needs to be addressed by the research community to better understand how it can be used to gain a more adequate user insight.

Acknowledgment

The authors are indebted to all workshop participants, and the University of Oslo Library for initiating the project.

References

[1] R. Buchanan, “Wicked Problems in Design Thinking,” Design Issues, vol. 8, no. 2, 1992, pp. 5-21.

[2] H. Collins, “Can Design Thinking Still Add Value?,” Design Management Review, vol. 24, no. 2, 2013, pp. 35-39.

[3] K. McCullagh, “Stepping Up: Beyond Design Thinking,” Design Management Review, vol. 24, no. 2, 2013, pp. 32-34.

[4] B. Nussbaum, Design Thinking is a Failed Experiment. So What’s Next? 2011. http://www.fastcodesign.com/1663558/design-thinking-is-a-failed-experiment-so-whats-next (last access 10. February 2015).

[5] J. Woudhuysen, “The Next Trend in Design,” Design Management Journal, vol. 6, no. 1, 2011, pp. 27-39.

[6] U. Johansson-Sköldberg, J. Woodilla, M. Çetinkaya, “Design Thinking: Past, Present and Possible Futures,” Creativity and Innovation Management, vol. 22, no. 2, 2013, pp. 121-146.

[7] T. Brown, Change by design: how design thinking transforms organizations and inspires innovation. Harper Business, New York, USA, 2009.

[8] T. Brown, and J. Wyatt, “Design Thinking for Social Innovation (SSIR),” Stanford Social Innovation Review, vol. 30, no. 5, 2010, pp. 29-53.

[9] R. L. Martin and K. Christensen, Rotman on design: the best on design thinking from Rotman magazine. Toronto [Ont.]: Rotman-UTP Publishing, 2013.

[10] A. Gasparini, “The Value of Emapthy in Desing Thinking,” in Innovation in HCI: What can we learn from Design Thinking? Workshop. In conjunction with NORDICHI 2014, Helsinki, Finland, 2014.

[11] E. Stolterman, “The Nature of Design Practice and Implications for Interaction Design Research,” Int. J. Dsign, vol. 2, no. 1, 2008, pp. 55-65.

[12] E. Spencer, The principles of psychology. Williams and Norgate, London, UK, 1881.

[13] S. New, and L. Kimbell, “Chimps, Designers, Consultants and Empathy: A “Theory of Mind” for Service Design,” Proc. Cambridge Academic Design Management Conference (CADMC 2013).

[14] R. Hogan, “Development of an empathy scale,” Journal of Consulting and Clinical Psychology, vol. 33, no. 3, 1969.

[15] I. Koskinen, K. Battarbee, T. Mattelmäki, Empathic design: user experience in product design. IT Press, Finland, 2003.

[16] M. Peeters, C. Megems, C. Hummels, A. Brombacher, “Experiential Probes: probing for emerging behavior patterns in everyday life,” Proc. International Congress of the International Association of Societies of Design Research (IASDR 2013).

[17] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, T.W. Malone, . Evidence for a Collective Intelligence Factor in the Performance of Human Groups. Science, vol. 330.6004, 2010, pp. 686-688.

[18] G. L. Kress, and M. Schar, “The Art and Science of Design Team Formation,” in Design Thinking Research, H. Plattner, C. Meinel, L. Leifer, Eds. Berlin: Springer Heidelberg, pp. 189-209, 2012.

[19] A. Polaine, L. Løvlie, B. Reason, Service design: From insight to implementation. Brooklyn, N.Y.: Rosenfeld Media, 2013.

[20] A. Culén and A. A. Gasparini, “Find a Book! Unpacking Customer Journeys at Academic Library,” in, The Seventh International Conference on Advances in Computer-Human Interactions, (ACHI 2014) IARIA, feb. 2014, pp. 89-95.

[21] A. L. Culén and A. A. Gasparini, “Student Driven Innovation: Designing University Library Services,” presented at the CENTRIC 2013, The Sixth International Conference on Advances in Human oriented and Personalized Mechanisms, Technologies, and Services, 2013, pp. 12-17.

[22] N. Cross, Design thinking: understanding how designers think and work. Berg, Oxford, UK, 2011.

[23] D. A. Schön, The reflective practitioner: how professionals think and work in action. Basic Books, New York, USA, 1983.

[24] S. Finken, A. Culén, A. Gasparini, "Nurturing Creativity: Assemblages in HCI Design Practices," Proc. Design Research Society, Umeå, Sweden (DRS 2014) pp. 1204-1217.

[25] P. Wright and J. McCarthy, "Empathy and Experience in HCI," Proc. of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2008, pp. 637-646.

Modified Betweenness Centrality to Identify Relay Nodes in Data Networks

Masaaki Miyashita and Norihiko Shinomiya,
Graduate School of Engineering,
Soka University,
Tokyo, Japan
Email: shinomi@ieee.org

***Abstract*—Several types of data networks require relay nodes to transmit data because the nodes would impact on services with the networks. Betweenness centrality is one of the measures that reveal important nodes for a network topology. However, the measure does not specialized in relay nodes of survivable networks. This paper proposes a evaluation method to designate the relay nodes in the network modifying betweenness centrality for relay nodes and survivability. In our simulation, we used two routing algorithms for the survivability of the network and compared the modified measure applying each algorithm with original betweenness centrality. The simulation results show that our approach estimates features different from the original does. The different implies that our method is effective to identify relay nodes.**

Keywords*–*Betweenness centrality; Graph theory; Data networks

I. Introduction

Several types of data networks including a wireless ad-hoc network require relay nodes to transmit data because the nodes would impact on services with the networks. One of the measures to reveal important nodes like the relay nodes is centrality which belongs to graph theory.

Centrality is known as a measure to capture characteristics of network topology in fields of social network, computer science, physics and biology [1], which, in detail, analyzes the importance of each node from some points of view. The importance is rated by a real-valued function on a node. Above all, betweenness centrality is applied to the network in which something such as packets or messages of the Internet flows between other nodes. For each node, betweenness centrality counts the number of shortest paths that the node lies on. Betweenness centrality assumes that every pair of nodes interchanges a message with equal probability in equal time intervals [1]. However, this assumption can be unsuitable for several types of networks. Therefore, Freeman et al. [2] suggests other betweenness centrality with not shortest paths but max flow.

Though a relay node is considered to be located closer to the middle of a route along which two nodes exchange data each other, suggested betweenness centrality may not be able to recognize such. In addition, the original betweenness centrality does not take consideration on survivability because of single path connecting two nodes.

Consequently, our study proposes the modified betweenness centrality based on two node disjoint paths to identify relay nodes in the data network. The measure is expected to give a ranking on the possibility that each node is to be a relay one of a network.

Section II describes the definition of betweenness centrality. Section III presents the modification of betweenness centrality. Simulation results are presented in Section IV. Section V concludes this paper.

II. Betweenness Centrality

According to Newman [1], betweenness centrality is defined as follows.

In a graph denoted as $G = (V, E)$, let σ_{st} be the number of shortest paths between two vertices $s, t \in V$, and let $\sigma_{st}(v)$ indicate the number of the paths through a vertex $v \in V$. Betweenness centrality C_B on v is defined by

$$C_B(v) = \sum_{\substack{s,t \in V \setminus \{v\} \\ s \neq t}} \frac{\sigma_{st}(v)}{\sigma_{st}}. \tag{1}$$

This formula supposes that if there exists multiple shortest paths connecting s and t, i.e. $\sigma_{st} \geq 2$, then one of the paths is chosen equally, so that $\sigma_{st}(v)$ is divided by σ_{st}.

III. Modification of Betweenness Centrality

This section shows the modification of betweenness centrality. A network are described as an undirected edge-weighted graph $G = (V, E)$. Let $w(e) \in \mathbb{R}^+$ be the weight of an edge $e = (u, v) \in E$. Denote $v \in p$ if a vertex $v \in V$ is included in a path p.

Because network flow relayed on a node could be regarded as several paths across the node which some data flows along, a path set is modeled as the flow. Considering endpoints and survivability. Since the graph G is undirected, it is unnecessary to distinguish the source and the destination of a path. Additionally, Not all nodes can send and receive data depending on the network. As a result, let a subset $V' \subseteq V$ include all endpoints contained in the graph.

On the other hand, to guarantee survivability of the network, the graph G is required to be 2-vertex connected. Thus, endpoints are certainly connected by two vertex disjoint paths. Denote $\psi = (p, p')$ as a pair of two paths p and p' that are composed of the same endpoints. Depending on the way to find a path pair of a couple of endpoints and the topology of the graph G, no pair may be found. Therefore, let Ψ_S be a set of ψ able to be discovered between two endpoints s and t, where a set $S = \{s, t\}$ $(s, t \in V', s \neq t)$.

Let $l(p) \in \mathbb{R}^+$ be the length of a path p which is the sum of the weights on edges in the path. Let $d_p(v)$ indicate the distance from either of endpoints of a path to a vertex v. Transmission distance of the network is limited to the upper bound $L \in \mathbb{R}^+$. If $l(p)$ holds $L < l(p) \leq 2L$, a relay vertex should belong to the set $\{v \in p \mid l(p) - L \leq d_p(v) \leq L\}$ to intercommunicate between the endpoints of the path. If $l(p) \leq L$, then p does not need the vertex. If $l(p) > 2L$, then at least

 ISBN: 978-1-61208-382-7

two relay vertices are necessary, so this paper assumes that $L \geq l(p)/2$, for all p for simplicity.

It is considered that the closeness of a vertex to the center of a path corresponds to the appropriateness of the vertex for the relay one. Thus, Define the function $f : [0,1] \to [0,1]$ that satisfies the following:

- $f\left(\frac{1}{2}\right) = 1$
- $f(x) = f(1-x)$
- $f(x) < f(x'), \quad \text{if } x < x' \leq \frac{1}{2} \text{ or } \frac{1}{2} < x' < x.$

The degree of the capability that a vertex $v \in V$ acts as the relay vertex on either path in a pair ψ is designated as the function $\delta_\psi : V \to [0,1]$ on v with the function f

$$\delta_\psi(v) = \begin{cases} f\left(\frac{d_p(v)}{l(p)}\right), & \text{if } \exists p \in \psi \text{ s.t. } v \in p, \\ & l(p) - L \leq d_p(v) \leq L \\ 0, & \text{otherwise} \end{cases}.$$

From the above, modified betweenness centrality C_M as a function on a vertex v is defined by

$$C_M(v) = \sum_{\substack{S \in \mathcal{P}(V' \setminus \{v\}) \\ |S|=2}} \frac{\sum_{\psi \in \Psi_S} \delta_\psi(v)}{|\Psi_S|}. \tag{2}$$

IV. Experiments and Results

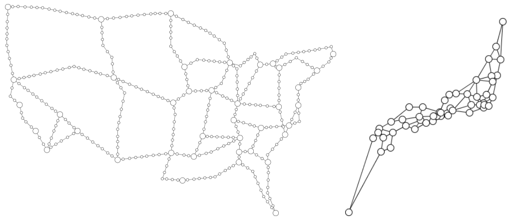
Two network models were analyzed by betweenness centrality and modified one. The results from each measures were compared.

Modified betweenness centrality is made with Constrained Shortest Path First (CSPF) and Vertex Disjoint Shortest Pair (VDSP) for the pair routing algorithm. Simplified algorithm of CSPF is given as follows:

1. Find a shortest path by Dijkstra algorithm in a graph.
2. Modify the graph to remove vertices in the path without endpoints.
3. Search the other shortest path between the same endpoints in the modified graph.
4. Obtain the pair of first path and second path.

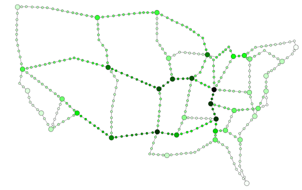
VDSP proposed by Bhandari [3] is an algorithm for finding two paths that have minimum total length. It enables to get the vertex disjoint pair in any 2-vertex connected graph.

Figure 1 shows network models for the evaluation of analysis methods. Model A in the figure limits endpoints sets, and model B does not. Model A possess 365 vertices. The



(a) Model A (b) Model B

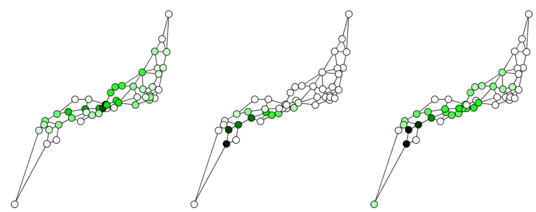
Figure 1. Two types of graph for experiments.



(a) Original betweenness centrality

(b) Modified with CSPF (c) Modified with VDSP

Figure 2. Visualized evaluates on model A.



(a) Original (b) with CSPF (c) with VDSP

Figure 3. Visualized evaluates on model B.

endpoints set V' is composed of big circles illustrated in the figure, of which the size is 39. Model B contains 49 vertices. Its endpoints set is the same set as own vertex set.

Figures 2 and 3 show visualization of evaluation results for the models by coloring each vertex the color of which brightness corresponds to each evaluate value of the vertex (palest = 0, darkest = max). These figures describe that locations of high value vertices in both of the modified measures are different from original measures on both model A and B. From this result, each modified centralities can reveal characteristics which existent betweenness centrality can not.

V. Conclusion

This paper proposed a method to identify relay nodes based on modified betweenness centrality. This method is constructed of a set of pairs of vertex disjoint paths, and the degree of closeness to the center of each path. In our simulation, the two types of modified measures by two algorithms CSPF and VDSP to determine a set of pairs were compared with original betweenness centrality in two different network models. The simulation results show that our approach indicates attributes the original does not. Thus, our method may be capable of designating relay nodes.

References

[1] M. Newman, Networks: an introduction. Oxford University Press, 2010.

[2] L. C. Freeman, S. P. Borgatti, and D. R. White, "Centrality in valued graphs: A measure of betweenness based on network flow," Social networks, vol. 13, no. 2, 1991, pp. 141–154.

[3] R. Bhandari, Survivable networks: algorithms for diverse routing. Springer, 1999.

User Interface Development of a COPD Remote Monitoring Application

A User-centred Design Process

Berglind Smaradottir, Martin Gerdes and Rune Fensli
Department of Information and Communication Technology
University of Agder
N-4604 Kristiansand, Norway
{berglind.smaradottir, martin.gerdes, rune.fensli}@uia.no

Santiago Martinez
Department of Psychosocial Health
University of Agder
N-4604 Kristiansand, Norway
santiago.martinez@uia.no

***Abstract*—The Norwegian Health sector is undergoing changes at all levels due to recent health reforms. Services traditionally offered by specialized health care are being transferred to primary health care managed by municipalities. In this context, telemedicine technology is introduced to facilitate new services that support communication across local borders, optimizing resources and increasing cost effectiveness. This study focuses on the user-centred design, iterative development and evaluation of the user interface of a mobile application to be used in a new telemedicine service for remote monitoring of chronic obstructive pulmonary disease symptoms. The research is a result of the cooperation between a municipality and a hospital as part of the European Commission project United4Health. Through a user-centred design approach, the tablet device application was developed based on information gathered in a workshop and group interviews where the end-users, patients and health professionals, described their preferred way of interacting with the telemedicine technology. User evaluations reported positive results on the ease of use and user satisfaction with the interaction with the application. Iterative application's user interface refinements were made through several end-users' evaluations, resulting in a fully developed system suitable for remote monitoring.**

Keywords-remote monitoring; patient empowerment; usability evaluation; telemedicine; user-centred design.

I. INTRODUCTION

The Norwegian Coordination Reform [1] urged health organisations to implement structural changes as the ultimate citizens' health care providers. Services traditionally offered by specialized health care national and regional institutions (e.g., follow up of chronic diseases managed by hospitals) were transferred to primary health care managed by municipalities. This new situation brought to light the need for an effective coordination and improved communication across borders of health care services.

Services associated with the new patient pathways required from municipalities an effective and efficient use of Information and Communication Technologies (ICTs). In this context, the European Commission funded research project United4Health [2] is developing technology for remote monitoring of chronic diseases and communication across the different levels of health care services. The Norwegian contribution to the United4Health project focuses on technologies that support remote monitoring of chronic obstructive pulmonary disease (COPD) patients after hospital discharge. The aim of the project is to evaluate the benefits of using technology for monitoring COPD patients that traditionally did not have the possibility of reporting their symptoms and health status after hospitalisation. Potential benefits would include reduction of hospital readmission rates with their correspondent diminution in cost, and benefits of quality of life improvements (already being investigated in other ongoing research from the same project). Research evidence shows that COPD patients are at an increased risk of readmission to hospital within 12 months [3][4] after hospital discharge.

In this study, a mobile telemedicine application was developed in a tablet device for remote monitoring of blood oxygen saturation (SpO_2) and pulse measurements. In addition, the application contained a questionnaire for daily report of COPD symptoms. Patients took measurements at home that were wirelessly transmitted to a newly established telemedicine centre assigned by a municipality partner where health professionals would remotely attend the patient.

A user-centred design (UCD) process was employed for the development and evaluation of the mobile telemedicine application. The application was designed with the active involvement of end-users from patient's union of cardiac and pulmonary patients, health professionals from the municipality and partner hospital. The study was led by a research group with ICT and health background. The application was validated from an operational and qualitative usability aspect. The research questions (RQs) of this study were:

RQ1: "How can a telemedicine application be developed for remote home monitoring purpose, including COPD patients and disease-related health professionals?"

RQ2: "What lessons from this study are transferable and applicable for the development of technology useful for other clinical pathways?"

Following this introduction, Section II gives an overview of research background about UCD and Section III outlines the research methodology employed. Section IV describes the results of the mobile application development. In Section V, the results are discussed and in Section VI the conclusion and future work are presented.

 ISBN: 978-1-61208-382-7

II. Research Background

UCD involves end-users in all the stages of a system's development [5][6][7]. It helps to understand users' needs and the context of use, which are key elements for the construction of a system framed within a clinical workflow [8]. In addition, the usability evaluation is necessary to analyse user's interaction and user satisfaction with the system [9][10][11]. Telemedicine systems often involve the interaction between multiple user groups through a system, e.g., a patient at home communicates using a device with nurse in telemedicine or health centre, or with general practitioner (GP) at his office. Communication in these scenarios of use is usually multimodal, that is, synchronous (e.g., videoconference) and asynchronous (e.g., data transmission and dispatch), what makes it crucial to know between whom, how and when the information transmission and personal contacts occur. Thus, an effective telemedicine application requires a detailed analysis of end-users' needs to inform system designers. In addition, the usability of such application is crucial for the continuous, efficient and satisfactory use of an application [7].

III. Methodology

The design of the telemedicine tablet application was performed as a part of the research project United4Health [2]. Qualitative methods were used for data collection and analysis. The UCD process was divided into two phases, as illustrated in Figure 1: (A) workshop with representative end-users, such as patients and health professionals; (B) iterative design of the tablet application for COPD remote monitoring. The latter was formed by a set of four sub-phases: design and implementation, functional test, user evaluation, and field trial. Each sub-phase's output informed the input of the next. The iterative system development included a sequence of concatenated stages where user requirements informed the design, implementation and functional test of the application.

Running commentary gathered in the two phases of the UCD process resulted in 18 hours of audio-visually recorded data, verbatim transcribed by the researchers. Transcripts were coded into categories and a qualitative content analysis [11] was made with the software QSR NVIVO v10.

A. *Workshop with End-users*

A one-day workshop with end-user representatives (i.e., health professionals and patients) was set up in October 2013 hosted by the University of Agder, Norway. The aim was to understand the context of use and to gather user requirements for the design of the tablet application for remote monitoring. In addition, the workshop was a source of information and familiarisation for end-users with the research team and health personnel working in the project. The participants were two members of the union of cardiac and pulmonary patients, mean age of 69 years; two nurses from the municipality and hospital, mean clinical experience of 6 years with COPD patients; and two technicians from hospital responsible for correct functioning and maintenance of the tablet devices, with a mean of 6 years of experience working

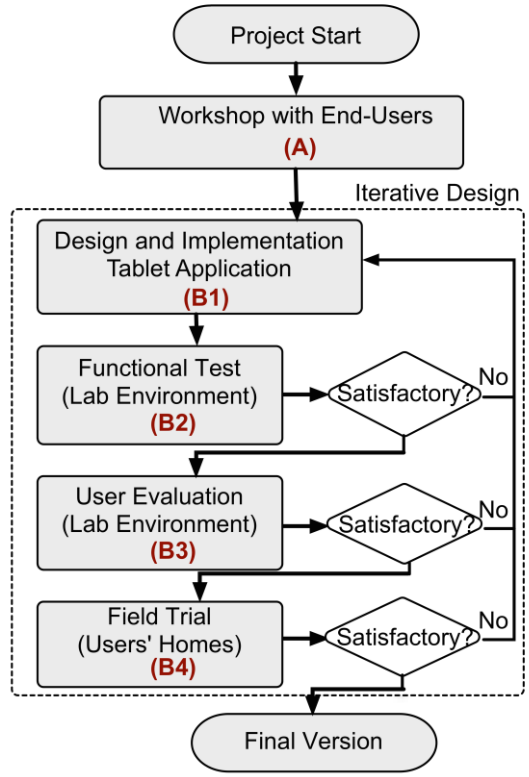


Figure 1. The User-Centred Design Process.

with medical technical equipment.

The workshop lasted 5 hours and was divided into two parts. In the first part of the workshop, participants were given an introduction to the research project United4Health. In order to understand the context of use of the system, a prototype demonstration of wirelessly transmitted measurements of SpO_2 and pulse was shown to end-users on a tablet. Additionally, a video-conference (software Cisco Jabber Video for Telepresence, v4.2) between a patient and a health care professional was tested. COPD patient representatives described their preferred way of interacting with the application at home and suggested ideas for the user interface's (UI) layout. Participants used colourful post-it notes and handmade sketches to describe application's functionalities and design.

In the second part of the workshop, participants described the workflow of remote monitoring of a COPD patient, such as taking measurements at home, transmitting measurements' values through the system to the telemedicine centre and illustrating the feedback given from telemedicine centre to COPD patient at home.

B. *Iterative Design*

The design of the application was carried out through the iterative execution of the following stages: design and implementation, functional test, user evaluation and field trial. A development team supervised by one of the researchers developed the system. An interaction designer hired by the team was in charge of the initial graphical user interface and interaction design.

 ISBN: 978-1-61208-382-7

1) Design and Implementation: The results from the workshop led the initial design and implementation of a Java native application. Java includes libraries for several low-level Application Program Interfaces (APIs), in particular for the Bluetooth connectivity and communication with sensor devices. In addition, using Java allowed the application to be used across different tablet devices. The outcome of the latter sub-phases informed additional user requirements included in the implementation of the User Interface Design (UID) and system's functionality.

2) Functional Test: The facilities of the Centre for eHealth and Healthcare Technology of the University of Agder, Norway, were used as a test bed for a functional test of the implemented application. It allowed to verify whether the system matched the requested functionality determined by users in the workshop and in user evaluations from other iterations.

3) User Evaluation: Two evaluations of the application's prototype were carried out with end users in the Usability Laboratory at the Centre for eHealth and Healthcare Technology, in January and February 2014. The facilities had two separate test rooms (referred to as "test room 1" and "test room 2") and one observation room. The infrastructure is further described in [12]. The user evaluations had the aim to provide end-user's feedback to the development team about system's errors and potential refinements. They consisted of a series of tasks using a think aloud protocol [9][10]. Group interviews were made at the end of the evaluations to complete the feedback.

a) Evaluation 1: 15 health care professionals from the municipality and hospital partner constructed a role-play scenario. In test room 1, which represented patient's home, health care professionals simulated the patient's use of tablet application and, at the same time, the interaction with the health care professionals placed in test room 2, which represented the telemedicine centre. The functionalities tested at a patient's home consisted of taking and sending patient's measurements (i.e., SpO_2 and pulse) and filling and sending a questionnaire to the telemedicine centre. In addition, a videoconference session between the patient and the telemedicine centre was performed and evaluated. There were three repetitions of the scenario with different users and the overall duration of the evaluation was 6 hours.

b) Evaluation 2: The evaluation was performed as a role-play with a simulation of the proposed use scenario of the new telemedicine application. It was carried out two weeks after evaluation 1 and included 9 end-users: 2 members of the patient's union (played the patient's role), 3 nurses from municipality (played telemedicine centre health professional's role), 2 nurses from hospital and 2 technicians from hospital. The test simulated the following interactions with the application: 1) user training of COPD patient in hospital with instructions from a hospital nurse; 2) COPD patient at home taking measurements, filling in questionnaire and sending it to the telemedicine centre; 3) videoconference between COPD patient at home and a health professional at the telemedicine centre. There were two iterations of the user evaluation, with a total duration of 5 hours.

4) Field Trial: A field trial was carried out with 6 diagnosed COPD patients (mean age 72.6 years). They tested the continuous functioning and interaction with the technology at home during a period of 7 days. The trial lasted 5 weeks. Each user made daily measurements, filled in a questionnaire and sent them to the telemedicine centre. In addition, videoconference between a user and a health professional at the telemedicine centre was tested. All these tasks were performed using a tablet device.

After each week of testing, user suggestions were incorporated in the improvement of the system.

IV. RESULTS

The results were obtained from transcripts of the audio-visually recorded data and annotations and observations during the UCD process. To ease the reading, the results of each phase are separately presented.

A. Workshop with End-users

The contributions from end-users in the workshop are grouped in 3 different categories: context of use, UID and patient workflow.

1) Context of Use: Patient representatives explained that individual's level of physical energy is regularly low and even simple actions, such as using a tablet device, may seem unachievable. This issue underlined the importance of designing an easy-to-use application that did not require much effort to be successfully used. Therefore, it was suggested that user interaction with the system must be minimal, with only the few necessary actions. One of them stated: "Usability is extremely important for the interaction with this application since COPD patients have little energy left on bad days".

2) User Interface Design: Patients agreed with the authentication method through a Personal Identification Number (PIN) mechanism, although they expressed having difficulties remembering numbers and they preferred to be able to choose their own PIN instead of using a pre-defined one. In addition, they requested to have the user's name at the top of the home screen after successful login.

Patients required seeing the results of their own measurements on the device's screen before sending them to the telemedicine centre. In addition, they asked for receiving immediate feedback when measurements were successfully delivered.

A time-span visualization of several days of measurement results was also suggested where patients could see measurements from previous days.

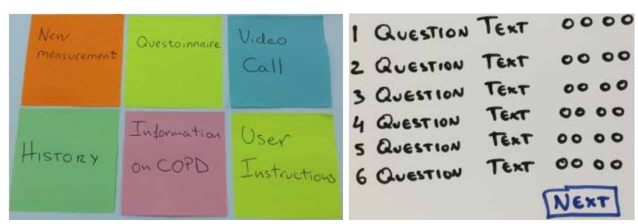


Figure 2. User's UI suggestions (left) and questionnaire's sketch (right).

Another request was the possibility of simultaneously seeing videoconference professional was required to guide the patient through any of the tasks.

For the interface's layout, patients chose not to have nested menus (e.g., one patient representative said: "you cannot ask elderly people to remember what is inside each menu") and instead, one touch area per action. Suggestions included 6 squared big-size touch areas, with readable and appropriate function's names. The 3 most important functions were placed at the top: "new measurements", "daily questionnaire" and "videoconference". The other 3 touch areas with less frequently used functions were placed at the bottom: "historical data", "information about COPD", and "user instructions", see Figure 2, left.

Further, it was concluded that the system was not to be used for emergency situations, so a written text saying "Call 113 for emergency" was required to be shown in the system's interface.

The answer to the daily questionnaire was suggested to be of multiple touchable selections (see Figure 2, right), and to have six questions visible on the screen at the same time because patients were afraid they would get tired of reading the questions one by one.

3) Patient Workflow: One of the most important findings of the workshop was the setup of the COPD patient's workflow for the use of the telemedicine application for remote monitoring, see Figure 3. In addition, instructions were required to be concise, to be on paper and digitally available in the system.

It is a common practice in the telemedicine centre to differentiate patient status by an easy-to-interpret colour scheme, called triage. Triage colour was represented in this case by a green colour for measurement values within the pre-defined cut-off values; and yellow and red ones for attention and alert respectively, activated when measurement values are outside the cut-off values. Patient representatives initially suggested that patients at home should be able to see the triage colour related to their own measurements in order to have a feeling of control of their own health. However, a "false" red measurement (e.g., cold finger can alter measurement readings) could potentially increase patient's anxiety. At the end, patient representatives agreed with the option that only health care professionals could see the triage's colour.

B. Iterative Design

1) Design and Implementation: In the sub-phase Design and Implementation, workshop's results were transformed into the initial Graphical User Interface (GUI) outline, see Figure 4, and user requirements of the application. Outcomes from further iterations' sub-phases contributed to refine the user requirements and improve the application implementation.

2) Functional Test: In each of the iterations, the application had to go through functionality test run by the development team. The identification of errors at this stage

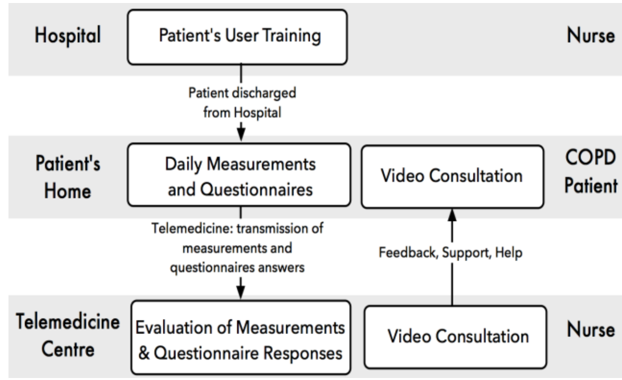


Figure 3. COPD patient workflow in remote monitoring.

proved to be relatively cost-effective to fix in terms of time and effort compared with further sub-phases.

3) User Evaluation: User evaluations in laboratory settings concerned the display of the questionnaire with the adequate number of questions per screen, from 6 to finally 1 to ease individual reading. A questionnaire's answers review was included to allow patient to double check before sending. Initially, a progress bar notified data transmission but it was insufficient for distinguishing between successful and unsuccessful data delivery. Therefore, a feedback notification pop-up window was included with text, a round face and a status colour code (green smiley face for successful delivery and red sad face for unsuccessful one).

In addition, user manual needed to be incorporated in the system, with intuitive images to guide on how to handle the measurement devices step-by-step. In this line, the GUI corresponding to the new measurement was improved by removing the unnecessary information load to perform the task. For instance, while measurement reader device showed correct measured values, wrong ones were displayed in the tablet screen and sent to the telemedicine centre. User evaluation helped to identify this issue.

In the group interviews, user comments about the tablet use were overall positive: "I think this will help us if we get worse; the tablet was easy to use with 5 or 6 functions and few things that should be touched to do measurements".

Some comments referred to the need of user training: "With some user training I think most people could use this, it was not complicated. If you forget how to do it, you can contact telemedicine centre". Patients also positively commented about the videoconference: "It was a good feeling to have videoconference with telemedicine centre. I think for users at home, it is good to see and hear the nurse".

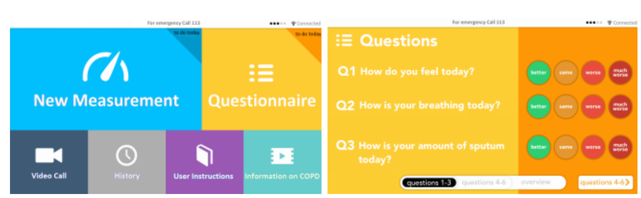


Figure 4. GUI of tablet application and the questionnaire.

4) Field Trial: Field trial results included automatic start of the application due to problems with touch initiation of the program icon (equivalent to mouse double-click). It was found that, ideally, the tablet application should report the battery level of the measurement device to the telemedicine centre and user. The videoconference image and sound quality was improved through software configuration changes. Furthermore, the sound quality was improved by the selection of optimal headphones and microphone setup for the users. All participants gave positive feedback about the printed user manual. In particular, they stated that the text was easy to understand and the pictures accurately illustrated what to do for each task.

Users' overall rating of the application was satisfactory concerning all interactions with tablet (equipment setup, device connection, measurements, questionnaire filling, data transmission, and videoconference): "I think the application is very well designed so you do not misunderstand anything. I consider this system user-friendly"; "This application was easy to use, because even an old person like me without computer experience could use it".

C. Final Version

The UCD process resulted in a tablet application that was evaluated as "satisfactory" in all sub-phases. A number of exemplary screen shots of the application are shown in Figure 5. The UI screenshot on top shows the main screen of the application, from where the user can initiate the main functions, such as "new measurements" and "questionnaire". The blue screen on the bottom left shows the start screen for carrying out new SpO_2 measurement, while the orange screen on the bottom right shows parts of the UI for the daily questionnaire. This application is being deployed on the trial devices of the Norwegian part of the United4Health project, and will support the new remote monitoring services provided by the municipalities involved.

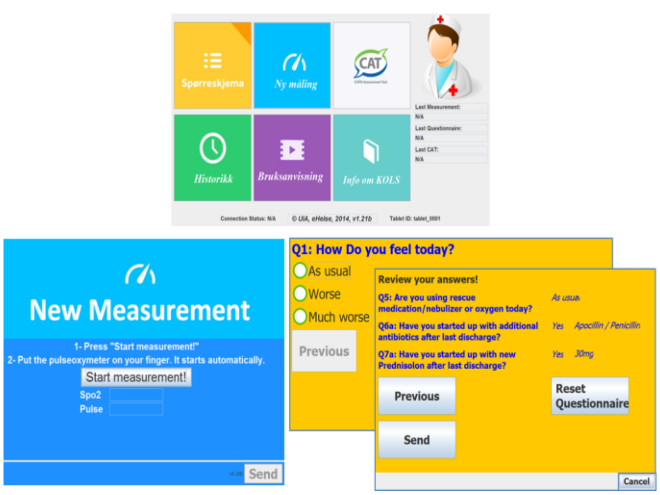


Figure 5. Screenshots of the system's final implementation.

V. Discussion

In this paper, the UCD process for the development of a tablet device application for remote monitoring of patients in home environment has been presented. A typical telemedicine application involves multiple users in number and type, such as patients, health professionals and administrative officers. This is why the involvement of those groups of end-users in the design of a new technical application is crucial to understand the clinical workflow where the solution will be deployed, its context of use and the interactions involved. The two research questions (RQs) formulated at the beginning of this paper are answered below based on the results from the study.

About the RQ1, which asked about the development of a telemedicine application for remote monitoring, it has been demonstrated that a UCD approach successfully included the user (i.e., COPD patients and health professionals) needs in the development of the application. For it, a workshop efficiently outlined user needs, context of use, and helped user groups to familiarise themselves with each other and the research team. In addition, the workshop established a useful starting point for the system's application development taking on board different kinds of user requirements, as aspects of GUI, interaction and functionalities.

The user evaluation was carried out both in a laboratory environment and at users' homes. The early evaluations in laboratory environment simulated a realistic user scenario in a controlled test environment, and enabled users to give feedback about GUI design and the interactions following the remote monitoring process. The field trial allowed studying the long-term and real-time usage of the technology by users at their home.

Several lessons were learned during the study that can be transferable for technology development for other clinical pathways (RQ2) involving chronic conditions included in the projections of global mortality for 2030, such as ischaemic heart disease and diabetes [13]. In particular, solutions to be installed in medical environments necessarily need to firstly involve all the user groups in the creation of the solution, and secondly, analyse how this solution can best fit in an existing clinical workflow or, if non-existent, build up such workflow in collaboration with the end-users. It is known that interoperability issues are one of the main problems nowadays in the deployment and use of technology in medical environments. In this way, the inclusion of a field trial provided useful information about the interactions between humans and technology, but also between the different technologies involved.

The research study of the UCD process had also some limitations such as a reduced number of end-users and user-scenarios were tested in a simulated environment. However, the simulated test environment allowed to create highly realistic scenarios under controlled conditions, and the field trial gave the opportunity to test the system in real-world settings.

 ISBN: 978-1-61208-382-7

VI. CONCLUSION AND FUTURE WORK

This study has been developed including end-users' (i.e., COPD patients and health professionals) needs, suggestions and preferences, in the design and evaluation of a COPD remote monitoring application. Positive results were reported after evaluation in laboratory settings, regarding ease of use of the telemedicine solution and user satisfaction. The continuous refinement of the application was the key to fully develop the system suitable for remote monitoring of COPD patients.

The benefits of giving the opportunity to COPD patients to report symptoms and health status after hospitalisation, together with actively including them in building the solution, are in line with the European Union (EU) Health Strategy, *"putting patients at the heart of the system and encouraging them to be involved in managing their own healthcare needs"* [14].

These facts, together with the simulation in high fidelity laboratory settings, and the field trial are significant contributing factors to the ecological validity of the research here presented. In a world where human-computer interactions progressively increase in number and complexity, real-time evaluations in real-world settings become crucial to understand not only the successful deployment, but the efficient and continuous use of technological solutions.

The proposed UCD process has been validated by the development of a telemedicine tablet application, successfully adopted by the 7th Framework Programme for Research and Technological Development (FP7) EU project United4Health, which focused on technologies that support remote monitoring of COPD patients after hospital discharge. As a result, over 200 patients in various municipalities in Norway will use the application.

In terms of future work, it is proposed to address research on appropriate identification and authentication methods for patients, more autonomous reasoning and decision support in the application, and integration of further devices to support other patient groups and clinical pathways associated with chronic diseases, such as hypertension and diabetes.

ACKNOWLEDGMENTS

The authors thank all participants and project partners for their contributions in the UCD process. We thank Joris-Jan van den Boom for contribution in graphic and interaction design and Baddam Sriramreddy, Mahdi Mahdavi Amjad, Navid Mokhtari, Saeed Zanderahimi, Volodymyr Shulgin and Yannick Schillinger in the development team.

This work was supported by UNIversal solutions in Telemedicine Deployment for European HEALTH care, 2013-2015 ICT PSP call identifier: CIP-ICT PSP-2012-3 and Point-of-Care Services Agder, sub-project financed by the Research Council of Norway, 2013-15, ref 227131/O70.

REFERENCES

[1] Norwegian Ministry of Health and Care Services, "The Coordination Reform, Proper treatment – at the right place and right time," Report No. 47 (2008-2009) to the Storting. [retrieved: January, 2015]. Available from: http://www.regjeringen.no/upload/HOD/Dokumenter%20INFO/Samhandling%20engelsk_PDFS.pdf.

[2] United4Health 2015, European Commission Competitiveness Innovation Programme, ICT Programme to Support Policy. [retrieved: January, 2015]. Available from: http://www.united4health.eu/ (the umbrella project) and http://www.united4health.no/ (the Norwegian project contribution).

[3] G. Gudmundsson et al., "Risk factors for rehospitalisation in COPD: role of health status, anxiety and depression," Eur Respir J, vol. 26, 2005, pp. 414-419, doi:10.1183/09031936.05.00078504.

[4] I. M. Osman, D. J. Godden, J. A. Friend, J. S. Legge and J. G. Douglas, "Quality of life and hospital re-admission in patients with chronic obstructive pulmonary disease," Thorax, vol. 52, Jan 1997, pp. 67-71, doi:10.1136/thx.52.1.67.

[5] Y. Rogers, H. Sharp and J. Preece, Interaction Design-Beyond Human-Computer Interaction. John Wiley & Sons; 2011.

[6] J. Lazar, Web Usability- a user-centered design approach. Pearson Education, 2006.

[7] J. Nielsen, Usability engineering. Elsevier, 1994.

[8] A. De Vito Dabbs et al., "User-centered design and interactive health technologies for patients," Comput Inform Nurs, vol. 27(3), May-June 2009, pp. 175–183, doi: 10.1097/NCN.0b013e31819f7c7c.

[9] A. W. Kushniruk and W. L. Patel, "Cognitive and usability engineering methods for the evaluation of clinical information systems," J Biomed Inform, vol. 37, Feb 2004, pp. 56-76.

[10] M. W Jaspers, "A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence," Int J Med Inform, vol. 78(5), 2009, pp. 340-353.

[11] J. Lazar, J. H. Feng and H. Hochheiser. Research methods in human-computer interaction. John Wiley & Sons, 2010.

[12] M. Gerdes, B. Smaradottir and R. Fensli, "End-to-end infrastructure for usability evaluation of eHealth applications and services," Scandinavian Conference on Health Informatics (SHI 2014) Aug 2014; pp. 53-59, ISSN(print): 1650-3686, ISSN(online): 1650-3740

[13] C. D Mathers and D. Loncar, "Projections of global mortality and burden of disease from 2002 to 2030," PLoS Med, vol. 3(11), Nov 2006, pp. 2011-2030, doi: 10.1371/journal.pmed.0030442.

[14] European Union. Official Journal of the European Union, vol. C 146, 22 June 2006, ISSN 1725-2423. [retrieved: January, 2015]. Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:C:2006:146:TOC.

Field Evaluation of a New Railway Dispatching Software

Isabel Schütz
Department of Railway Engineering
TU Darmstadt
Darmstadt, Germany
schuetz@verkehr.tu-darmstadt.de

Anselmo Stelzer
Department of Railway Engineering
TU Darmstadt
Darmstadt, Germany
stelzer@verkehr.tu-darmstadt.de

***Abstract*—In this paper, we present a program which assists in choosing the right evaluation methods. The Test Selection Program contains a four step method for choosing evaluation methods: describing important boundary conditions of the study, defining important evaluation criteria, choosing evaluation methods and deciding on the right order in time. As a practical example, connection dispatching will be introduced and used to illustrate the functionality of this program.**

Keywords-evaluation methods; connection dispatching; usability; railway engineering.

I. INTRODUCTION

The domain of railway engineering is safety-critical, so the main focus is on ensuring safe operations. The secondary focus is on punctuality to increase and maintain customer satisfaction and service quality [1].

When developing safety-critical software, the main focus is mostly on safety. Beyond that, usability nowadays does not play a big role [2].

In the last few years, the focus has been shifting more and more to customer oriented dispatching, for example through good customer information or connection dispatching. Especially in the field of connection dispatching, which we will concentrate on in this paper, suitable software to display the connection status or connection conflicts is not yet known. At the moment, general software is used which poorly supports connection dispatching [5]. To comply with requirements given by contracting traffic authorities, e.g., regional transport authorities, this needs to be changed. Moreover, the importance of intermodal connection dispatching rises and is therefore integrated into traffic contracts, often combined with a contractual penalty [4].

The consideration of intermodal connections increases the dispatchers' workload compared to today's situation. The increasing requirements cannot be met with the contemporary personnel shortage. Furthermore, training dispatchers takes a long time because of the need to gain experience in everyday work. Without well-designed software which supports the dispatcher in doing his job of connection dispatching, this kind of workload cannot be sufficiently dealt with. A clear and efficient way of displaying all connections is important to facilitate handling them simultaneously and also to be able to automate several steps within this process of connection dispatching. This is expected to reduce the dispatcher's workload. For these reasons, a project has been established to support the dispatcher in connection dispatching. Its objective is to support dispatchers by displaying planned connections and connection conflicts together in a concise, newly developed software interface. The project contains a field study for which suitable evaluation methods needed to be found. Since the work of a railway dispatcher is not safety-critical (in contrast to the work of a traffic controller), it offers optimal possibilities to establish usability in the field of Railway Engineering [5]. We will use this project as a practical example to illustrate a new process of choosing the right evaluation methods.

Several other evaluation projects with different partners – mainly expert evaluations and evaluations in a simulation environment – have been conducted [8]. During these projects, the importance of using suitable evaluation methods was pointed out, as was the necessity of reasonable combination in the right order in time.

Since the introduced project's field study will end at the beginning of 2015, in this paper we concentrate on the following issues: How to choose the most suitable evaluation methods and how to combine them reasonably and in the right order in time. For this, in Section II the field of railway dispatching in general, and the process of connection dispatching in particular, are introduced. Then, in Section III, the focus is on the tool we developed to support an optimal choice of suitable evaluation methods. Thereafter, in Section IV, special attention will be paid to the specific example of the field study conducted with dispatchers to show how we chose the evaluation methods and why we combined them in the actual way and sequence we did. We finish our contribution with a conclusion and a description of future work in Section V.

II. RAILWAY DISPATCHING

Dispatchers are well-known in different realms, for example rescue services or logistics. Likewise, in the railway system, we distinguish between different dispatchers. In this paper, we will concentrate on the dispatcher of the train operating company (further referred to as "the dispatcher"). It is his job to coordinate activities, such as passenger information, disposition of rolling stock and personnel, communication, coordination and connection dispatching. He concentrates on maintaining and increasing the customer satisfaction. In doing so, a consideration of the feasibility and the consequences of a performed measure is necessary. The dispatcher – in contrast to a traffic controller – has no

 ISBN: 978-1-61208-382-7

responsibility for safety, as he is not using any interlocking systems [3].

A. Definition of a Connection

A connection is the possibility for the passengers to change from an arriving feeder train to a departing distributor train within a certain interchange time. Possibly, also connections from and to different modes of transport can be considered and are referred to as "intermodal connections" [3].

B. The Process of Connection Dispatching

To travel from one location to another, interchanges between trains are often necessary in a railway network as dense as in Germany. To change from the feeding train to the distributing train under consideration of the physical distance between the platforms and the local conditions inside the station, a certain interchange time and also an additional buffer time have to factored into the planning process of a travel. Normally, this carefully considered time is enough to successfully switch trains without any further work for the dispatcher.

The dispatcher's work in connection dispatching begins when trains are delayed. In this case, a connecting train might not be reached within the above-mentioned predetermined interchange time. In some cases, trains can wait for each other without generating any issues within the so called waiting-time arrangements which permit the distributor a few minutes delay from the original departure time.

If the remaining time between the arrival of the feeding train and the departure of the collecting train under consideration of the waiting time is not sufficient for the passenger to switch trains, the dispatcher has to perform several activities:

He has to decide whether this connection can be reached or not. This decision is often based on his experience gathered about interchanging passengers or his knowledge of the location which allows him to predict that an interchange can be performed faster than assumed initially, e.g., because of the interchange taking place on the same platform. If the dispatcher decides upon securing the connection, he has to confer with the railway infrastructure company and possibly with other train operating companies whether it is possible to let the distributing train wait longer than the initial waiting time. Both can reject keeping this connection when they fear negative consequences along the following route of the train or further connection conflicts at other stations. In the case of having decided not to keep the connection, the dispatcher has to inform the passengers about alternative connections. In both cases, he has to inform the train drivers, his dispatching colleagues and the travellers on the train [3].

C. The Newly Developed Dispatching Software

Nowadays, the process described in the previous section is mainly done by hand. That is why a prototype software has been developed which is intended to support the dispatchers during this process by displaying connections and connection conflicts [7]. The connections are arranged in a matrix with the feeder train on the y axis and the collecting train on the x axis. Connections are presented within the cells of the matrix. Thus, it is possible to see all important information at a glance. By clicking the cell, more detailed information can be retrieved and further actions can be taken [5].

For the visualisation, each connection is assigned to a category – green, yellow or red – comparable to a traffic light. The coloured cell, which also displays additional information concerning the connection, represents this category. It advises the user on how to proceed with a connection (conflict) [5].

The prototype software developed was initially evaluated by experts. Several Focus Groups were hosted to gain input on the interface and its functionality, but also several cognitive walkthroughs were employed on that account. The experts consisted of special dispatchers, who also conduct trainings, and of usability experts. Subsequently, the prototype software was tested in a first user study within a simulation environment close to reality – the so-called Eisenbahnbetriebsfeld Darmstadt (EBD) – with prospective users to confirm the design and its suitability for use [8].

The EBD is a research facility for railway operations which provides the complete chain of railway operation and dispatching. It embodies a realistic simulation environment where tracks and trains are models, but interlocking technology, dispatchers' software and auxiliary equipment, such as phones and walkie-talkies, are real. In this user study, the EBD was used to prove that this prototype software enables the dispatchers to handle a realistic testing scenario. In this way, the prototype software was tested for the first time under realistic conditions and could indeed prove its benefits [9].

Apart from making sure that this prototype software can be used in the field, further improvement of the user interface was to be achieved before testing in the field – using the feedback of the tests in the EBD. Moreover, one major aim of this test was to ensure that all essential functions have been implemented as well as keeping the amount of errors small to prevent users in the field study from being frustrated. Since field testing is comparably expensive, an assurance was necessary that all results are valid and can be used for the analysis and improvement of the prototype software [8].

III. TEST SELECTION PROGRAM

Considering these above-mentioned preliminary studies, we wanted to make sure that the field study and its results are not distorted by malfunctioning prototype software or by user frustration. Also, crucial for the success of a field study is collecting the right evaluation data by using the most suitable evaluation methods. To ensure the choice of suitable evaluation methods and reasonable combinations thereof in the right order in time, a program was developed during a six-month study [6].

The Test Selection Program is designed to assist in choosing the most suitable evaluation methods given a specific evaluation context. It selects suitable methods based on the criteria entered by the user, employing underlying filtering algorithms. Its functionality could be proven in several studies and it is integrated in an ongoing updating process.

All evaluation methods are characterised by a descriptive profile. This profile is the basis for the underlying algorithms choosing the most suitable evaluation methods. As can be seen from Figure 1, it is displayed on the left side of the program mainly comprised of toggle buttons, but also of two toggle button groups with a checkbox each. When clicking a toggle button, it changes its colour to red and stays pushed in to indicate a selected criterion, as can be seen in Figure 2. Clicking the button again releases it and thus de-selects the property. The toggle button groups are first greyed out until the respective checkbox has been selected (see Figure 2). These visual cues allow to easily identify the choices the user has taken so far [6].

On the right hand side (see Figure 1), a list is displayed, arranged in a tree structure. Initially, it contains all evaluation methods that are included in the program. As soon as buttons are clicked, the list reduces by the methods not meeting the selected criteria (see Figure 2). So, after having entered all boundary conditions and criteria of the study, the user can instantly see the most suitable evaluation methods [6].

For further technical or functional details about the program see [6].

IV. CHOICE OF EVALUATION METHODS

The choice of evaluation methods based on the Test Selection Program described in Section III can be divided into four steps:

The first step is to describe the important boundary conditions of the evaluation to take place (cf. Section IV.A). The second step is to define important evaluation criteria (cf. Section IV.B). Thirdly, these boundary conditions and criteria are entered into the program and instantly yield an optimal choice of evaluation methods (cf. Section IV.C). Since in most cases there is more than one suitable evaluation method, the user has to decide on how many evaluation methods to use, how to combine them and also which subtype of a method to use, if applicable. In a fourth step, the right order in time has to be chosen (cf. Section IV.D).

During a six-month research study [6], many scientific references concerning existing methods and procedures were analysed, compared and integrated into one process to choose evaluation methods, and a relevant workflow was defined. It embodies a clear structure and helps to avoid missing an important step. Beyond the original definition of the process, it has been and is constantly being refined during other research projects.

In this paper, we will describe this process based on the specific example of the user interface for displaying connections and connection conflicts in the context of connection dispatching (see Section C).

A. Step 1: Describing Important Boundary Conditions of the Study

To find the optimal combination of evaluation methods, the boundary conditions of the evaluation need to be described first.

For our study, it can be stated that we are rather late in the development process and therefore have a working prototype software which should be used during the evaluation, and not only a paper prototype or a mock-up. Thus, it is possible to test in the field with real users to get a realistic testing scenario and not in a laboratory or online.

The focusing aspects of the evaluation were worked out with the project partner: design and functionality with special regard to usability and acceptance of the prototype software (and not performance or a comparison of two prototypes, for example) to ensure an optimal support of the dispatchers. The prospective users had an active role in the evaluation, but they were not involved further in the software development process and are therefore not part of the decision making (meaning a more passive contribution).

In summary, these are the important boundary conditions for our study:

- Working prototype software
- Study takes place in the field
- Passive contribution of the user
- Late in the development process
- Focus on design, functionality, usability and acceptance of the prototype software

B. Step 2: Defining Important Evaluation Criteria

After this first step of describing important boundary conditions, additional criteria of the evaluation need to be defined.

At the beginning, it was decided that the evaluation methods should generate diverse data to get a comprehensive understanding and a broad spectrum of results. ”Diverse data” refers to subjective and objective data as well as qualitative and quantitative data. The data should be generated either directly during the interaction of the user with the prototype software, or indirectly after having used the prototype software. This depends on the evaluation method and especially on the possibility of integrating such an evaluation method into daily work without distracting the dispatchers too much.

Due to the advanced project progression, evaluation methods with a low to medium expenditure of time and a low to medium effort for analysis and interpretation were needed. Methods with high expenditure of time, high effort for analysis and interpretation were not considered further.

Based on the decision to generate diverse data and because it is not really possible to generate subjective and objective data with the same evaluation method, two different search profiles were created to find all suitable evaluation methods. However, some of the parameters are equal for both profiles:

- High degree of detail
- Low to medium expenditure of time
- Low to medium effort for analysis and interpretation

 ISBN: 978-1-61208-382-7

1) Description of Search Profile 1

The first search profile was supposed to generate subjective, qualitative data and – to make sure that the dispatchers are not distracted during their work – indirect data generation was chosen.

2) Description of Search Profile 2

The second search profile was intended to generate objective and either qualitative or quantitative data gained directly during the usage of the prototype software.

C. Step 3: Choosing evaluation methods with help of the Test Selection Program

All these boundary conditions (see Section A) and criteria (see Section B) are to be entered into the Test Selection Program which was particularly developed to easily choose from many evaluation methods, and which has been described in a previous section.

1) Results for Search Profile 1

For Search Profile 1 (cf. Section B.1) the program delivered the following methods:

- Focus Group
- Interview
  - Half-structured Interview
  - Plus-Minus-Method
- Diary Studies

Because of the great effort to coordinate the interviews and the difficult time availability of the dispatchers being busy in their work, we decided against conducting single interviews. That is why, for the considered project, the methods Diary Studies and a Focus Group were selected.

2) Results for Search Profile 2

For Search Profile 2 (cf. Section B.2), we gained the following methods as results:

- Observation
  - Participatory
  - Non-participatory

We decided to use the Participatory Observation to be fully integrated in the dispatching process and to gain as much objective data as possible. Using this evaluation method, it is possible to interact with the dispatchers and thus prevent them from feeling uncomfortable.

D. Step 4: Deciding on the Right Order in Time

After having chosen the most suitable evaluation methods in Section C, the next step is to decide on the right order in time.

It was decided to start with the Diary Studies, then continue with the Observation and at last, conduct the Focus Group. The reasons for choosing this order were the following: Using the Diary Studies as a first evaluation method allows to ascertain in advance which items to concentrate on in detail during the Observation. This will help the observer to be more focused during the subsequent Observation. Since it was infeasible to discuss remaining questions which arose while reading the Diaries or during the Observation, the Focus Group was chosen as the last evaluation method. With the first two evaluation methods, enough input for the discussion will be gained. Moreover, key items can be discussed in detail and remaining open questions can be clarified. Also, users are able state their opinion and their improvement proposals orally and directly to the persons in charge. With the combination of these three methods in the described sequence, a fully detailed overview of the design and functionality of the prototype software as well as input for improvement can be obtained.

V. CONCLUSION AND FUTURE WORK

Choosing the most suitable evaluation methods, to combine them reasonably and in the right order in time is important for the success of a study since this is the basis for the data generation and thus for undistorted, reliable results. Unfortunately, this preparation is already challenging due to the variety of evaluation methods and the numerous possibilities to combine them as elaborated on in [6]. These encountered challenges are addressed with the help of the presented four-step method (see Section IV) which is integrated in the Test Selection Program described above (see Section III). It aims at supporting the choice and combination of evaluation methods.

For the field study described in this paper (see Section II.C) the software could successfully be used. The next step for the evaluation in the above-mentioned field test is to generate – with the help of the combination of evaluation methods described in the section above – all necessary and also detailed data to improve the prototype dispatching software, to adapt it to the dispatchers' needs in the best way possible such that a better support in connection dispatching is given and thereby the dispatchers' satisfaction can be increased.

Apart from this, there is ongoing research to further improve the process of choosing the most suitable evaluation methods (see Section IV) and in doing so, also the software supporting this process (see Section III). The latter was well proven first in a simulation environment [8], in the field of expert evaluation [8] and finally also in the here presented field test (cf. Section IV).

Before conducting the aforementioned second turn of evaluations, an investigation concerning participating users has to be done:

Since dispatchers located at two different sites participated in the evaluation, a decision has to be made whether to let

- The same dispatchers or
- A completely new set of dispatchers or
- A combination of previous and new dispatchers or
- All previous dispatchers and additional new dispatchers participate in the second evaluation.

A key question to be addressed is which users to omit, all from one or several from each location. In the case of omitting dispatchers, questions to be addressed are the arrangement of the exclusion, which person should be in charge to decide, how to prevent distortion of results caused by preferring power users or sceptics and decision-making without sufficiently knowing the dispatchers.

These issues should be addressed before starting a second session of field tests for the described project. The results we

 ISBN: 978-1-61208-382-7

could obtain so far already have shown possibilities to further improve user interfaces with respect to usability to grant better support for people working in the field of railway control and dispatching.

Further research also concerns the transferability to other fields of research, e.g., railway traffic controllers, air traffic controllers or even different industries. The conditions on these areas seem similar, but a practical proof of application of the presented tool and method is missing.

REFERENCES

[1] DB Netz AG, “Guideline 408 – Driving Trains and Shunting” (“Richtlinie 408 – Züge fahren und rangieren”). Karlsruhe: DB Kommunikationstechnik GmbH, 2012.

[2] A. Kauppi, “A Human-Computer Interaction Approach to Train Traffic Control”. Uppsala: Department of Information Technology, 2006.

[3] C. Kendziora and A. Böhme, “Guideline 615 01 – Traffic Management Passenger Transport” (“Richtlinie 615 01 – Transportleitung Personenverkehr”). Frankfurt am Main: DB Fernverkehr AG/DB Regio AG, 2010.

[4] Ministerium für Infrastruktur und Raumordnung Land Brandenburg, “Financing Public Transport in the State of Brandenburg” (“ÖPNV-Finanzierung im Land

[5] Brandenburg”). Potsdam: Land Brandenburg, Ministerium für Infrastruktur und Raunordnung, 2007.

[6] T. Schnick, A. Wolters, and A. Stelzer, “Visualisation of Connection Conflicts in Dispatching” (“Visualisierung von Anschlusskonflikten für die Disposition”). Deine Bahn, vol. 41, no. 6, pp. 26-29, June 2013.

[7] I. Schütz, “New user interface concepts for railway management” (“Neue Bedienoberflächenkonzepte für die Betriebsführung von Eisenbahnen”). Darmstadt: Department of Railway Engineering, 2012.

[8] A. Stelzer, A. Oetting, and F. Chu, “Connection Dispatching - an Algorithmic and Visual Support for the Dispatcher”. The 13th World Conference on Transportation Research – Rio de Janeiro 2013, July 15-18, 2013, July 2013, ISBN: 978-85-285-0232-9.

[9] A. Stelzer, I. Schütz, and A. Oetting, “Evaluating Novel User Interfaces in (Safety Critical) Railway Environments”. Human-Computer Interaction. Applications and Services, 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, vol. 8512, June 2014, pp. 502-512, doi:10.1007/978-3-319-07227-2_48.

[10] C. Streitzig and A. Stelzer, “TU Darmstadt – Research, Training & More Besides”. EURAILmag, issue 26, 2012, pp. 152-159.

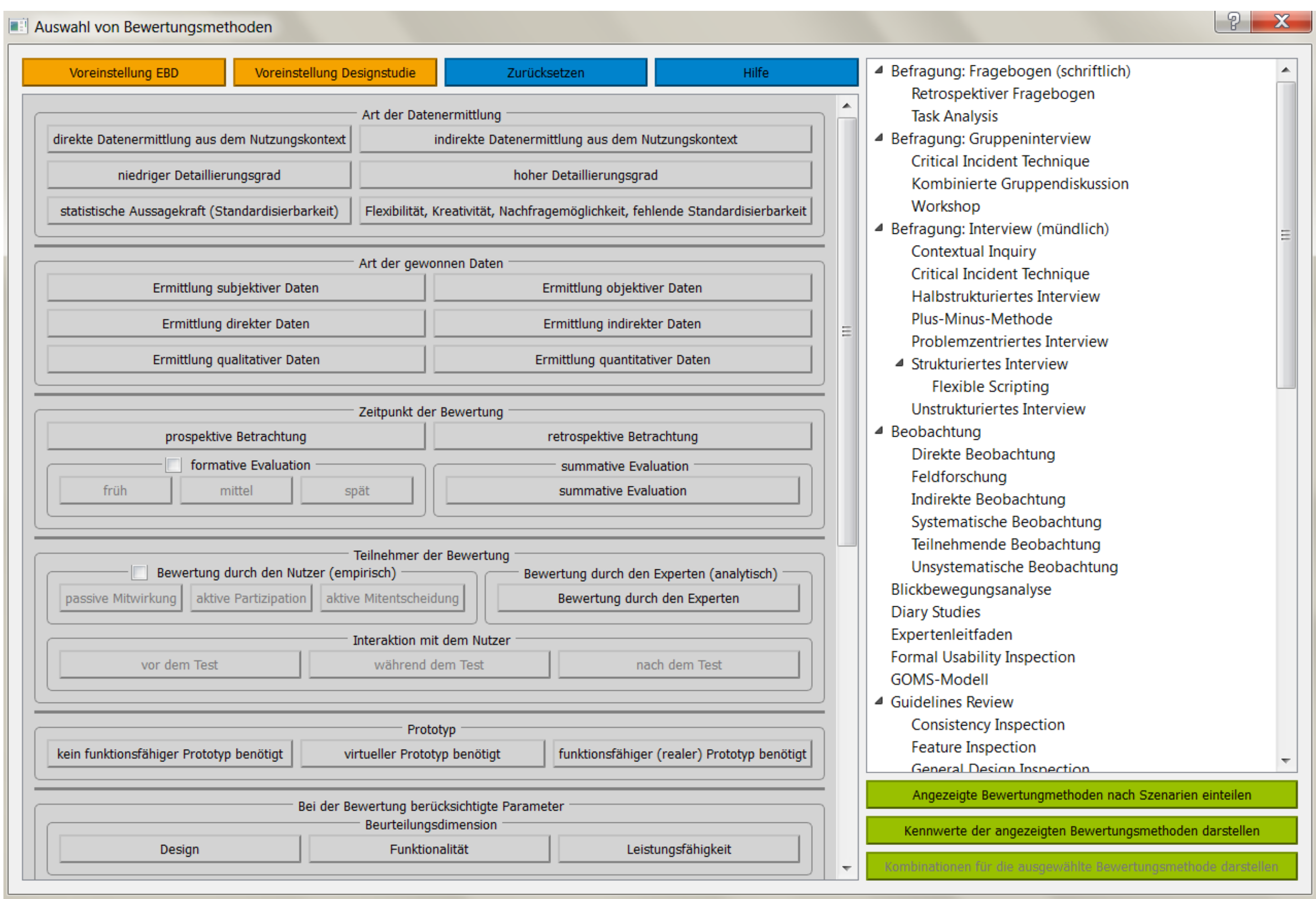


Figure 1. The Test Selection Program. (Source: own representation)

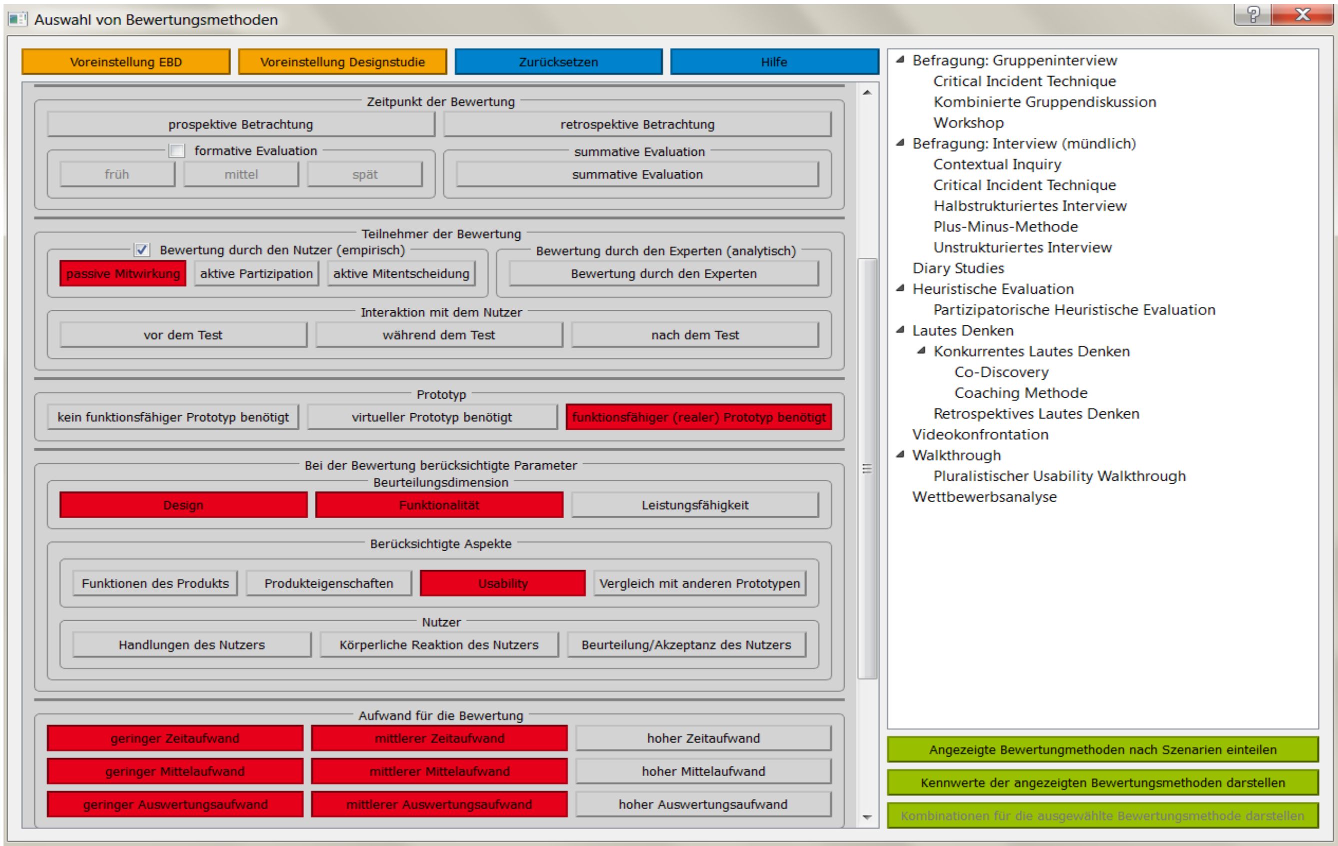


Figure 2. The Test Selection Program. Visible are some selected buttons and an accordingly filtered list of evaluation methods. (Source: own representation)

Inversus

The Sensitive Machine

Luís Leite
IT / FEUP
Porto, Portugal
luis.grifu@gmail.com

Verónica Orvalho
IT / FCUP
Porto, Portugal
veronica.orvalho@gmail.com

***Abstract*—Inversus is a digital interactive installation that explores the relationship between user actions with common objects: lamps, speakers and fans. It is an interactive machine that shifts the conceptual understanding a user traditionally has about a specific object, making people wonder: why should a lamp be used only to illuminate? In fact, lamps, speakers or fans are commonly used as output interfaces, imagine now, what would happen if we turn the output into an input interface? This machine explores an inversion on this relationship by using lamps as light sensors, speakers as pressure sensors and fans as blowing sensors.**

Keywords-HCI; audiovisual instrument; DIY; user experience; interactive installations.

I. Introduction

Electronic devices are designed in general with a specific function. Today, many electronic devices are getting smarter with generic purposes, sensing the environment to respond to all kinds of inputs providing a wide range of capabilities. Devices that are designed for one main purpose can be used with a different one depending on the user intention. One can use a smart television to make a phone call, or use a smart phone to watch television. There is a shift in the way devices and objects are designed and in the way users relate to them. We live in a sensorial age where sensors are included in every day objects. But, instead of replacing all the objects and devices with smart new ones, we might try to recycle them. With the “Do It Yourself” (DIY) approach, ordinary people have the power to change and be creative to find new uses for their old devices. Our project challenges participants to discover and explore new applications for traditional one-purpose machines. Inversus, encourage the participants to rethink the original functionality of the objects into other potential uses, searching for new capabilities and combinations. [1]

Inversus is a sensitive machine that captures human interaction to produce sound and visual kinetics. It is a performing instrument that gives life to a mechanical flower, which spins when someone blows into the machine producing an animated shadow similar to the shadow puppetry effect (Figure 1). A virtual marionette lives inside the machine that reacts to the pressure of four sensitive pads; this marionette is rigged with virtual strings that are mapped to the pads that make them squash and stretch producing animation, like pulling the strings from a marionette. Based on human interaction the machine generates animation in the physical and in the virtual space in the mechanical and in the digital domain enhancing the user experience. This audio-visual instrument made from a washing machine spins a

Figure 1 – 3D simulation of the machine with the shadow of the flower projected on the wall.

physical colorful wheel after sensing the interaction, mixing all the media elements as a metaphor for the real washing machine. Users are familiar with these common objects and recognize their original function, but when they start interacting with Inversus they realize that it as a different purpose. They experience a strange reaction to a familiar object and explore its new potential making a dialog with the machine.

The remainder of the paper is organized as follows. Section II presents the design principles and describes how the idea was developed. Section III describes the implementation of the system and framework. Section IV addresses the impact of the interaction in participants. Section V presents the conclusion.

II. Design

The idea of building a machine from recycled materials was influenced by the toy hacking spirit and by an ecological thinking. Inversus was designed based on the way we play musical instruments with three types of interaction: i) striking a drum with the hands, ii) moving the hands above a theremin, iii) blowing air to a flute with the mouth. These gestures were adapted to present distinct interaction experiences based on the sensitivity of each sensor divided into three categories: Touch – by touching the color pads the

 ISBN: 978-1-61208-382-7

user triggers sounds that respond to the pressure (like playing drums) and animates a virtual marionette that reacts to the pressure and represents the mechanical body;

Motion – by moving user hands above Light Emitting Diodes (LED) it produces a continuous sound (sound keeps playing until the hand moves away from the LED) similar to the interaction with a theremin instrument;

Blow – by blowing a fan, the user changes the frequency of sounds and spins a mechanical flower producing a shadow animation in the wall that represents the organic body.

In our methodology we searched for electronic materials that could be recycled into sensors. And although we did not seek total accuracy the diversity of the sources brought variable results. Our challenge was to create a balance between all the sensors and build an instrument that could be played by any user distinguishing the sensitivity in the different interaction methods.

Inversus was made from the fowling recycled materials: the chassis of a washing machine was used to build the case and to provide a link to familiar objects; the pressure sensors were made with piezoelectric from small speakers; LED´s from toys were used for building the light sensors; the blowing sensor was made from a computer fan. In this way all the objects were employed with a different purpose from their original design.

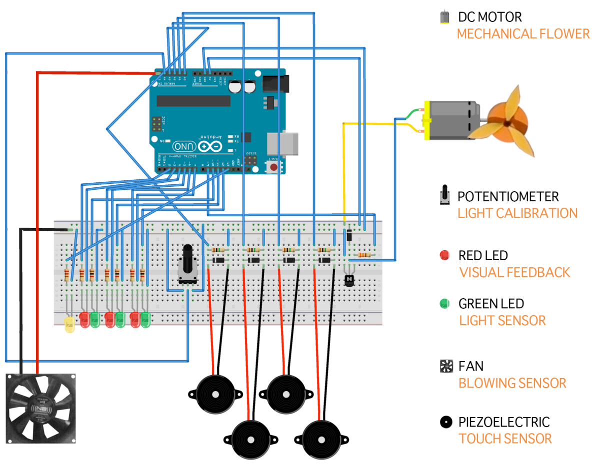


Figure 2 – Orthographic top view from the connection sketch with the input sensors and the output motor.

Using LED´s as photodiodes we can build low-cost light sensors [2]. A LED is basically a P-N junction, an interface between two types of semiconductor material that conducts electricity in two different sides. If a photon (light particle) penetrates the region around the P-N junction it can potentially interact with one of the atoms transferring it´s energy into the atom´s electron creating an electron hole pair. The electric field across the region causes a tiny current to flow, known as the photo current. Because the junction area of the LED is small the result current is also small. Adding up photocurrents over time makes the diode look like a capacitor and will increase the result current.

III. Machine Development

To implement this solution we used the Arduino microcontroller, which provides configurable pins making it possible to change the direction of the current flow. We wired up the LED´s to the Arduino connecting the anode (positive) of the LED to a digital pin and the cathode (negative) to an analog pin (Figure 2). First, we charge up the capacitor by making the anode negative and the cathode positive in a short amount of time. Then, we take a reference measurement, reading the analogue voltage and waiting for the photocurrent to be integrated. Finally, the voltage is measured again and subtracted from the reference value. LED´s are sensible to thermal noise and are not precise as light sensors, in particular in dark environments. A potentiometer was included to calibrate the sensitivity of the LED´s to increase the accuracy of the readings face to the light environment. For pressure sensing, we used a piezoelectric element [3] from small speakers attached to foam pads [4]. Foam is a good material to attenuate the pressure impact and to expand the area of interaction. We connected the "piezos" through resistors and diodes to the analog input. The musical note velocity increases proportionally to the pressure that is made in the pads like in a real drum. The blowing sensor was a PC fan connected with two-wires, one connected to the ground and the other connected to the analog input. Based on the blowing strength, two DC motors increase or decrease the rotation of a mechanical flower and of a colorful wheel.

Our framework is based on serial and network communications that link hardware to software. Sensor readings from Arduino are mapped to control and note messages via MIDI protocol, used in musical-based applications, and sent through the serial port. Messages from the serial port are then routed to a virtual MIDI port using the software hairless-midiserial that allows MIDI-based applications to capture the messages. The messages are then mapped to Ableton Live, a sound application that triggers and shapes the sound. To route the messages into animation-based applications, we use the visual programming environment Pure Data to convert the MIDI protocol into Open Sound Control (OSC) network protocol, which is available in many multimedia applications. The OSC messages are then received in Animata, a real-time animation software, and mapped to a skeleton of a virtual marionette that reacts to the pressure of the Pads.

IV. Interaction

We design Inversus to generate an emotional response to participants, motivating them to play and discover cognitively the interrelations between their actions and the result in the environment.

By providing a multisensory experience the participants have the feeling of embodiment in the interaction with the machine, which is enhanced by the relationship between action, cognition and the environment [5]. The distinct interaction methods presented in this installation provided an opportunity to understand how users establish their

 ISBN: 978-1-61208-382-7

communication with an unknown interface. We used some design conventions, such as shape, color or spatial distribution of the controllers that provided some interaction clues.

Figure 3 – Inversus installation presented in a art exhbition: Cheia, Póvoa do Varzim, Portugal.

Inversus was available for public interaction in two artistic exhibitions during two months (Figure 3). About 300 people attended the event. We performed a pilot study with 10 participants. We observed in detail the participants interaction and behavior with Inversus. From the observations, we detected interesting trends in embodiment interaction [6] that open new lines for further research. We found that most of the participants took some time to analyze the device before trying it, beginning their interaction by touching the pads seeking a relationship between the strength of their actions and the intensity of the sound and animation. After this understanding, most of the users that were observed moved their attention into the blinking LED´s by touching them. And although, not touch sensitive the light occlusion occurs producing an effect in the environment inducing the user in how to interact, establishing a relation between their hand position and the sound and the light of the LED. Participants approached the fan wheel in the last stage using their hands, and again, the effect produced in the environment pointed to the correct way of interaction. Changing their interaction behavior and by blowing to the fan, they realize that the fan wheel spins faster producing a more intense modification in sound, as well as making the mechanical flower to spin, producing a shadow animation in the wall. After the explorative phase and knowing how the interface works, participants expanded to a multimodal interaction combining the controllers like playing drums. By making it´s own cognitive observations and by establishing relations between their actions and the environment the participants become embodied with the a strange interface.

V. Conclusions

This interactive installation demonstrates how to "hack" electronic components used as output devices found in everyday objects and applied them as input devices in a different application domain. Recycling technology in an innovative way presents interesting challenges to users that can increase the feeling of empowerment and can contribute to novel interaction methods and applications. Two examples of the diversification of use of specific purpose technology can be found in game controllers such as the adaptation of the Nintendo WII remote as a low-cost interactive whiteboard [7], or in the adaption of Microsoft Kinect as a 3D scanner [8].

Inversus challenges the participants to interact with a machine that looks familiar, but presents unknown options and functions. It gives the participant the possibility to re-think the interaction fundamentals behind everyday objects. The participant is invited to discover the relations between the input sensors and the audiovisual output, and to realize the sensitive capabilities of inverted devices from recycled materials. As a result of the interaction experience, we hope to instill in the participants the desire to explore and create new communication challenges by re-defining the concept, usage and interaction mechanism of common objects.

Acknowledgment

This work is partially funded by FCT Fundação para a Ciência e Tecnologia (BD/51799/2011), Instituto de Telecomunicações, Faculdade de Engenharia, Universidade do Porto. Thanks to Marcelo Lafontana, Sílvia Fagundes, Rui Torres, Leonel Cunha and Jacqueline Fernandes.

References

[1] F. Blum, "Digital Interactive Installations: Programming Interactive Installations Using the Software Package Max/MSP/Jitter." VDM Verlag, 2007.

[2] S. E. Hudson, "Using Light Emitting Diode Arrays As Touch-sensitive Input and Output Devices", Proc. of the ACM Symposium on User Interface Software and Technology, NY, USA, 2004, pp. 287–290.

[3] J. Noble and N. Joshua, "Programming Interactivity: A Designer's Guide to Processing, Arduino, and Openframeworks", O'Reilly Media, Inc, Sebastopol, USA, 2009.

[4] Mschaff, "Ardrumo: Mac OS X Virtual MIDI Interface for Arduino Prototype Boards", Mschaff. [Online]. Avaiable from: <http://inivent.com/ardrumo/> 2014.11.15

[5] A. N. Antle, "Lifelong Interactions: Embodied Child Computer Interaction: Why Embodiment Matters", Interactions, vol. 16, no. 2, NY, USA, 2009, pp. 27–30.

[6] P. Dourish, "Where The Action Is: The Foundations of Embodied Interaction",The MIT Press, Cambridge, USA, 2001.

[7] J. Lee, "Low-Cost Multi-point Interactive whiteboards Using the Wiimote.", Johnny Chung Lee. [Online]. Avaiable from: <http://johnnylee.net/projects/wii/> 2014.11.20

[8] S. Izadi, et al., "KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera" , Proc. of the ACM Symposium on User Interface Software and Technology, NY, USA, 2011, pp. 559–568.

Instruments for Collective Design in a Professional Context: Digital Format or New Processes ?

Samia Ben Rajeb, Pierre Leclercq
LUCID, University of Liège
Liège, Belgium
e-mail: {samia.benrajeb, pierre.leclercq} @ ulg.ac.be

***Abstract*—This article reports on different collective practices and their instruments observed in the context of professional design activities. Based on interviews and in situ observations of customs in six architecture, engineering, and design agencies, it shows the diversity of typologies of collective activity, identifies the main factors of collaboration, and concludes on the needs for the instrumentation of professional practice. Is it a matter of building higher and higher performance digital formats for a shared modeling of the project or to lean instead towards the creation of new processes of group management and remote work by several people? From the results of this observation of collective practices in design, this article allows one to highlight the real needs of the agencies and to help their teamwork. These needs consist of: 1) facilitating reflection, 2) managing changing dynamics, 3) allowing the reflective exploration by several persons between space and time, and finally 4) guaranteeing a common progressive and exploratory strategy between actors whose expertise and commitment differ within the same process.**

Keywords-collaborative design; professional practice; observations; tools and processes.

I. INTRODUCTION

Faced with competition, tight deadlines for delivery, and qualitative and regulatory demands, which are increasingly complicated and difficult, the architectural and design agencies are innovating concerning interdisciplinary approaches, associating different skills that are necessary for the realization of the project [1][2]. Nowadays, the designer does not work alone on a project but with other experts, bringing together architects, engineers, designers, sociologists, economists, etc. The process has thus become collective, reuniting different skills which must be applied starting with the first phases of the project design.

Several researchers are indeed interested in group design and have proposed tools to facilitate the sharing of communication [3][4]. Since the development of the Internet, new technologies have been conceived, stemming from the CSCW (Computer-Supported Cooperative Work) scientific field to support group work [5].

Called groupware, these tools are used today to guarantee the coordination, to manage the tasks and to enable the cooperation between several actors separated geographically. But most of them only partially meet the specific needs of the agencies and their design processes whose synchronous sharing of annotations, graphic interactions, and information management is essential [6]. This sharing is even more indispensable in daily activity because of qualitative requirements and increasingly coercive regulations. In addition, most of these agencies are faced with strong competition and tighter and tighter deadlines even though their activities involve different skills due to design, architecture, engineering, ecology, ergonometry, sociology, etc.

This is why it seems essential to clarify the specificities of the real activity of design in agencies, in terms of tools as well as work methods, before even beginning to compare the computer support and/or solving the problems that can be linked to it. This choice is based on the working hypothesis where the project considers design as a unique and complicated activity gathering together different viewpoints and having to answer to several choices, problems and constraints linked to the project. Our study is based on an analytic approach inspired by the methodological approach of “action research” [7], which consists in going into the field in order to get directly into a real industrial context to observe and study some work tools and procedures.

This article presents in sections II, III, and IV, the latest developments and the question of research that are the framework of our study and the methodology on which it is based. Section V follows by exposing different real practices in architectural, engineering and design agencies, which present the particularity of leaning on a collective, multidisciplinary, and multi-sites activity. Their practices, procedures and tools are put forward in section VI, so as to point out, through the diversity of observed collective situations, the collaborative operations put into practice, as well as the different co-spaces that make up their work environment.

II. LATEST DEVELOPMENTS AND RESEARCH QUESTIONS

Collective activity has been the object of many research projects. It is usually compared to individual activity [8] and may take several forms depending on the field in which it is examined. To define the different forms of collective activity, some researchers have distinguished between them according to the objectives sought by the protagonists of the project and the specificity of the tools used to work in groups [9][10]. Other researchers have highlighted the influence of the number of actors and their hierarchical relations in their activities [2][5][11][12]. Others have shown that collective activities change according to the task, the kind of exchange, and the tools used by the actors to work together [13][14][15]. Others have specified the collective activity according to the time reference and the space of exchange, and they have even proposed a classification of the tools, which allows them to support this work in groups (groupware) via the time/space matrix first set up by Johansen [16] and then repeated by Ellis, Gibbs and Rein [17], and Gaver [18]. Other researchers have also proposed other kinds of classification of groupware according to their objectives in the conception: cooperation, coordination or communication [19].

Faced with this abundance of classifications of collective activities and their tools, our beginning premise is to focus our research neither on the comparison of performances of computer-assisted design nor on the solution to their problems. The primary aim of our study is to clarify the collective design activities set up in the agencies by simply relating their specificities and their real needs without preconceptions or the imposition of any tool or group-work method.

III. Research Framework

This work fits partially in the frame of the program "Creation: actors, objects and contexts". Called CoCreA and financed by the National Agency for Research in France, this project groups 3 research laboratories in different fields (Limsi-CNRS, Map-Maacc and LUCID-ULg). Its objective is the development of new understanding of creative collaboration in architecture, examining the customs, their tools for sharing, as well as the implications in the actual design activities [20]. This study framework has enabled us to investigate the professional context on which our research question is based, examining the actual collective design activities and their need for assistance.

IV. Methodology

For the study of collective design activities, several protocols have been set up. According to Ericsson and Simon [21], two complimentary trends of intervention and analysis are to be distinguished between:

- The "retrospective protocols," which concern the study of projects regardless of the situation in which they evolve;
- and the "concurrent protocols," which take an interest in the analysis of the real activity, taking into account the particularity of the designers, their tasks, and the environment in which they work and collaborate.

Our study joins more with the second trend and is more particularly based on the "action research" methodological approach, whose first objective is to examine the practices in their actual context, and group together action and reflection, theory and practice, with the participation of individuals and their communities [22]. Based on ethnographical studies and faced with the complexity of collective design activity, we have chosen this approach because, according to Brydon-Miller et al. [7], it is better suited to specify the action and the know-how of the experts in their context: between actors, social hierarchy, the urgency of the reports, regulations, procedures to be respected, deadlines to meet, diversities of situations in which the objects to be designed evolve and their protagonists, etc..

In order to build up a specific knowledge of this activity, we have been thus based on on-site observations and interviews. This data involves several agencies having a collective design activity and meeting the same criteria: they are multi-disciplinary, multi-site and/or work with other design agencies, and declare that they have set up real collaborative practices in their activity (architecture, engineering, or design agencies).

We have been able to accumulate observations and interviews from different meetings of operators working in the agencies. We have then organized our analytic basis on 5 themes:

- The way the agency functions: the methods and habits of the agency, the different sectors that it deals with, the operators that it brings together depending on their expertise, competencies, and knowledge in the project;
- the kinds of collaboration and the characteristics of the collective activities: the face-to-face or remote collaborative practices of the agency, inter-agency collaboration or collaboration with external actors (sub-contractors, consultants, general companies, etc.), the kinds of meetings that they set up, etc.;
- the tools and procedures of exchange and their roles in the collaboration and in the design process: in this case we mainly based our research on the cases of remote collaboration highlighting the kind of project concerned, the tools used, the procedures and methods chosen or adopted to work together remotely, as well as the consequences these practices have on the design methods;
- the exchange methods and principally, the graphic representation: the role of the graphic representation as object between the collaborators, the tools used to share, the annotation and/or modification of these representations, etc.;
- the expectations and perspectives: the problems that designers raise during their remote collaborative activities and their suggestions for ideal instrumentation in order to efficiently aid the face-to-face or remote collaboration in design activities.

These corpus have thus enabled us to identify, on one hand, the different cognitive operations put into play and, on the other hand, to highlight some procedures, methods and tools set up in these practices. Our objective is to identify the means and modes of collaboration used by the designers in the preliminary phases of the project.

V. Analysis Compared of the Collective Activities in Agencies

Our observations and interviews were used to construct a chart of collective activities observed in the 6 following agencies: the AIA agency (between Paris, Nantes and Lyon), ORA-ITO (interviews of project leaders in Paris), Mikado

Architecture (Lille), Art & Build (observations in the main agency in Brussels), Architecture Studio (Paris), and Gehry Partners (at Gehry Technologies in Paris). This corpus, in its diversity, enabled us to highlight several kinds of collective activity, methods and tools used to be able to work with several persons. The kinds which are shown are based principally on the dichotomy Space/Time as set up by Johansen [16] in the field of CSCW; see table 1. We observe several recurring tools, such as face-to-face meetings, or by telephone, videoconference, e-mails, post-its, internal network exchanges, electronic plan boxes, etc.

TABLE I. SYNTHESIS OF GROUPWARE USED IN AGENCIES FOR GROUPWORK [16]

Collaborative work	**Same time**	**Different times**
Same place	Syncronous : real co-attendance meeting	Asynchronous + delayed : - post-it™ - email - file server
Different places	Synchronous + distributed - phone call - visioconference - screen sharing	Asynchronous + distributed: - mail and email - file server - wiki

Nevertheless, differences in the practices have also been observed. These differences also depend on the project and the degree of complexity that is involved. They show a range of kinds of collective activities and a diversity of solutions, sometimes pre-existent and often put in place by diverse procedures or by diversion of tools. Some of these tools and procedures are the same in all of the agencies such as the mail or the establishment of a graphic chart. But particular cases are also highlighted by the specific activity of each agency. We synthesize, in the following section, the principal diversions which are set up to respond to, as well as possible, the constraints of distance and/or the differences of the teams that have to work together in the context of project design.

A. *The case of AIA*

In the context of Design/Construction cooperation, the agency has had to use a system of computer screen sharing to manage their regular inter-agency meetings. In the observed situation, this tool is coupled with a telephone, hooked up to a speaker, to speak over the distance. The collaborators can thus see, at the same time, the same image, all the while having the possibility to point to spaces (via the cursor) and share some annotations (realized as well as possible via a mouse). The collaborators can consider, comment or annotate the documents shared by one of the two designers. The latter assumes the role of "chairperson" because he is the only one to decide which space can be shared and used by all players (see VI.C : as We-space) or kept in hidden notes as private space (I-space).

To manage the diversity of actors in the same project and their possible replacements, AIA has set up a procedure involving the updating of a "thematic notebook". This "thematic notebook" serves as a point of reference to everyone joining the group on the way. It is a sort of project logbook which enables everyone to understand different decisions taken in the course of the project, its evolution, and references. Ranking the data to be taken into account and essentially serving the construction of a common ground, it shows the work done during the conception, and makes the choices of the team and the architects' organizational schemes explicit in relation to the project. It also regroups technical notes, and the minutes of the meetings which can also serve as a support for the contracting authority.

B. *The case of ORA-ITO*

Under the direction of Ora-ito, the agency makes people with different skills intervene: industrial designers, graphic designers, architects, etc. It also works very closely with several industrialists for the manufacturing of the products, integrating even the packaging and the shops which expose them. In spite of the multi-disciplinary activity involving several protagonists found on different sites, this agency has clearly expressed its lack of interest for remote collaboration activities. "If we need to work together, it is enough to gather everyone around a table" declares one of the project-leaders of the agency.

C. *MIKADO Architecture*

To be able to devise a complicated cooperation with tight deadlines, the Mikado agency is associated with other external agencies (one in Paris and the other in London). In this particular situation the actors do not have the opportunity to get together often enough to work together on the project. They have thus chosen an online game of virtual reality (like Second life®). Each of the actors invented an avatar to create a model shared online for their project in this virtual space that they manipulate together over a distance. Each time a problem comes up, they get in contact by telephone, then they get connected, each one via his computer, to work together, directly on the shared virtual model. These actors have had to face many difficulties because the 3D model manipulating tools offered by the game remain too limited and basic to assume efficient oral and graphical exchanges in real time.

D. *ART & BUILD*

In the framework of designing a collective living project which is close to one of branch offices of the head office Art & Build, the two sites use a new system to share graphic annotations in real time, called Collaborative Digital Studio [23]. Developed in the LUCID Laboratory of the University of Liège, this system has been loaned to the enterprise to support their activity of inter-agency remote design. In contrast with the experience of Mikado, Art & Build is satisfied with the use of this new communication and remote collaboration system. At any time of the day, they can call, get connected at a distance and work together in real time on the graphic documents that they have just produced. In this way they can better coordinate. Nevertheless, the observation of their use also shows that the actors do not make use of this system as much as supposed, because, in fact, the system has not changed the frequency of the trips made by the collaborators between one site of the agency and the other. Rather it enables other kinds of meetings that are more spontaneous and shorter. According to the users, this is not due to bad understanding or appropriation of the system, but rather to a habit which is not yet part of their daily routine: "We know that the system was there, but we forget to use it each time that we have to communicate or work with the other agency: the reflex is to take the train and go there directly…"

E. *Architecture Studio*

The particularity of this agency is that the designers have different nationalities with different architectural cultures. To manage this kind of situation, all of the project teams must accept the suggestions from the others and adapt to the multicultural differences in the name of the project. In this context, the agency is also faced with language and reference-synchronization problems, and the knowledge of the others. To try to eliminate the effects of hierarchy and to let everyone express themselves as they please, the agency has set up a procedure based on the codification of shared drawings according to their colors. In this way four kinds of drawings have been defined and codified: 1/the red drawing for all of the existing data that is important to keep in mind for the project design: it cannot be modified, nor questioned by the designers; it represents the only fixed and certified element in the project; 2/ the green drawing contains the remainder of the elements that compose the project such as the walls, the closets, the opening, etc.; 3/ the black drawing is for the elements that concern the existing plan except those already put in the red drawing; 4/ the blue drawing which transcribes all the annotations and information about constraints to be integrated in the project design. According to the designers, even the traces left by the scratches of a razorblade on the tracing paper are important because they give an idea of the history of the project. This way, all the drawings represent the basis of the design of the model and are the result of collective decisions managed by this kind of procedure.

As the agency also works in collaboration with other branch offices in different time zones, it has had to invent a system to share digital annotations in real time. It had set up a system diverted from shared screen projection that it threw together as well as possible with the tools.

F. *Gehry Partners*

Working with non-standard shapes for the technical design and the construction of the Louis Vuitton Foundation pavilion, the Gehry Partners agency uses programs originally developed for the aeronautic sector [24]. In this context, the team in Los Angeles and the team in Paris have given themselves 2 distinct roles in the design process: the former mostly takes care of the formal and functional aspects of the project, the latter focuses more specifically on the technical aspects and the structural calculations of the building. The first 3D model was made by the Los Angeles team. It serves as a digital model so that the head architect, F.O Gehry, can refine and transform his project according to choices related to his own pertinences. Then the team in Paris takes back the first model, extracts the geometry and builds a new model with defined and shared parameters in which they insert their choices for the structure and the technical calculations of the building. This second model was developed from the moments of collaboration in the presence of the different engineers and architects from Paris who work with other design and calculation tools before integrating their decisions in the model with the shared parameters. This model is then approved by the chief coordinator, who is also in Paris, and then put in the 3D model with the shared parameters. It is not only visible by the whole team in Paris, but also accessible and consultable in Los Angeles. The head agency can also survey the evolution of the technical model at any time and distance. From the first transformation or modification of the building, the head of the project can decide to back up or to transform the model so that it suits him. The shared program of the 3D model serves more to validate choices, to evaluate them in relation to decisions inserted in the model by other people, and to coordinate the team work rather than a collaborative design, to the management of the negotiation and the questioning by consensus of the project.

VI. DIVERSITY OF COLLECTIVE SITUATIONS VS. NEED FOR ADEQUATE TOOLS: BETWEEN TOOLS AND PROCEDURES

The following observations enable the identification of the principal concepts documenting the question of remote collaboration in an actual situation of architectural design, including:

- The kinds of collective activities: between collaboration and cooperation,
- the collective operations: between design, collaboration and tools,
- the work co-spaces: I-space, We-space, and Space-between.

These observations highlight the need for the construction of common referential operatives faced with the diversity of each of these ideas as well as the part played by negotiation, evaluation and the questioning by consensus in the process of collective design. Is this need better orchestrated by digital supports which are more and more reliable for shared establishment of a project model, or by setting up of new group management practices and work done by several people?

A. *Diversity of the shape of collective activities*

To respond to the complexity of a project, several regular or spontaneous meetings are organized. Some concern the organization of the agency and its branch offices, others, deal more specifically with the project. Regular meetings are essential for the organization of the group and its coordination. They enable the collective decision-making and choices concerning the project and to synchronize each one's tasks in the design process. These meetings give control, coordination and steering of the collective design activity. They tend to reduce misunderstandings and build the shared group consciousness, which is necessary in any collaborative situation.

However, to respond rapidly to some constraints of the project, the designers also resort to spontaneous meetings. These meetings usually take place in co-presence, with face-to-face discussion about certain aspects not mentioned in the formal meetings. Meetings are also held over distances via videoconference systems, screen-sharing or diversion of a group of tools (for example on-line virtual reality games used by Mikado). These meetings generally try to respond to an immediate need to deal with questions in common. They mark the milestones of the project: (1) moments when the designers collaborate, think and make decisions, by negotiation and consensus about choices which concern either the project or the organization of the group, and (2) moments when each one focuses on their own tasks for the same shared objective. This passage from a moment of collaboration to a moment of cooperation implies, on one hand, dynamics of learning (which tend to gradually cross during the exchanges [25]), and missions to share and divide between different actors, but also communication tools in common. On the other hand, the complementarity of the moments when the designers cooperate and other moments when they collaborate, just as the passages from one to another are important to manage during collective activities (see figure 1).

When they cooperate, the designers do not need to see each other, each one doing his task then waiting for the validation (or not) from the coordinator. When they collaborate, the designers synchronize their knowledge (cognitive synchronization) and try to build a mutual consciousness of their activities, tasks and contexts to respond jointly to the project needs [26]. Nevertheless, this mutual consciousness implies a sharing of knowledge related to the context of the project design and to the tasks of the actors through a cognitive as well as temporal-operational [27] synchronization.

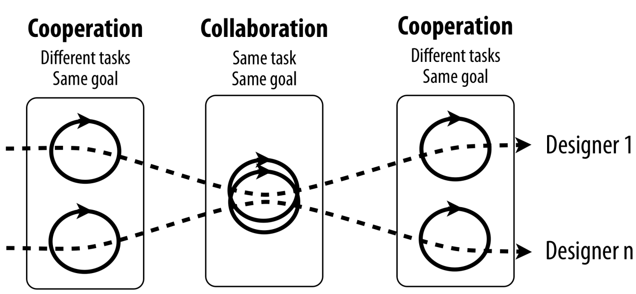


Figure 1. Collective design process

It is important that the tools take into account these passages between the different kinds of collective activities in the design. In fact, the actors need to structure their activity and their tasks according to the project needs and its level of complexity. Some parameters, as important as the cognitive synchronization or the need to put together a common referential operative, are only partially managed by the groupware currently used in the agencies. This is why the coordinators have such a dominating place in the process. Also, to partially meet these needs, the designers are often obliged to divert a group of tools or call on production groupware which often proves to be not adapted to the preliminary design phases.

B. Diversity of collective operations

Based on observations of these different collective design activities, our analysis, based on the field of applied architecturology [28], have shown that specific cognitive operations are put into play. Some operations are linked to the design itself of the project, others are linked to collaboration strategies and the work done together and others are specifically linked to the appropriation and use of the tool which is used to work together. If we focus only on the operations linked to collective activity, we enumerate several categories [29]: pooling, interpretation, discussion / evaluation / reconsideration, automation / cooperation, validation, conflict solving, decision / prescription, building group work strategy / coordination. Putting these operations into play is transmitted through speech as well as by drawing, regardless of the kinds of collective activities in which they intervene: regular or spontaneous meetings, taking place face-to-face or over a distance, simultaneously or not.

During these meetings, the actors share graphical representations which can have three roles: that of mediation, of translation and/or representation [9]. These representations are considered either as a "closed" intermediary object, which cannot be argued with or modified, or as an "open" object, which can be discussed and questioned during the meetings. This intermediate object is transformed and evolves between the moments of collaboration and the moments of cooperation. The shift from one moment to another demands that the representations must be at least standardized, adjustable and multipurpose in order to be understood by the other collaborators. It is in this way that the collective operations such as: pooling, interpretation, discussion / questioning the decision / instructions are primordial in the choice of tools or procedures to be set up for collaborating. Different examples coming from our observations can be cited:

- Pooling has been enabled by the codification of the colors in the activity of Architecture Studio, thus backing up the representation between collaborators by sharing open intermediary objects;
- the interpretation has been enabled by sharing a 3D model fixed between remote Gehry collaborators, thus supporting the translation between the collaborators by the sharing of closed intermediary objects;
- the discussion / questioning has been enabled by the sharing of a 3D model built via virtual space between remote collaborators of MIKADO, thus supporting the mediation by the sharing of open intermediary objects;
- the decision / instructions have been enabled by a thematic notebook set up by the AIA, thus supporting the representation between the collaborators through the sharing of closed intermediary objects.

 ISBN: 978-1-61208-382-7

C. Diversity of working co-spaces

The idea of space is decisive in collective design activities [30]. In the case of cooperative activity, it is easier to work at a distance than in the case of collaborative activity where the sharing of space and the context of the work in real time proves to be more necessary. This shared work space is perhaps not physical, in rooms or in offices: it can also concern hybrid spaces, bringing together at the same time virtual and physical environments. This hybrid space implies the setting up of intermediary work spaces that are short-lived and that are created according to the needs of the designers and their negotiation strategies.

We suggest, in figure 2, to distinguish three kinds of intermediary space composing the joint work environment [31]: the I-space (representing the personal work space), the We-space (representing the shared work space) and the Space-Between (representing the work space built between designers separated from the group).

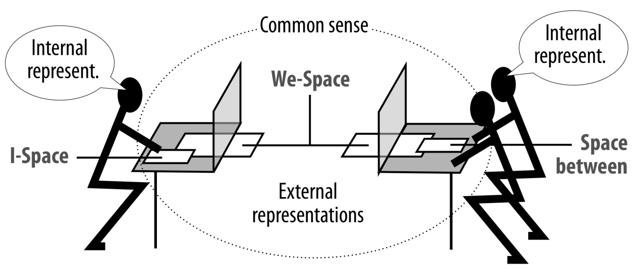


Figure 2. I-Space, We-Space and Space-between as intermediary work spaces.

The role assigned by the designers to this or that space can change according to the objectives, needs, choices related to the negotiation process and the arguments between the actors. The relation between the individual and the situation in which he evolves is emphasized - cognition situated [32] – by the relation that the actor has with his space, his tools and the other actors who surround him in a remote collaborative environment. If, for strategic reasons, a designer decides not to discuss a solution with the collaborator in the same office, the shared document becomes a private space, or a Space-between.

Thus it is important that the intermediary tool offers this flexibility between the intermediary spaces. In fact, the designers need to structure their information and their display interface, to be able to navigate easily from one document to another and to create work methods that are adapted to their situation. All these parameters are only partially managed by the screen sharing system (as that which is used at AIA to respond to support design/construction collaboration) as it imposes the designation of a "chairperson" who is the only one who can make manipulations on his screen. The collaborators, being only observers, can just annotate the documents and point to certain zones of the work, only if the chairperson enables them to. Thus, this tool adds value to the individuality of the designer/chairperson at the expense of equal sharing between collaborators. This even has an influence on the negotiation process, evaluation, and questioning by consensus between the collaborators. In fact, when the actors get together for a meeting (from a distance or face-to-face), each one arrives with his/her solutions, points of view, and references (I-Space) that he shows to his collaborators. He/she evokes them during the meeting (We-Space), according to the themes that have already been defined either by the circulation of the agenda, or just simply by a demand (formal or not) generally provoked by the project director about a particular problem. The agenda is a classic procedure but is necessary to guarantee the mutual awareness of each one's tasks and the evolution of their role in the process. During these meetings, according to the project or objectives, negotiations take place between the actors to defend their choices or their own objectives. This exchange of viewpoints is a part of the construction of new shared knowledge to result in a compromise between the actors and their common objective. The different propositions generated are often followed by opinions and arguments from one group and another to justify them or to put forward others. Taking different forms – analytical, comparative or analogic [33] – these evaluations take place at key moments in the design making it possible to develop an iterative process of collaboration and to introduce the next subject to deal with. They are dictated by permanent research of compromise, where three categories of interaction are revealed [26]:

- Interactions dedicated to the collaborative design process of the architectural object – linked to an awareness activity (orientation of the building, its dimensions, the site and the functionality, etc.);
- interactions dedicated to the situation of the collaboration – linked to social awareness (definition of the design context, work procedures, sharing methods, communication tools, etc.);
- interactions dedicated to the actors – linked to action awareness (knowledge of the actors, experience, competence, logical actions, roles, tasks, organization, coordination, etc.).

Most of the existing tools manage one or another of the operations used in negotiation, evaluation, or questioning by consensus without real assistance enabling them to be linked. Let's take the example of the Chantier.com tool [34] set up by AIA to present an entry point on the Internet to exchange documents between different project actors. The uploading and downloading of the files are the principal function defining this space. However, the accumulation of this data and certain versions can lead to confusion because, in the observed version, there is no real hierarchy between the data except for their listing and their date of insertion in the site. They also create complications through a lack of coherence between the tools and the way teams function. Just like this example, most of the groupware currently used in the agencies, only partially enable synchrony between the actors collaborating from a distance, increasing the misunderstanding between them and decreasing their interactions. Nevertheless, taking charge of the heterogeneity between the actors, their specialties, and their references can be managed by procedures and, sometimes, imposed by norms. The procedures are often unique to the group of designers according to their tasks, their pertinences, preferences, knowledge or personal experience. When they are imposed according to the norms, they must also allow for some gradual evolution of the collaborative process to not restrict the group. The organization of the data to be treated in a hierarchy, by the actors, during the project is also necessary to manage the negotiation and evolution process during the collaborations. Deciding on the level of priority of one point of view or another enables one to consider each criterion in the design and to decide to reject, or to suspend the incompatibility of others.

VII. CONCLUSION.

This report on practices has allowed us to qualify the context in which different groupwares, group support and strategy are integrated and set up by the agencies in order to aim for better management of knowledge and more efficient and productive interaction. It is clear that there are no methods or tools that are perfectly adapted to the context, especially in the case of synchronous and remote collaboration. The research attempts that develop specific tools for collective activity are not yet compatible with the constraints and the reality of the practice in the agency. These discrepancies can be explained because the tools are often developed for other activity sectors without really focusing on the specificities of each of them. And in the case where they claim to be adapted to a particular design activity, they are only adapted to the advanced phases of the process where the choices concerning the project have been previously defined by the group of actors.

This way, we can conclude by listing the need of agencies to go from tools that are thrown together towards the construction of strategies and procedures which enable:

- The management of the processes of negotiation, evaluation, calling into question the first phases of design: the objective being to encourage reflection;
- to take into account the multiplication of exchange places and the passage between them (I-Space, We-Space, Space-Between): the objective being to manage a dynamic in motion;
- to enable the diversity of representations and their transformations from one format to another: the objective being to enable reflexive exploration;
- to assure the synchrony between actors collaborating remotely, as well as the passage between moments of collaboration and moments of necessary cooperation in the current activities of the agencies: the objective being to allow a strategy that evolves, explores, and is flexible implying the object to be planned, the group and the tool.

REFERENCES

[1] A. Farel, “Conception d'un bâtiment : l'organisation d'un travail collectif,” Concevoir, Inventer, Créer, L'Harmattan, Paris, pp. 51-63, 1995.

[2] L.L. Bucciarelli, “Between thought and object in engineering design, ” Design Studies, vol. 3. 23, pp. 219-231, 2002.

[3] W.J. Mitchell, “Challenges and opportunities for remote collaborative design,” Collaborative design and learning: Competence building for innovation, International series on technology policy and innovation, Praeger, Conn, Westport, pp. 5-12, 2004.

[4] M.L. Maher, M. Paulini, and P. Murty, “Scaling up: From individual design to collaborative design to collective design,” Proceedings of Design Computing and Cognition‘10, Stuttgart, 2010.

[5] J. Grudin, “Computer-Supported Cooperative Work: History and Focus,” Journal, IEEE Computer, 27 (5), pp. 19-26, 1994.

[6] S. Ben Rajeb, “Modélisation de la collaboration distante dans les pratiques de conception architecturale : Caractérisation des opérations cognitives en conception collaborative instrumentée,” PhD thesis in Architecture, ENSA Paris-la-Villette, 2012.

[7] M. Brydon-Miller, D. Greenwood, and P. Maguire, Why action research? Action Research, vol.1. 1, pp. 9-28, 2003.

[8] L. Karsenty and B. Pavard, “Différents niveaux d’analyse du contexte dans l’étude ergonomique du travail collectif,” Réseaux, vol. 85, pp. 73-99, 1997.

[9] A. Jeantet, “Les objets intermédiaires dans la conception. Eléments pour une sociologie des processus de conception,” Sociologie du travail, vol. 3, pp. 291-316, 1998.

[10] B. Maggi, (1996). “La régulation du processus d’action de travail,” Traité d’ergonomie, Octarès, Toulouse, pp. 637-662, 1996.

[11] P.A. Gloor, Swarm Creativity: Competitive Advantage through Collaborative Innovation Networks, Oxford University Press, Oxford, 2006.

[12] D. Fay and M. Frese, “Self-starting behavior at work: Toward a theory of personal initiative,” Motivational psychology of human development: Developing motivation and motivating development, Elservier, Amsterdam, pp. 307-337, 2000.

[13] B. David, “IHM pour les collecticiels,” Réseaux et Systèmes Réparties (RSR-CP), vol. 13, Hermes Science, pp. 169-206, 2001.

[14] Y. E. Kalay, “Architecture's New Media. Principles, Theories, and Methods of Computer-Aided Design,” The MIT Press : Communication, MIT Press, Cambridge, pp. 83-198, 2004.

[15] A. M. Maier, “A grid-based assessment method of communication in engineering design,” PhD thesis, Department of Engineering, Selwyn College, University of Cambridge, Cambridge, 2007.

[16] S. Johansen, “Statistical analysis of cointegrating vectors,” Journal of Economic Dynamics and Control, vol. 12, pp. 231-254, 1988.

[17] C. A. Ellis, S. J. Gibbs, and G. Rein, “Groupware: some issues and experiences,” Communications of the ACM, vol. 1. 34, pp. 39-58, 1991.

[18] W.W. Gaver, “The affordances of media spaces for collaboration,” Proceedings of CSCW'92, ACM, New York, PP. 17-24, 1992.

[19] M. Favier, F. Coat, J. C. Couron, and J. Trahand, Le travail en groupe à l’âge des réseaux, Collection : Gestion, Série : Production et techniques quantitatives appliquées à la gestion, Economica, Paris, ISBN: 2717837671 - 9782717837674, 1998.

[20] F. Darses, A. Mayeur, S. Ben Rajeb, C. Lecourtois, F. Guéna, S. Safin, and P. Leclecrq, Rapport final, Lot 2.4, Livrable projet ANR CoCréA, 2011, unpublished.

[21] K. A. Ericsson and H. A. Simon, Protocol Analysis: Verbal Reports as Data, MIT Press, Cambridge, 1993.

[22] P. Reason and H. Bradbury, Handbook of action research: Participative inquiry and practice, Sage Publications, London, 2001.

[23] J. M. Burkhardt, F. Détienne, L. Moutsingua-Mpaga, L. Perron, S. Safin, and P. Leclercq, “Multimodal collaborative activity among architectural designers using an augmented desktop at distance or in collocation,” Proceedings of ECCE'08, the 15th European conference on Cognitive ergonomics: the ergonomics of cool interaction, ACM, New York, 2008, ISBN/EAN : 978-1-60558-399-0.

[24] A. de Boissieu, F. Guéna, and C. Lecourtois, “Modélisation paramétrique partagée : Le cas de l’utilisation de Digital Project lors de la conception du Pavillon de la Fondation LouisVuitton pour la Création (Gehry Partners) sous l’angle des opérations de découpage,” Actes du Scan'10 : Espaces collaboratifs, Marseille, 2010.

 ISBN: 978-1-61208-382-7

[25] A. Hatchuel, “Apprentissages collectives et activités de conception,” Revue Française de Gestion, vol. 99, pp. 109-120, 1994.

[26] J. M. Carroll, D. C. Neale, P. L. Isenhour, M. B. Rosson, and D. S. McCrickard, “Notification and awareness: synchronizing task-oriented collaborative activity,”, International Journal Of Human-Computer Studies, vol. 58, pp. 605-632, 2003.

[27] F. Darses, F. Détienne, and W. Visser, “Les activités de conception et leur assistance,” Ergonomie, Presses Universitaires de France, Paris, pp. 545-563, 2004.

[28] P. Boudon, P. Deshayes, F. Pousin, and F. Schatz, Enseigner la conception architecturale : cours d’architecturologie, Editions de la Villette, Paris, 2000.

[29] S. Ben Rajeb, “Conception collaborative distante : étude architecturologique pour la caractérisationdes opérations cognitives,” Interfaces numériques, vol. 2. 3, Lavoisier, 2013.

[30] L. Suchman, “Constituting shared workplaces,” Cognition and communication at work, Cambridge University Press, Cambridge, pp. 35-60, 1996.

[31] S. Ben Rajeb and P. Leclercq, “Using Spatial Augmented Reality in Synchronous Collaborative Design,” Lecture Notes in Computer Sciences, Springer, vol. 8091, pp. 6-15, 2013

[32] N. Bonnardel, “Activités de conception et créativité : de l’analyse des facteurs cognitifs à l’assistance aux activités de conception créatives,” Le travail humain, Presses Universitaires de France, vol. 72, 2009, http://www.cairn.info/article.php.

[33] G. Martin, F. Détienne, and E. Lavigne, “Confrontation of viewpoints in a concurrent engineering process,” Integrating design and manufacturing in mechanical engineering, Kluwer Academic Publishers, London, pp. 3-10, 2002.

[34] Wapp 6: conject group. *DOC 6: Piloter les informations de vos projets de construction*. [Online – last access date : January, 2015]. http://www.wapp6.com/doc6.php.

Icons++: An Interface that Enables Quick File Operations Using Icons

Xiangping Xie
Department of Computer Science
University of Tsukuba
Tsukuba, Ibaraki, Japan
Email: xie@iplab.cs.tsukuba.ac.jp

Jiro Tanaka
Department of Computer Science
University of Tsukuba
Tsukuba, Ibaraki, Japan
Email: jiro@cs.tsukuba.ac.jp

***Abstract*—In graphical user interfaces, when users want to operate on a file, they usually double-click the file icon to launch the associated application and open the file. Several file operations are available in the context menu by right-clicking, such as printing and deleting. However, if the user wants to perform some simple operations, such as copying the file contents, opening the file by an application and then selecting a menu, can be cumbersome. Furthermore, operations in the context menu are limited. Thus, in this paper we present Icons++, a user interface which allows users to perform the file operations they want in a quick way by using icons. Through the use of Icons++, users can take a quick look at the file contents, and at the same time they can perform often-used file operations with only one click, without opening the file by a relevant application. In this paper we present our design of Icons++ and the user studies we performed in order to evaluate it. Studies' results show that using Icons++ is 53% faster than using an application to execute the same task, and our interface is preferred by participants.**

Keywords*— *icon interface; preview; file operation; file manager.

I. Introduction

Clicking a file icon to launch an application that opens the file is commonly used for accessing files in Graphical User Interfaces (GUIs). File icons allow users to browse files [1], and figure out the file types. Thus, it can be said that icons have an excellent browsability. On the other hand, applications allow users to perform a wide variety of file operations, so users can edit files using those applications. It can be said that applications have an excellent operability.

As can be seen from shortcut keys or mouse gestures, users have a constant need for carrying out often-used operations in an easy manner [2]. As a method of taking a shortcut to operations on menu selection, Appert et al. investigated using stroke gestures as command shortcuts [3]. However, it imposes memorizing gestures' patterns on users. Some researches accentuate items in a menu, e.g., Ephemeral Adaptation [4] and Bubbling Menus [5], though these techniques require tracking many steps on menu selection to reach the desired item. In addition, the following instances show that users also have a need for knowing the file contents without actually opening the file: first, using thumbnails of file contents as file icons is common. For example, a thumbnail for the first page of a PDF file is used as the icon of that PDF file. Also, in the current User Interface (UI) of Gmail, there is a preview function, which shows the contents of the attached files without using an application to open the files.

Therefore, we propose a new user interface which displays the file contents and allows users to execute operations on files, so that they can perform common file operations in a quick way. In this paper, we introduce an interface called Icons++ that has both browsability and operability. Icons++ is located between the icon and the application (Figure 1).

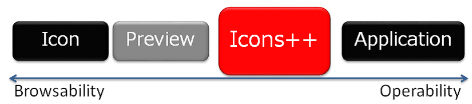


Fig. 1. Icons++ is placed between the icon and the application

As shown in Figure 1, we assign browsability and operability to the horizontal axis. File icons allow users to look at many files at the same time and know the file types, so the use of icons increases the browsability. Preview function shows further file information and allows users to scroll pages, so Preview is next to Icon on our axis. Since applications allow users to change file contents and operate on files significantly, the Application contributes to a high operability. Our proposed interface, Icons++, lies midway between the Icon and the Application.

The following section describes our Icons++ interface and interaction. Section 3 shows a usage scenario to reveal the usefulness of Icons++. In Section 4, we explain how we implemented Icons++. We then present two studies and a questionnaire about Icons++. The studies show that Icons++ is useful on file operations. Finally, we review related work on hovering, using icons for operations, and visual feedback on icons and thumbnails.

II. Interface Design for Icons++

In order to allow users to operate on files in a quick way, we first display the file contents in what we called "Contents View", so that users can confirm whether it is the right file. In Contents View, there are "Operation Icons" which make file operations possible with only one-click. Moreover, in order to find the file contents easier, users can add a mark on thumbnails in Contents View.

A. Contents View by Hovering

Users can activate Icons++'s Contents View by using hovering. Hovering is the action that occurs when the mouse cursor

 ISBN: 978-1-61208-382-7

is placed over the object. Using hover, users do not need to memorize complicated operations like mouse gestures. Moreover, it has the potential to make the user notice Icons++'s functionalities naturally. Hover is uncompetitive with other mouse events, such as a right-click to open a pop-up menu or a double-click to open the file in the application. Thus, it allows users to use existing mouse events as usual.

In Icons++, when users hover over a file icon for about one second, Contents View is automatically displayed near the file icon (Figure 2a). Users can look at each page of the file in Contents View. Contents View disappears when the mouse cursor leaves the file icon. Contents View becomes fixed when users click the file icon. Application icons become available when users move the mouse cursor on top of Contents View (Figure 2b left). Selecting one application icon, the file will be opened by the application (Figure 2b right).

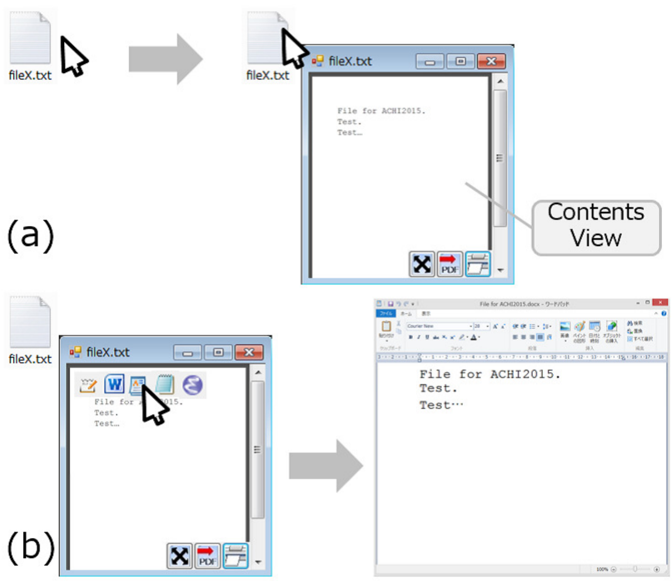


Fig. 2. (a) Hover over the file icon to display Contents View (b) Hover over the top of Contents View to show application icons and open the file by an application

Thus, by using hover, users can access Icons++ from a file icon, and can then access applications from Icons++. Therefore, it can be said that Icons++ is located between file icons and applications, and it links the file icons with applications. Unlike the traditional way, one file type can be linked with multiple applications.

B. File Operations by Using Operation Icons

Using Operation Icons in Contents View of Icons++, users can perform some often-used file operations with only one-click, as opposed to using many steps when using an application. As shown in Figure 3, there are three Operation Icons at the bottom of Contents View. Users click an operation icon, then the file operation is executed (Figure 3).

File operations include making a different type of file (e.g., PDF or plain text) based on the original file, showing in full-screen, displaying as slide show, printing, etc. Figure 4 shows some examples for Operation Icons. Operation Icons

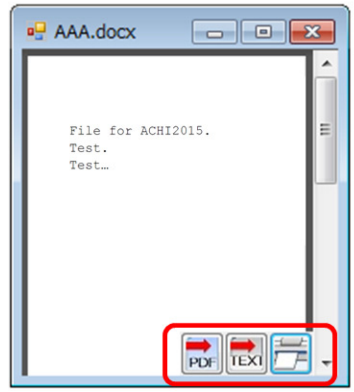


Fig. 3. Operation Icons on the bottom of Contents View

Fig. 4. (a)Full-screen (b)Slide show (c)Make PDF (d)Make TEXT (e)Print

in Contents View are dynamically changed according to the file type. With regard to visibility for the file contents, the maximum number of Operation Icons in Contents View is 6, which is the number of icons that can be shown in one line.

Besides the default Operation Icons in Icons++, users can make icons for custom operations. Custom Operation Icons are made when users record the operations they want, then press the custom operation icon to repeat the recorded operations.

Thus, Icons++ has an improved operability, because, with only one-click, Operation Icons enable users to perform file operations, which they would traditionally perform in multiple steps.

C. Mark File Pages by Dog-ear and Folded Page

In Contents View of Icons++, we make it possible for users to add a mark on the thumbnail of the page, so that they can easily find the file or page they want. As shown in Figure 5, Dog-earing is used for the thumbnails of important pages (first page in Contents View in Figure 5). Folded page is used for the thumbnails of insignificant pages (second page in Contents View in Figure 5). For example, if a report file has a part that the user wants to review, the user can add a dog-ear on the pages that have that part. On the other hand, if the title page of the report has little information, the user can fold the thumbnail of the title page to make it less noticeable.

We explain the creation of dog-ears and folded pages below.

1) Dog-eared Page: Dog-earing is an action that folds a corner of a page to create a triangle shape. In Icons++, clicking the upper right corner of the thumbnail adds a dog-ear. Single-clicking adds a small dog-ear, while double-clicking creates a big dog-ear. Clicking the area which already has a dog-ear can remove this dog-ear by flattening the dog-eared area. The page that has a dog-ear will be the page that users see first in Contents View, thus finding the important page becomes easier.

 ISBN: 978-1-61208-382-7

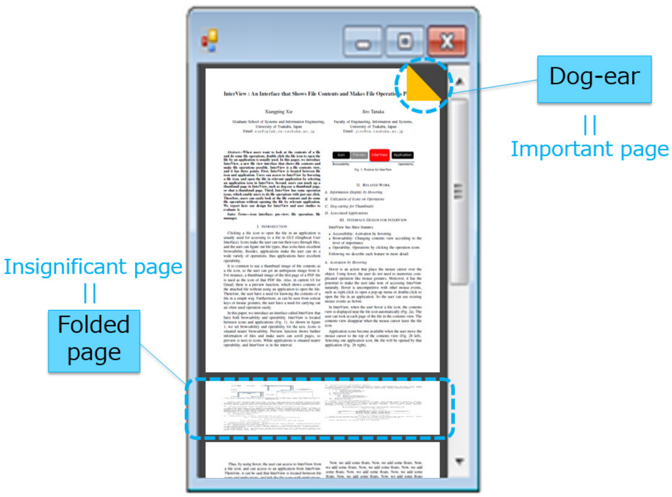


Fig. 5. Dog-ear and folded page for thumbnails in Contents View

2) Folded Page: Folded page is a thumbnail which is a quarter in size of the original size. Double-clicking the thumbnail of the page transforms it into a folded page. Similarly, double-clicking a folded page takes the thumbnail back to the original size.

We can thus state that Icons++ has an improved browsability because users can check the file contents at a glance in Contents View, and they can set the importance for any page using dog-ears and folded pages.

III. Usage Scenario

In this section we will introduce a scenario for using our interface. Let us assume that, after a meeting, user A wants to send the minutes of the meeting, recorded as plain text, via e-mail. Taking into account possible character corruption, user A is going to change the minutes written in plain text into a PDF file and attach it to the e-mail.

In the traditional way, user A should perform the following procedures: (1) use the context menu by right-clicking the minutes file icon in the file manager, (2) open the file in WORD, (3) click menu in WORD, (4) select "Save As", (5) change file type to PDF, (6) make a PDF file, (7) attach the PDF file to the e-mail.

When using our Icons++, the procedure becomes as follows: (1) hover over the minutes file icon in the file manager to show Contents View, (2) click the file icon to make Contents View fixed, (3) select "make PDF" Operation Icon, (4) make a PDF file, (5) attach the file to the e-mail. Thus, user A can perform the file operation in a shorter way without going through an application.

IV. Software Implementation

We created a prototype as a file manager for Icons++ in C# under Windows 7. We made it based on the close image of the file manager in Windows, hence it is expected to be easy for the user to adapt to Icons++.

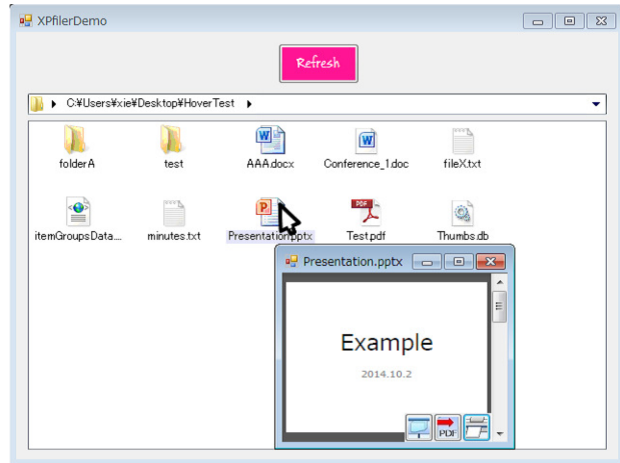


Fig. 6. UI for the prototype of Icons++

A. UI for the Prototype

The user interface for the prototype of Icons++ is shown in Figure 6.

Figure 6 shows the user hovering over a PowerPoint file icon with the mouse cursor; the Contents View for this file becomes visible near the file icon.

B. Contents View

Icons++ keeps JPEG files corresponding to each page of the user's files. Therefore, when the user hovers over a file icon, the corresponding JPEG files are immediately displayed, so that the user can instantly check the contents of that file.

When there is a newly-created file, and the user hovers over that file icon in Icons++, our system will generate JPEG files corresponding to that file before Contents View is displayed. If the file is large, e.g., a PowerPoint file that has 20 slides, then it takes a little while to generate JPEG files corresponding to that file.

In Icons++, we also have a folder which stores dog-eared JPEG files and folded JPEG files. If a file is hovered over in Icons++, the system first checks the contents of this folder. If there is a dog-eared JPEG file for the hovered file, Icons++ displays the dog-eared page as the starting page in Contents View.

Contents View has four layers, as shown in Figure 7.

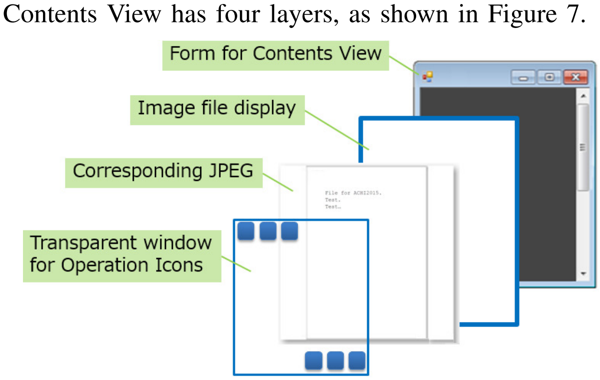


Fig. 7. Four layers for Contents View

 ISBN: 978-1-61208-382-7

When Icons++ shows a JPEG file in Contents View, the system first shows a form window. Then it adds a layer of image file display. It then adds a JPEG file corresponding to the hovered file on the image file display. Finally, the system adds a transparent window for Operation Icons on the top in a way that allows the user to perform file operations.

Because of the use of JPEG for showing the contents of the file, Contents View of Icons++ can prevent character corruption, regardless of the format of the original file.

C. File Operation

We conducted a user survey on file usage before implementing the prototype. We asked 9 students (1 female, 8 males, ranging from 20 to 24 years old) about the functions in file operations that they often use and that they would like to use easily. Based on their responses, we implemented 6 operations that had the highest number of responses (some are shown in Figure 4).

We preliminarily registered commands which correspond to Operation Icons in Icons++. When users click an Operation Icon in Contents View, the system activates the pre-registered command which corresponds to that Operation Icon, and then executes the file operation.

In the prototype of Icons++, the targeted file type is a document file, such as plain text, source code (e.g., program file, LaTex file), Microsoft office file, PDF, etc.

To enable the customization of file operations, Icons++ allows users to record operations they want to easily use. A sequence of shortcut keys and menu selections in applications will be recorded in the system. When users select a custom operation icon in Contents View, recorded operations will be repeated automatically.

V. Evaluation

To better understand the benefits of Icons++ and compare Icons++ with the UI of an existing file manager, we conducted two user studies and a questionnaire survey.

Seven computer users (2 females, 5 males), ranging in age from 22 to 27 (mean 24), participated in the studies.

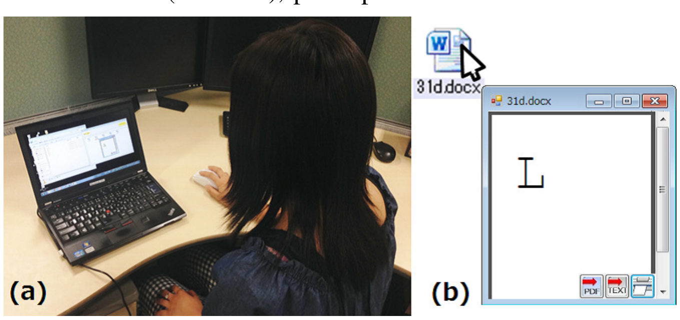


Fig. 8. (a)A participant taking part in the study (b)File with just one alphabet letter

A. Procedure

We first asked the participants to use Icons++ for 3 minutes without any explanation. Then we explained the use of Icons++ and all of its functionalities. After that, we let participants use Icons++ for about 3 more minutes so that they get familiarized with its usage. After the participants got used to Icons++, we asked them to perform two user studies (Figure 8a). Finally, we asked them to fill out a questionnaire survey. We will describe the two user studies and the questionnaire survey in detail in the following subsections.

B. User Study 1: Finding a File Containing a Designated Alphabet Letter

The goal of this study is to compare Icons++ with the file manager of Windows 7. The task of this study is to find a file containing an alphabet letter designated by the system.

We prepared files containing just one alphabet letter (Figure 8b). In total, we had 3 approaches to find the file: (1) using double-click to open the file in the application, (2) using the preview function in the file manager of Windows 7, and (3) hovering over the file icon in Icons++. Participants had to perform 5 trials for each approach. In each trial, participants had to find one file with an alphabet letter designated by the system amongst 5 files. Thus, participants performed the file finding task in 5 files $\times$ 5 trials $\times$ 3 approaches (in total, 75 trials per participant). The system calculated the time for each finding task.

C. User Study 2: Look at a PowerPoint File in Slide Show

The goal of this study is to compare Icons++ against the traditional model for file operation. In this study, participants performed the same operation using two approaches: (1) using the method they usually use, (2) using Operation Icons in Icons++. The task for this study is to look at a particular PowerPoint file in slide show. There was no limitation on the methods they usually use. These methods could be using the menu or the icon in PowerPoint, using a shortcut key, using the context menu by right-clicking on the file icon, etc. Each participant performed this task once for each approach. The system calculated the time for each approach.

D. Questionnaire Survey

After all user studies were completed, we asked participants to answer a questionnaire about the usability of Icons++. We used a five-point Likert scale, in which 1 corresponds to “strongly disagree” and 5 corresponds to “strongly agree”.

First, we asked the participants about accessibility in each approach of User Study 1 (i.e., “whether they know how to use the interface without explanation”). Next, we asked the participants about the usability of Icons++ and whether they want to continue to use Icons++: (1) Is Contents View by hovering easy to use? (2) Are file operations by Operation Icons easy to use? (3) Is the overall UI for Icons++ easy to use? (4) Do you want to continue to use Contents View by hovering? (5) Do you want to continue to use file operations by Operation Icons? (6) Do you want to continue to use the overall UI for Icons++? In addition, we offered the participants the opportunity to write their comments freely.

VI. RESULTS

Figure 9 shows the average completion time for 7 participants with each approach in User Study 1. The left side in Figure 9 shows the mean of total time for 5 trials by 7 participants for each approach, and the right side in Figure 9 shows the mean time of each trial by 7 participants for each approach.

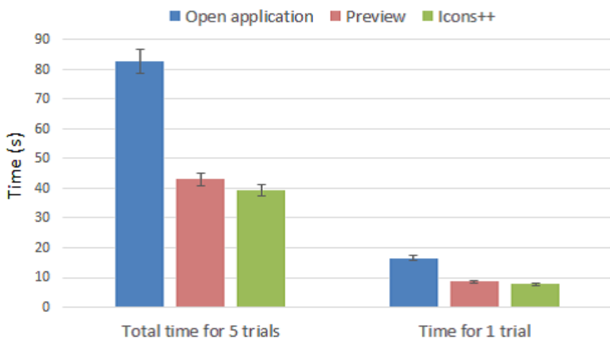


Fig. 9. Average completion time by 7 participants for trials with opening file by an application, preview, and Icons++

The mean time to perform 5 trials with hovering over the file icon in Icons++ is 39.2 seconds (s.d. 5.7), 53% faster than the time opening the file by an application (mean 82.7, s.d. 18.1): $T(12) = 5.6$, $p < .001$. Thus, based on the results, the time for trials with Icons++ shows a significant difference compared with the method of using double-click to open in an application.

On the other hand, the mean time to perform 5 trials with hovering over the file icon in Icons++ is 9% faster than the time using preview in the file manager of Windows (mean 43.1, s.d. 8.6): $T(12) = 0.7$, $p = 0.2$. Hence, there is no significant difference in time with Icons++ compared with the method of using preview function in the file manager.

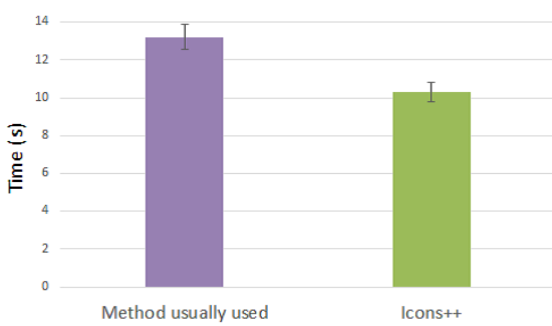


Fig. 10. Average completion time by 7 participants to open a PowerPoint file in slide show with the method they usually used and with Icons++

Figure 10 shows the average completion time for 7 participants for User Study 2. The time to open a PowerPoint file in slide show with Icons++ is 10.3 seconds, 22% faster than the time for using the method they usually used, which is 13.2 seconds: $T(12) = 0.9$, $p = 0.19$. In terms of methods the participants used, some of them opted for menu selection on PowerPoint, some used the icon at the bottom of the window of PowerPoint, and some used a shortcut key (i.e., F5). No one used the context menu by right-clicking on the file manager.

Based on the results for the questionnaire survey, participants responded more favorably to the question "Hovering over the file icon in Icons++ is a more intuitive manipulation" (7 above neutral, none below) than for the equivalent question with double-click to open in an application (6 above neutral, 1 below) and with preview function (4 above neutral, 3 below). Thus, although the time for the same task shows little difference in using the preview function and using hover over in Icons++ (as shown in the result of User Study 1), participants preferred Icons++ to the preview function in terms of operability.

Regarding the UI of Icons++, participants also felt it was easy to use Contents View by hovering over (6 above neutral, 1 below), file operation by Operation Icons (7 above neutral, none below), and the overall UI for Icons++ (6 above neutral, 1 below). When asked whether they want to continue the use, almost all participants agreed to Contents View by hovering over (6 above neutral, 1 below), file operation by Operation Icons (7 above neutral, none below), and the overall UI for Icons++ (7 above neutral, none below).

According to the participants' comments, we found that they felt positive about Icons++. Some participants said that Icons++ is access-friendly because they do not need to learn how to use it. Some participants also pointed out that they found Operation Icons useful because they could perform operations easily. However, there were a few comments about the design improvement of Icons++. Two participants indicated that Icons++ had a drawback: Contents View covered up the lower file icons when it appeared. One participant mentioned that if he really wanted to look at the file contents, Contents View of Icons++ is too small to be legible. Another participant expressed that although Icons++ is useful, the automatic appearance of Contents View is a nuisance if he does not have the intention of knowing the file contents.

VII. DISCUSSION

Compared with existing approaches, such as launching applications, or using the preview function, Icons++ is more intuitive. Participants were able to find out how to use Icons++ without any explanation. One of the main reasons why Icons++ has a better usability is the quick display of the file contents by hovering, which is faster than opening files by an application.

Since hovering is an action that naturally precedes a click (to select) or a double-click (to open the file), Icons++ has an advantage: participants can naturally discover the display of the file contents when hovering over an icon. Results for two user studies and a questionnaire indicated that users preferred simple ways of use, such as hovering, or icons with only one click.

However, the fast display of Contents View is likely to be impeditive when the user has no intention of knowing the file contents. Therefore, there is still some room for improving the design of Icons++ to make it more practical. Some possible solutions could be setting a time limit on the appearance of

Contents View, or making Contents View translucent. In our approach, we appended a slider to Icons++ window, which can turn On or Off the display of Contents View, and which can allow setting its on-screen time.

VIII. Related Work

Many researches use hovering for displaying information near the hovered object. For example, Fitchett et al. proposed a UI named Hover Menu [6] in the file manager, which shows the list of commonly accessed items inside hovered folders. Terry et al. presented Side Views [7], which provided a dynamic, on-demand preview of a command applied to a copy of the data when a menu was hovered over. CommunityCommands [8] by Matejka et al. and Share and Share Alike [9] by Voida et al. display information about the hovered file, such as notes written by co-workers or members who share the file. These researches are related with Icons++ in terms of displaying information by hovering. However, Icons++ is different in that Icons++ allows users to perform file operations and to access an application in Contents View.

Several researchers utilize icons on operations. Touch-Display Keyboards [10] by Block et al. shows an environment using icons for file operations by projection on the keyboard of the computer. Users can select operations by pressing a key. Sikuli [11] by Yeh et al. and the research by Chang [12] used graphical icons in screenshot-based scripts. Users can use screenshot patterns, shown as icons in the scripts, to perform mouse and keyboard events. Unlike these, Icons++ uses the mouse action, which is a conventional method. In addition, Icons++ does not only perform actions on one file, but also allows the users to perform file operations on two or more files in the same file manager's window.

In order to allow users to quickly find an item or a file they want, there are some techniques proposed in existing research. One of these techniques is to add some visual feedbacks on the file icon or file thumbnails for accentuating the target. For example, Fitchett et al. showed Finder Highlights [13], a file browser which has a function of highlighting file icons. There are proposals that use dog-ears to emphasize the target, such as NiCEBook [14], WebView [17], research by Kaasten et al. [15], and research by Hoeben et al. [16]. Unlike these works, we not only add a visual feedback (i.e., dog-ear) for accentuating the object, but we also propose a method to make information less noticeable (i.e., folded page).

Some applications and functions are related to our approach in Icons++. For example, the Quick Look function on Mac OS is similar to Icons++ in that it shows file contents. However, there is no Operation Icon in Quick Look function; it exists for browsing only. Furthermore, the way to access Quick Look is to press the space bar on the keyboard. We cannot state that pressing the space bar is an intuitive manipulation, because, if unknown beforehand, it is difficult to find it naturally. Another example is the preview function in Gmail. The difference between Gmail preview and Icons++ is the tackling of character corruption. In Gmail preview, sometimes characters are not displayed correctly, while this does not happen in Icons++, because Icons++ uses JPEG files.

IX. Conclusions and Future Work

We presented Icons++, an icon based user interface for viewing file contents and manipulating files. Icons++ is able to show file contents quickly when users hover over a file icon. Moreover, it allows users to perform file operations by Operation Icons with only one click. We conducted two user studies and a questionnaire survey for comparing Icons++ against existing approaches. Based on these results, it was found that, first, using Icons++ for checking the file contents was 53% faster than using an application to open the file. Second, using Operation Icons to perform the same file operation was 22% faster than the common methods employed by users. Furthermore, the participants' comments pointed out several issues needing improvement.

In our future work, we expect to address the issue of the file contents not being clearly viewed because of the relatively small scale. Also, we will assess some solutions for the problem of the lower file icons being covered up by Contents View.

References

[1] D. Barreau and B. A. Nardi, Finding and Reminding: File Orgamization from the Desktop. ACM SIGCHI Bulletin 27.3, 1995, pp. 39-43.

[2] J. Tidwell, Designing interfaces. O'Reilly Media, Inc., 2010.

[3] C. Appert and S. Zhai, Using Strokes as Command Shortcuts: Cognitive Benefits and Toolkit Support. In Proc. CHI '09, 2009, pp. 2289-2298.

[4] L. Findlater, K. Moffatt, J. McGrenere, and J. Dawson, Ephemeral Adaptation: The Use of Gradual Onset to Improve Menu Selection Performance. In Proc. CHI '09, 2009, pp. 1655-1664.

[5] T. Tsandilas and M. C. Schraefel, Bubbling Menus: A Selective Mechanism for Accessing Hierarchical Drop-Down Menus. In Proc. CHI '07, 2007, pp. 1195-1204.

[6] S. Fitchett, A. Cockburn, and C. Gutwin, Improving Navigation-Based File Retrieval. In Proc. CHI '13, 2013, pp. 2329-2338.

[7] M. Terry and E. D. Mynatt, Side Views: Persistent, On-Demand Previews for Open-Ended Tasks. In Proc. UIST '02, 2002, pp. 71-80.

[8] J. Matejka, W. Li, T. Grossman, and G. Fitzmaurice, Community-Commands: Command Recommendations for Software Applications. In Proc. UIST '09, 2009, pp. 193-202.

[9] S. Voida, W. K. Edwards, M. W. Newman, R. E. Grinter, and N. Ducheneaut, Share and Share Alike: Exploring the User Interface Affordance of File Sharing. In Proc. CHI '06, 2006, pp. 221-230.

[10] F. Block, H. Gellersen, and N. Villar, Touch-Display Keyboards: Transforming Keyboards into Interactive Surfaces. In Proc. CHI '10, 2010, pp. 1145-1154.

[11] T. Yeh, T. Chang, and R. C. Miller, Sikuli: Using GUI Screenshots for Search and Automation. In Proc. UIST '09, 2009, pp. 183-192.

[12] T. Chang, Using Graphical Representation of User Interfaces ad Visual References. In Proc. UIST '11, 2011, pp. 27-30.

[13] S. Fitchett, A. Cockburn, and C, Gutwin, Finder Highlights: Field Evaluation and Design of an Augmented File Browser. In Proc. CHI '14, 2014, pp. 3685-3694.

[14] P. Brandl, C. Richter, and M. Haller, NiCEBook-Supporting Natural Note Taking. In Proc. CHI '10, 2010, pp. 599-608.

[15] S. Kaasten and S. Greenberg, Integrating Back, History and Bookmarks in Web Browsers. In Proc. CHI '01, 2001, pp. 379-380.

[16] A. Hoeben and P. J. Stappers, Flicking through Page-based Documents with Thumbnail Sliders and Electronic Dog-ears. In Proc. CHI '00, 2000, pp. 191-192.

[17] A. Cockburn, S. Greenberg, B. McKenzie, M. Jasonsmith, and S. Kaasten, WebView: A Graphical Aid for Revisiting Web Pages. In Proc. OzCHI '99, 1999, pp. 15-22.

Designing an Adaptive User Interface According to Software Product Line Engineering

Yoann Gabillon and Nicolas Biri and Benoît Otjacques

Luxembourg Institute of Science and Technology (LIST)
5, avenue des Hauts-Fourneaux
L-4362 Esch/Alzette
Email: {yoann.gabillon, nicolas.biri, benoit.otjacques}@list.lu

***Abstract*—An adaptive User Interface (UI) is a UI that is able to adapt itself to a change of the context of use (user, platform, environment). Designing an adaptive UI remains a difficult and time consuming task that needs the use of common and variability parts between the different UI adaptations. Software Product Line (SPL) engineering is a software engineering approach that aims to develop a collection of similar software systems by using software assets and a variability model: the Feature Model. Dynamic Software Product Line is an adaptation of the SPL approach in order to design an adaptive software system. This paper proposes a method to design an adaptive UI according to a DSPL process. This method is implemented through the UI ADAPTOR prototype. This first implementation underlies the several benefits of the proposed method.**

Keywords–Adaptive User Interface; Software Product Line; Dynamic Software Product Line; Feature Model; Context of use.

I. INTRODUCTION

With the increasing amount of devices and mobile platforms as well as new user profiles, designers need to design a software system adapted to the current context of use. A software system encompasses functional core and User Interface (UI) parts. An adaptive User Interface (adUI) is a UI that is able to adapt itself or to be adapted, automatically at run time, to a context change. A context of use varies according to the properties regarding the *users*, the *platforms* and the *environment* of interaction. Because there are a huge amount of possible combinations, designing an adUI is a difficult and time consuming task.

Model-based approaches for the UI development (MBUID) are mainly used in adUI development in order to decrease the development costs. MBUID approaches promote the modelling of adaptation rules to make the adUI evolve according to a change of context [1] [2]. The designer must understand and identify the adaptation rules (variations) and its effects [3], i.e., the adaptation rules must be transparent (understandable and reusable) for developers. However, the huge amount of possible combinations make it difficult to understand and lead to a lack of reusability and design errors.

Software Product Line Engineering (SPLE) aims to develop a collection of similar software systems from a shared set of software assets. This approach is used to design a set of software systems that encompass common and variability parts by using a model of variability. Indeed, the variability model helps the designer to understand variations and reuse concepts and tools to check consistency of variations.

In order to decrease the development cost of MBUID approaches and to increase the understandability/reusability of adaptation rules, this paper provides a method to design an adUI according to SPLE. This method allows to reuse concepts and tools proposed by SPLE approach to manage and understand variabilities and its effect.

The paper begins with a background presentation of SPLE including the whole process and feature model. Based on related work on adUI design and on UI design by SPLE (Section 3), the paper proposes a method to design adUI according to SPLE in Section 4. In Section 5, the proposed method is implemented through a prototype called UI ADAPTOR. Based on UI components, this prototype composes a UI adapted to the current context of use. UI ADAPTOR aims to collect experimental lessons and to underlying the many benefits of the proposed method.

II. BACKGROUND

The *software product line (SPL)* development method separates the two following processes: the *domain engineering* and the *application engineering* [4]. Firstly, *domain engineering* is the process which is responsible for establishing the reusable artefacts and thus for defining the commonality and the variability of the product line. All types of software artefacts needed to develop the final products may be developed: requirements, design, realisation, tests, etc. Secondly, *application engineering* is the process which is responsible for deriving product line applications (products) from the artefacts established in domain engineering. A large part of application engineering consists of reusing artefacts of the domain engineering and binding the variability as required for the different applications.

The *Dynamic SPL (DSPL)* development method [4] is an adaptation of the SPL method in order to produce only one adaptive product, instead of a set of products. The DSPL process also separates the requirement engineering that is made at the design time and the domain engineering that is made automatically this time at runtime (i.e., dynamically) according to the current context of use.

Mostly, *Feature Models* (FM) are widely used to model the variability of requirements during the whole SPL and DSPL process. The *Feature Model* was first proposed by Kang and

 ISBN: 978-1-61208-382-7

al. as a part of Feature Oriented Domain Analysis (FODA) study [5]. Since then, feature models have been very popular in software product lines and have been widely accepted and used by the academic and industrial communities, although several dialects exists [6]. A FM is a tree that defines relationships between the features. The four following relationships specify the hierarchical decomposition of a feature into its sub-features: *Mandatory*, *Optional*, *Alternative* and *Or*. *Mandatory* means that the feature is included in every product that includes its parent. *Optional* means that the feature may or may not be included in a product that includes its parent. *Alternative* means that every product that includes the parent must include exactly one feature from the group. *Or* means that every product that includes the parent must include a non-empty subset of the group. Constraints are used to specify cross-tree relationships between features. A constraint consists of a boolean expression. For example, "$A \Rightarrow B$" means that if B is selected during the configuration phase, A must be selected as well.

III. Related work

This section presents a related work overview of approaches in adaptive UI development and in UI development according to SPLE.

A. Adaptive User Interface

According to [7], different software development approaches have already been investigated to design an adUI. For example, MBUID [1] [2] [8] promotes the modelling of transformation, Aspect-oriented programming [9] pushes forward insertion of aspects at runtime or Component-based programming [10] focuses on the (re)composition of available components. These approaches are based on the same principle: an adaptive system is in charge of adapting a UI by adaptation rules according to the current context of use. As reference example, the SERENOA project [3] [2] [11] identifies the main models needed to design an adUI: the UI, the rules of adaptation, the context of use and the adaptor. The adaptor is in charge of applying rules on the UI model according to the context of use to produce a new UI model expressing the adapted UI. The UI model is defined according to the four standard levels of abstraction in MBUID: Task Model, Abstract User Interface, Concrete User Interface and Final User Interface [12].

However, according to [3], a main issue is the lack of inspectability and understandability of adaptation rules that leads to possible side effects and lack of reusability.

B. User Interface development by Software Product Line

Even if the design of adaptive systems by SPLE has already been investigated [13] [14], this work focuses on the adaptation of the functional core part. In contrast, the literature concerning UI design according to SPL engineering focuses on the design of a set of UIs. [15]–[17] propose to model variants of UI features such as different whole UIs. As a consequence, UI variations cannot be reused because they are not traceable or saveable. [18] propose to model variations of interactors. Each interactor that varies is a variant. For example, a interaction variation can have two alternative variants: JComboBox or JList. This modelling has the advantage to be easily inspectable and traceable. [19] [20] argue in favour of the use of the standard levels of abstraction to model variability into the feature model. However, the selection of UI variations is based on functional features. Unfortunately, UI variations are necessary to improve usability even if functional features do not change [21]. In contrast, [22] propose a methodology to design a set of UIs from the selection of UI variants based on a feature model containing the four abstraction levels.

IV. Method to design Adaptive UI by DSPL

An overview of the SPLE-based Adaptive UI design method is sketched in Figure 1. We have adapted the DSPL process to design an adaptive UI. The main idea is to prepare UI components and to model their variations according to the targeted context of use during the domain engineering at design time. Then, at runtime, the UI adaptor selects and composes UI components according to the current context of use.

Firstly, at design time during the requirement engineering, context variations and UI variations must be modelled as features of the FM. The context features represent context variations such as recommended by [23]. The UI features represent variations between UI adaptations such as recommended by [22]. For example, the screen size variation can be expressed as two features: smallScreen and largeScreen. The UI variations can be defined at different abstraction levels. For example, two interactors can be expressed as two features: slider and textfield. In order to express the link between the context features and the UI features, we can add constraints. For example, the slider that is replaced by the textfield for space reasons can be expressed as two constraints: "slider ⇒ largeScreen" and "textfield ⇒ smallScreen".

Secondly, the UI components must be designed corresponding to each feature.

Thirdly, at run time during the application engineering, the appropriate context feature and UI features must be selected according to the current context of use. This selection must be made automatically to promote a context-aware adaptation of the composed UI. Many tools have been developed (such as Feature IDE [24]) in order to automatically selects the UI features according to the constraints and the selected context features.

Finally, the components corresponding to the selected features are composed to automatically design the adUI.

In order to promote context-aware adaptation of the composed UI, an adaptive system in charge of recomposing the adUI. From a context change detection, the UI adaptor updates the context of use model and, in consequence, selects the context feature. The context feature selection leads to a new UI feature selection and a new UI composition to produce the new adUI.

V. Applying the method: Implementation

Example application. Lets suppose a user, Orson, who want to visualize the consumption data of his house. To achieve his goal, he uses an adaptive visualization software called "*VisuData*" (the Figure 2 shows a possible UI of this software). *VisuData* provides four graphics in order to visualise a set a data: a horizontal barchart, a vertical barchart, a pie chart and a bubble chart. The *VisuData*'s user could use three data filters. The first, called "DataDisplayed", allows to choose the data displayed. For example, Orson could select the quantity

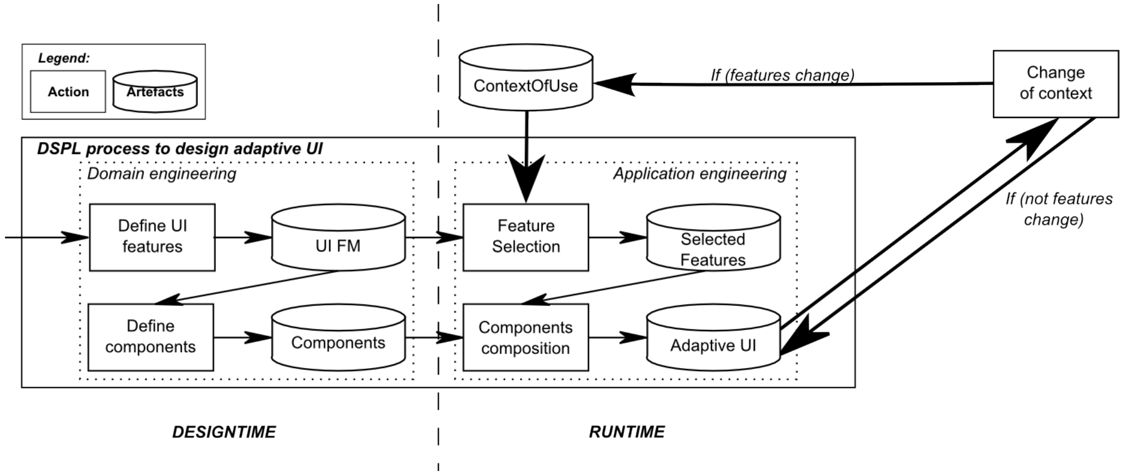


Figure 1. Overview of the Dynamic SPL process to design an adaptive UI.

of electric energy or the price consumed by the radiators. The second, "TimeLine", allows to select the time interval of the data displayed. For example, Orson could display the consumption between 2010 and 2012. The last filter ("Filter 1D") allows to select the value interval of the data. For example, Orson could choose to display the radiators that consume between 0 and 2000 dollars.

As requirements, the aim is to design this software adapted to three different contexts (for convenience, the design of the three adaptations according to three platforms are presented but the benefit of the approach increse as well as the number of adaptations):

1) if the user has a **personal computer with a large screen**, the graphic and the filters are displayed in a same frame (see Figure 2). The graphics selection is made by icons (see the"$SelectGraphic\_Icon$" red square in Figure 2). The selected data are displayed as a list widget (see the "$DataDisplayed\_Radio$" square in the Figure 2). The two other filters (see "$TimeLine\_Slider$" and "$Filter1D\_Slider$" red squares in Figure 2) use a slider to select intervals.
2) if the user has a **smartphone with a small screen** (see Figure 3), the graphics and the filters are displayed in two tabs. Because they use less space, the icons and the list are replaced by a combo box (see "$SelectGraphic\_ComboBox$" and "$DataDisplayed\_ComboBox$" in Figure 3), the Sliders are replaced by two textfields (see "$TimeLine\_2TextField$" and "$Filter1D\_2TextField$" in Figure 3).
3) if the user has a **personal computer with a large screen** and a **smartphone with a small screen**, the graphic is displayed on the PC (such as in the Figure 2 without the three filters) and the three filters are displayed on the smartphone (such as the second screen in the Figure 3 without the tab menu).

Prototype implementation. The prototype UI ADAPTOR is implemented in JAVA and JAVAFX. The two platforms are simulated according to their screen size (1280x1024 for the large screen and 640x960 for the small screen).

Feature model. These requirements are modelled by the feature model of Figure 4 according to the methodology proposed by [22]. The UI variations depend on the task. For example, the "$filter1D$" feature expresses the choice for the designer to use a slider or two textfields to design the filter 1D component. Each constraint informs the designer how the UI feature can be selected. For example, the "$Filter1D\_Slider$" feature is selected if there is large screen and no other platform available. Consequently, this feature is selected when Orson has a personal computer (case 1 of the requirements). The context of use variations depend on the screen size and the number of platforms. For example, a screen size can be large or small. The "$LargeScreen2$" feature is in red because it can be selected. Indeed, when Orson has a second platform, it is a smartphone such as specified by the requirements.

Components. The components are implemented in JAVA FX. The component model is not the focus of this paper, the component model is defined as in [10]. Each component corresponds to a concrete features of the feature model (Figure 4) and is represented by a red square in the Figures 2 and 3. The UI of these components are underlined in a red square in the Figures 2 and 3. A last component is developed corresponding to the "$VisuData2\_Frame$" feature. This component encompasses two frames in charge of displaying the graphic on the large screen and the three filters on the small screen.

Context-aware adaptation life-cycle. UI ADAPTOR supports the context-aware adaptation life-cycle. Firstly, the context of use perception is simulated by the designer, i.e., the context features are selected manually. Secondly, from the selected context features, Feature IDE deduces (thanks to a SAT solver [24]) the selected UI features automatically. Thirdly, from the list of selected features, the UI ADAPTOR composes the appropriate component in order to produce a composed UI displayed according to the corresponding size of the screen.

VI. CONCLUSION AND LESSONS LEARNED

This paper has proposed a method to design adaptive User Interfaces according to SPL engineering. The experience allows to learn lessons and benefits concerning the proposed methods.

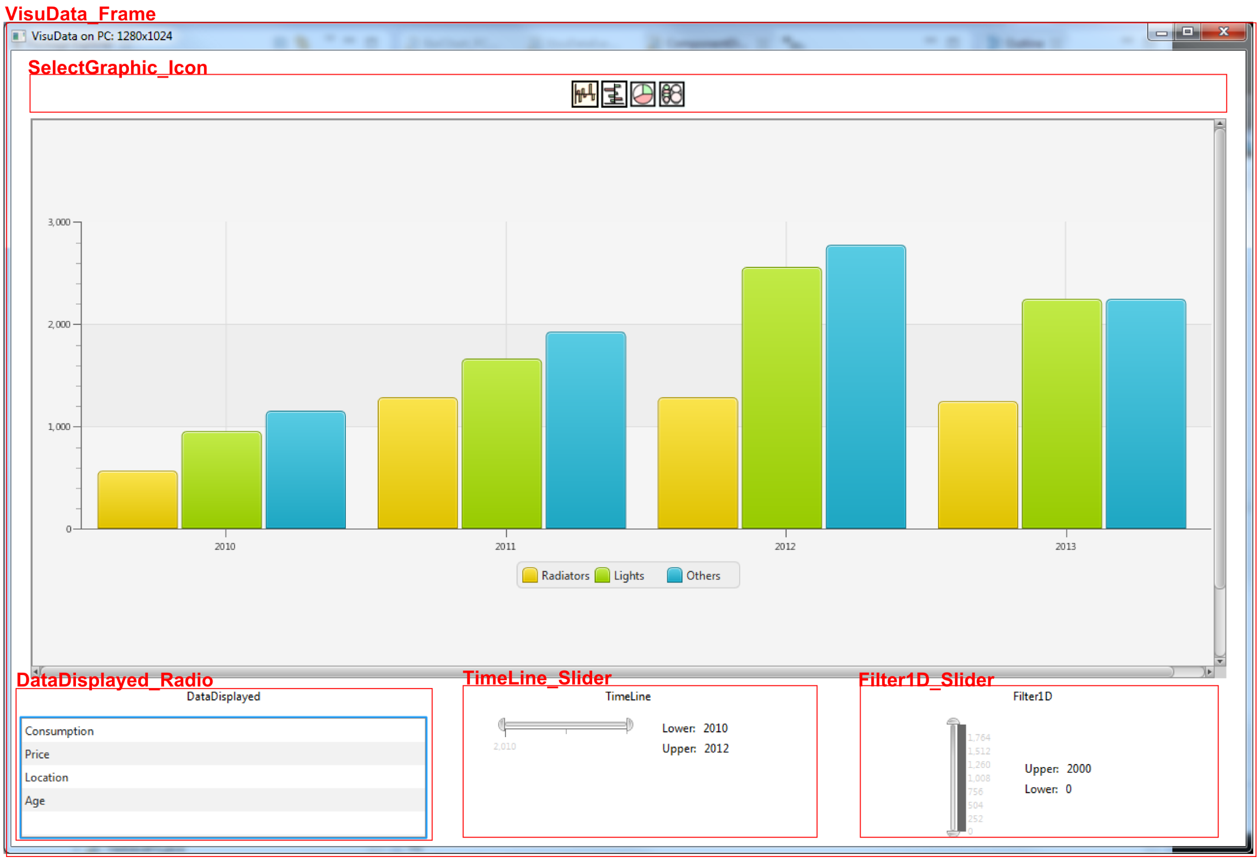


Figure 2. The composed User Interface adapted to a large screen. The red squares and the red labels are added to identify the UI components.

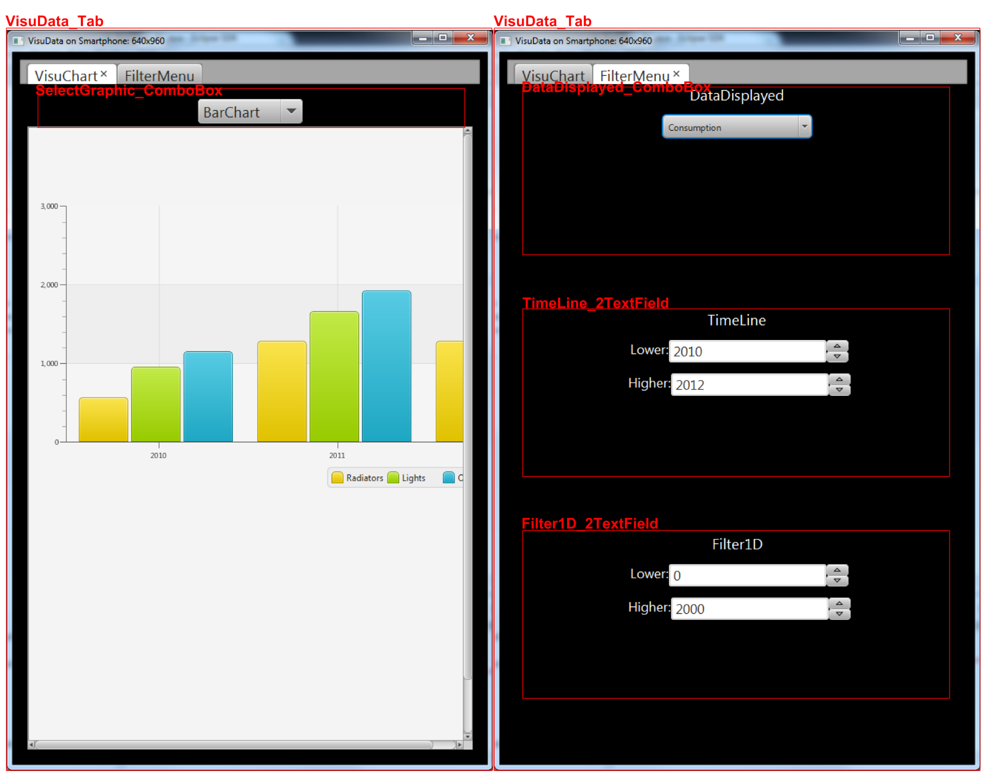


Figure 3. The composed User Interface adapted to a small screen. The red squares and the red labels are added to identify the UI components.

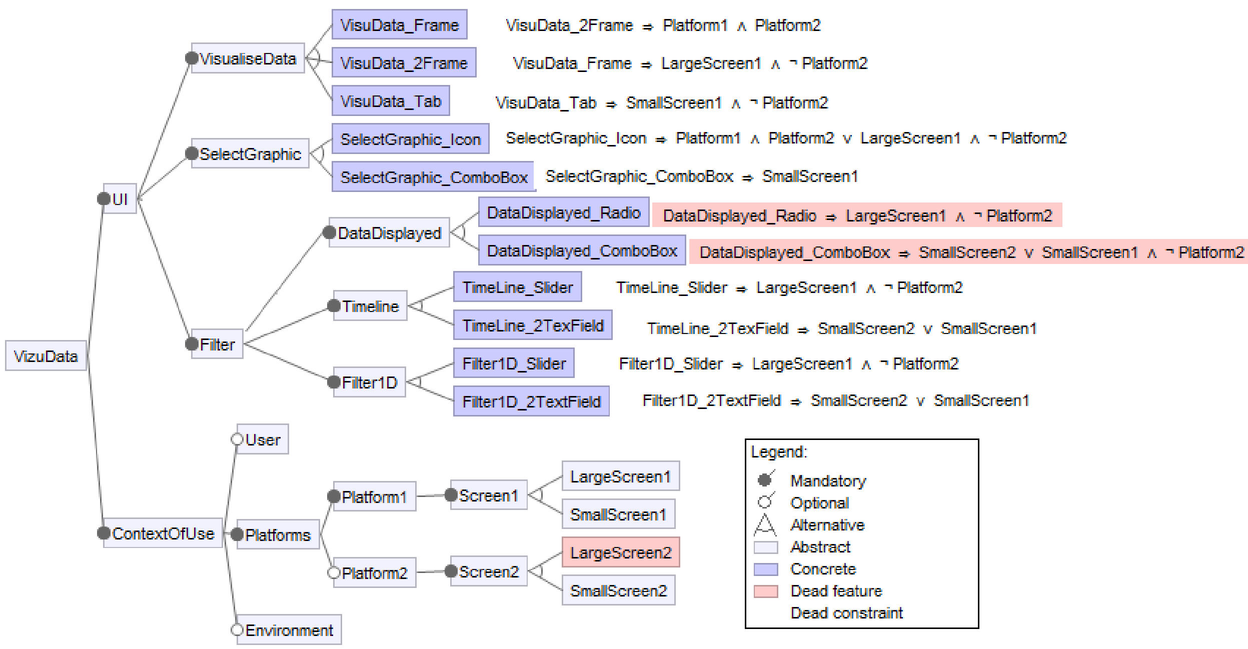


Figure 4. Feature Model of VisuData

Feature model. Because the feature model has a visual representation compared to adaptation rules, it increases transparency and understandability of the adaptation rules used by producing a formal and visual model of variability. Because the FM is formal, the FM also increase reusability between different adUIs and different designers. It is also particularly important to maintain and reuse adUI in order to design new adaptations to add new contexts. The constraints on the feature model increase the validity of the adaptation. For example, the dead features (LargeScreen2) and the inconsistent constraints are detected by Feature IDE. Moreover, the context of use can be easily modelled because there are other FM editors (such as CVL [25]) that handle the cardinality of the features. In consequence, these editors allow to model only one platform features without listing the number of platforms used.

Components. The use of components increase reusability and enhancement of quality because the UI components are reviewed and tested in many adaptive/composed UIs. Component-based programming can also be combined or replaced by other programming paradigms. For example, a component can be seen as a set of elements of the standard abstraction levels [12]. Moreover, when a UI component is maintained, the change can be propagated to all adUI with which the component is being used.

Process. The flexibility of the method allows to reuse pre-existing development approaches to design adUI. For example, the adaptation rules can be defined as aspects or models. The features can model the variability of a component, a task, an interactor but also the adaptation rules. In addition, the proposed method allows to improve the cost estimation of the adUI to design based on previous experience [4].

We plan to increase the number and type of artefacts by adding tests to design more complex adUI. We plan to design a complete framework to design AdUi, i.e., including context detection and multi programming languages (HTML and JAVA) in order to evaluate the designer effort and adaptation performance.

REFERENCES

[1] J.-S. Sottet et al., “Model-driven adaptation for plastic user interfaces,” in Human-Computer Interaction–INTERACT 2007. Springer, 2007, pp. 397–410.

[2] V. G. Motti and J. Vanderdonckt, “A computational framework for context-aware adaptation of user interfaces,” in RCIS, R. Wieringa, S. Nurcan, C. Rolland, and J.-L. Cavarero, Eds. IEEE, 2013, pp. 1–12.

[3] G. Meixner, F. Patern, and J. Vanderdonckt, “Past, present, and future of model-based user interface development.” i-com, vol. 10, no. 3, 2011, pp. 2–11.

[4] K. Pohl, G. Bockle, and F. Van Der Linden, Software product line engineering. Springer, 2005, vol. 10.

[5] K. Kang, S. Cohen, J. Hess, W. Novak, and A. Peterson, “Feature-oriented domain analysis (foda) feasibility study,” DTIC Document, Tech. Rep., 1990.

[6] P.-Y. Schobbens, P. Heymans, J.-C. Trigaux, and Y. Bontemps, “Generic semantics of feature diagrams,” Computer Networks, vol. 51, no. 2, 2007, pp. 456–479.

[7] I. Jaouadi, R. Ben Djemaa, and H. Ben Abdallah, “Interactive systems adaptation approaches: A survey,” in ACHI 2014, The Seventh International Conference on Advances in Computer-Human Interactions, 2014, pp. 127–131.

[8] F. Paterno, C. Santoro, and L. D. Spano, “MARIA: a universal, declarative, multiple abstraction-level language for service-oriented applications in ubiquitous environments,” vol. 16, no. 4, pp. 1–30.

[9] A. Blouin, B. Morin, O. Beaudoux, G. Nain, P. Albers, and J.-M. Jzquel, “Combining aspect-oriented modeling with property-based reasoning to improve user interface adaptation,” in Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems. ACM, 2011, pp. 85–94.

[10] Y. Gabillon, M. Petit, G. Calvary, and H. Fiorino, “Automated planning for user interface composition,” in Proceedings of the 2nd International Workshop on Semantic Models for Adaptive Interactive Systems: SEMAIS’11 at IUI 2011 conference. Springer HCI, 2011, pp. 1–5.

[11] V. G. Motti, D. Raggett, and J. Vanderdonckt, “Current practices on model-based context-aware adaptation,” in CASFE, 2013, pp. 17–23.

[12] G. Calvary, J. Coutaz, D. Thevenin, Q. Limbourg, L. Bouillon, and J. Vanderdonckt, "A unifying reference framework for multi-target user interfaces," Interacting with Computers, vol. 15, no. 3, 2003, pp. 289–308.

[13] S. Hallsteinsen, M. Hinchey, S. Park, and K. Schmid, "Dynamic software product lines," Computer, vol. 41, no. 4, 2008, pp. 93–95.

[14] C. Parra, X. Blanc, and L. Duchien, "Context awareness for dynamic service-oriented product lines," in Proceedings of the 13th International Software Product Line Conference. Carnegie Mellon University, 2009, pp. 131–140.

[15] P. Trinidad, A. Ruiz-Cortés, J. Pena, and D. Benavides, "Mapping feature models onto component models to build dynamic software product lines," in International Workshop on Dynamic Software Product Line, DSPL, 2007, pp. 51–56.

[16] K. Geihs et al., "A comprehensive solution for application-level adaptation," Software: Practice and Experience, vol. 39, no. 4, 2009, pp. 385–422.

[17] H. Arboleda, A. Romero, R. Casallas, and J. Royer, "Product derivation in a model-driven software product line using decision models," in Proceedings of the Memorias de la XII Conferencia Iberoamericana de Software Engineering, CIbSE, vol. 2009, 2009, p. 59.

[18] K. Garcés, C. Parra, H. Arboleda, A. Yie, and R. Casallas, "Variability management in a model-driven software product line," Revista Avances en Sistemas e Informática, vol. 4, no. 2, 2007, pp. 3–12.

[19] A. Pleuss, G. Botterweck, and D. Dhungana, "Integrating automated product derivation and individual user interface design," Proceedings of VaMoS, vol. 10, 2010, pp. 69–76.

[20] A. Pleuss, B. Hauptmann, D. Dhungana, and G. Botterweck, "User interface engineering for software product lines: the dilemma between automation and usability," in Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems. ACM, 2012, pp. 25–34.

[21] Q. Boucher, G. Perrouin, and P. Heymans, "Deriving configuration interfaces from feature models: A vision paper," in Proceedings of the Sixth International Workshop on Variability Modeling of Software-Intensive Systems. ACM, 2012, pp. 37–44.

[22] Y. Gabillon, N. Biri, and B. Otjacques, "Designing multi-context uis by software product line approach," in Proceedings of the International Conference on Human-Computer Interaction (ICHCI'13). World Academy of Science, Engineering and Technology (WASET), 2013, pp. 628–637.

[23] A. S. Karataş, A. H. Doğru, H. Oğuztüzün, and M. Tolun, "Using context information for staged configuration of feature models," Journal of Integrated Design and Process Science, vol. 15, no. 2, 2011, pp. 37–51.

[24] C. Kastner et al., "Feature ide: A tool framework for feature-oriented software development," in Software Engineering, 2009. ICSE 2009. IEEE 31st International Conference on. IEEE, 2009, pp. 611–614.

[25] K. Czarnecki, P. Grunbacher, R. Rabiser, K. Schmid, and A. Wkasowski, "Cool features and tough decisions: a comparison of variability modeling approaches," in Proceedings of the sixth international workshop on variability modeling of software-intensive systems. ACM, 2012, pp. 173–182.

 ISBN: 978-1-61208-382-7

Intelligent Shop Window

Producing Dynamic Situated Augmented Reality using a Large See-through Screen

Reo Suzuki
Graduate School of Science and Technology
Seikei University
Tokyo, Japan
dm146207@cc.seikei.ac.jp

Yutaka Takase, Yukiko I. Nakano
Faculty of Science and Technology
Seikei University
Tokyo, Japan
{yutaka-takase, y.nakano}@st.seikei.ac.jp

***Abstract*— The purpose of shop windows is to attract people's attention and increase the consumers' desire to buy the merchandise. This paper proposes an "intelligent shop window" that can display merchandise information and advertisements on a large see-through display overlaid on real items or people in the shop. The system decides the target item to be advertised based on the position of the customers/shop clerks and the color of the clothes that a clerk is wearing. Then, it displays the information about the item on a window-like see-through screen through which passers-by can see the information overlaid on the situation inside the shop. As compared to the mannequins in a shop window, our system can present information dynamically. Therefore, it is expected to attract greater interest from passers-by and increase the number of customers visiting the shop.**

Keywords - augmented reality; shop window, overlaid display, Kinect sensor.

I. INTRODUCTION

Pedestrians on a street look at the shop windows of the small shops and department stores. Shop windows attract people's attention and have the effect of increasing the consumers' desire to buy. Through the shop window, people would like to see whether the shop sells items matched with their preference, and who the customers are (age, status, sense of fashion). Then, if the shop looks good, they step into the shop. Therefore, the information inside the shop is very important for the customers to decide whether they go into the shop or not.

By virtue of the development of larger displays, production technologies, and human sensing technologies, many studies on interactive displays have been conducted. These display systems can recognize the environment of public places and change the displayed information accordingly in real time. For example, such systems were used in exhibitions at museums [1][2]. Other systems can project information onto the facade of buildings [3][4].

When designing information display systems for shop window interfaces, the essential roles of a shop window should be considered. Doorn et al. [5] suggested that the shop window has three roles as a marketing tool.

1. It expresses the style of the shop;
2. It creates the appropriate atmosphere and attracts people;
3. It informs people about the available products and their function.

The objective of many of the previous systems was to attract pedestrians and allow them to interact with the system. The Dynamically Transparent Windows system [6] utilizes foil that becomes transparent when energized. By making a specific small area transparent according to the movement of the pedestrians, the system changes the displayed products and areas in the shop that are visible to the pedestrians. With the goal of providing an ambient response to the users' input, in the Persuasive Interactive Mannequin system [7], mannequins displayed in the shop window gaze at the customer looking into the window. Moreover, Müller et al. [8] reported that the users' degree of interest differed according to the method of giving feedback via an interactive display that can interact with passers-by in public places.

However, many previous systems were more interested in recognizing the presence and motions of passers-by, but not in recognizing the situation inside the shop. Our study, on the other hand, focuses more on the environment inside the shop because the information inside the shop is useful for the users to decide whether to step in the shop or not. Thus, our "intelligent shop window" attracts the attention of passers-by by changing the display information according to the environment inside the shop. The passers-by are able to observe the situation inside the shop through the window. In addition, they can see overlaid information related to the people and items inside the shop. When no people are close to the window, the shop window displays promotion materials that represent the style of the merchandise inside the shop. The system can also show the details of the items when people are close to the window.

The rest of this paper is organized as follows: Section 2 describes in detail our proposed system. Section 3 describes the implementation of the system. Following this, Section 4 presents an example of application. And finally, Section 5 summarizes this paper and mentions future work.

II. PROPOSED SYSTEM

We propose a display system that resembles a conventional large shop window adjacent to the sidewalk through which users can see enhanced information about inside a clothes shop. Basing its decision on the clothes that the customers and shop clerks are wearing and their position in the shop, the system selects an item that is popular or recommended and overlays the advertisement information

 ISBN: 978-1-61208-382-7

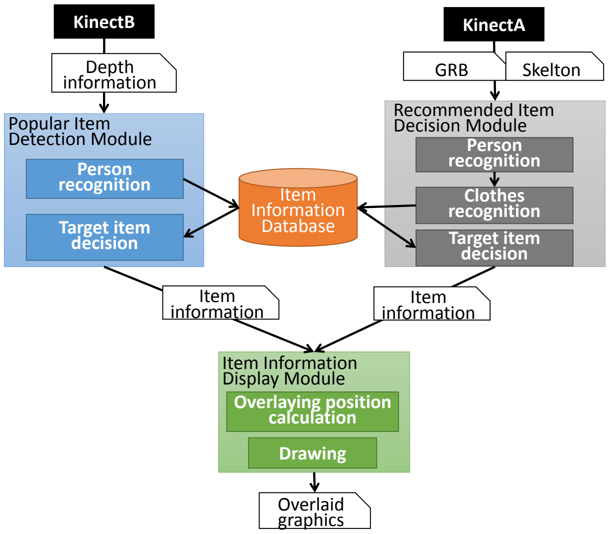


Figure 1. System architecture

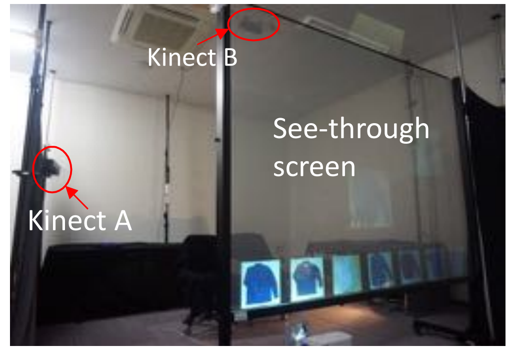


Figure 2. Environment

related to the item on the shop window. The advantage of this system is that the advertisement information is automatically generated by recognizing the shop clerks, and it is not needed to be manually changed/registered every day. Moreover, this system enables the passers-by to observe that many customers are gathered around a popular item and see the information about it, which is overlaid on the window. They can also see recommended items by looking at information overlaid near a shop clerk who is wearing the item. This function allows the passers-by to get better idea about inside the shop as well as the information about popular items.

Our proposed system consists of a popular item detection module, recommended item decision module, item information database, and item information display module. The system architecture is shown in Figure 1. A detailed explanation of each module is presented below.

A. Popular Item Detection Module

We define a "popular item" as an item around which many people gather or which they stop to pick up and examine. In order to detect a popular item according to this definition, the people and the item close to which people are standing must be recognized. It is also necessary to determine whether a person is only passing by or examining the item. If the person is examining the item, the system retrieves the item data by searching the product database and sends the information to the display module. However, when many people are present in the shop, the information overlaid on all the customers can be a source of irritation. In such a situation, our system identifies crowded locations based on the recognition of the position and movement of people and prioritizes the most crowded location when displaying the overlaid information. For people recognition, depth information sensed by a Kinect sensor is used.

B. Recommended Item Decision Module

As an advertisement of the shop, the system displays the information about special items that characterize the sense of fashion of the shop. For this purpose, our shop window system recognizes special items that shop clerks are wearing (we assume that the shop clerks wear the special items as advertisement), and displays the information about the special item when a clerk passes near the shop window. The item information is displayed according to the position and movement of the shop clerk. To achieve this, the module recognizes a person inside the shop passing by the window and recognizes the item that he/she is wearing using RGB color information and a simple pattern matching technique. Then, using these data as keys, the module searches the item information database, and sends the results to the information display module. Note that this function enhances the information inside the shop, and make the special items attract the attention from the passers-by.

C. Item Information Database

This database is used for detecting popular items and recommended items. For both these purposes, the item name and its RGB image must be registered. In addition, for detecting popular items, the layout of the shop, that is, the display location of each item, is required. The system measures the depth image around the locations where the popular items are displayed in order to recognize a crowd. For detecting recommended items, the RGB values are required so that the clothes that the shop clerks are wearing can be recognized.

D. Item Information Display Module

When the target items have been determined using the popular item detection module and the recommended item decision module, and information about the items has been obtained from the database, the item information is sent to the item information display module, which displays the information on the window.

The information should be appropriately overlaid so that the advantage of window interfaces that passers-by can see inside the shop is not impaired. The system satisfies this requirement using the following two methods.

Figure 3. Depth image for the person recognition

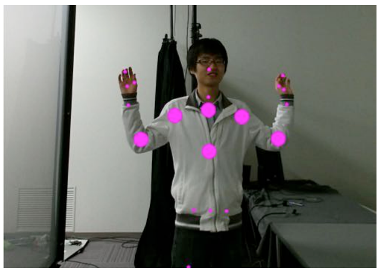


Figure 4. Positions of six joints for Clothes Recognition

- It determines the overlaid display by considering the number and position of target items;
- It prioritizes the items around which many people gather and those which a shop clerk is wearing.

III. IMPLEMENTATION

A. Environment and Equipment

We described the implementation of the proposed system in the previous section. The environment and the equipment are shown in Figure 2. We set up a 4 m × 4.25 m area as a shop. In this area, a 100-inch see-through screen (2.0 m × 1.5 m) was located on one side of the area as a shop window. Since real glass is not suitable for projection purposes, we used a see-through screen glued onto a large thick acrylic board. For person and clothes recognition, we used a Kinect for Windows v2 open Beta (hereinafter, referred to as the Kinect sensor) and a Kinect SDK 2.0. For person recognition, we set up a Kinect sensor on the ceiling (2.9 m in height) and obtained the depth image as viewed from the top (Kinect B in Figure 2). For clothes recognition, a second Kinect sensor (Kinect A) was set behind the screen.

B. Person Recognition

In order to implement the popular item detection module, the people gathered around an item must be recognized. We used a Kinect sensor attached to the ceiling to obtain the depth image as viewed from the top. The depth image resolution is 512 × 424 pixels, and the pixel data are processed by OpenCV library. Figure 3 shows an example of the depth image. There is a person at the center of the picture. We can recognize him/her with the depth information. For detecting crowded locations, a sensing area is predefined for each item, and the system monitors the depth data in this area. Before the system is initiated, that is, when no person is present in the area, the depth image of each sensing area is measured in order to calibrate the sensor. When the system has been initiated, if a certain amount of change from the initial depth information is detected, we count a sequence of the pixels, the depth data of which changed from the original values. If the number of pixels in a given sequence is larger than a threshold, the system recognizes that a person has entered the area. If the number of pixels is smaller than the threshold, the result is discarded as noise. It sometimes occurs that a person only passes by the item and stays inside the area for a very short time. In this case, the number of frames during which the person stays inside the area is counted. If the number of frames is larger than a threshold, i.e., the person stays inside the area longer than the threshold, the system recognizes that this person is examining the item.

When the variation in the depth data is sufficiently large, the system assumes that many people are remaining in the area and the item is popular. As the popular items should be prioritized over other items in the overlaid display, in addition to the position of the target item, the system measures the size of the area where the depth image variation is large and sends the information to the item information display module.

C. Clothes Recognition

In the recommended item decision module, a function for recognizing the clothes that a shop clerk walking near the shop window is wearing is required. In the current implementation, we used RGB color data sampled from several points on the clothes and applied them in a simple pattern matching technique.

First, using the Kinect SDK, the system obtains the skeleton (position) and RGB color information for 25 joints. As shown in Figure 4, we used RGB information for six joints: SpineMid, ShoulderLeft, ElbowLeft, ShoulderRight, ElbowRight， and SpineShoulder. The sensing area of the Kinect sensor is restricted such that persons outside the shop window are not recognized as being inside the shop. When a person is detected and tracked, the RGB values of the six joints are measured. Then, using a Sum of Squared Difference (SSD) method, these values are compared with the RGB value for each item in the database. If the SSD value is smaller than a threshold for more than two data points among the six, the system determines that the input data match the candidate item. Then, to obtain the position of the person who is wearing the recognized clothes, the position of the SpineMid joint is obtained, and its distance from the Kinect sensor is calculated based on the position data. Finally, the information about the item and the distance between the Kinect sensor and the clerk are sent to the item information display module.

D. Overlaid Display

In the item information display module, the results of the person and clothes recognition are processed, and the item information is then displayed. To avoid obstructing the view

Figure 5. A person examining an item

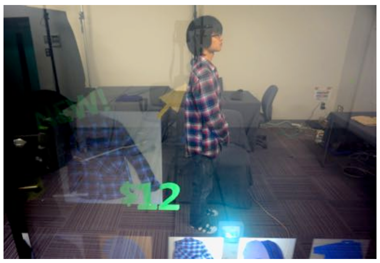


Figure 6. A shop clerk wearing a recommended

of the passers-by into the shop, only three items, which were selected as the advertisement targets in the previous 30 sec, are displayed.

For displaying the item information that is sent from the popular item detection module, the information is overlaid on the real target item in the shop. Thus, the display position is determined based on the position of the item. For displaying information about the recommended items, the information must be overlaid such that the passers-by can easily identify who is wearing the recommended item. We use the position of the clerk wearing the clothes for determining the display position.

IV. Example

This section shows how the implemented mockup system works using a simple example. Figures 5 and 6 show snapshots of our intelligent shop window. There are five items in the shop. At the bottom of the window, the graphics of all items are displayed repeatedly.

Figure 5 shows that the system detects a person who is examining an item, and overlays the advertisement about the item. In this example, the graphics of a shirt and an advertisement phrase "ON SALE" are overlaid close to the person.

Figure 6 shows an example of a recommendation. The system recognizes a shop clerk and his clothes, and then, it displays the advertisement of the item by overlaying it on the clerk. In this example, a yellow arrow is drawn to clearly identify the person wearing the recommended clothes.

V. Conclusion and Future Work

In this paper, an intelligent shop window that can display information and advertisements of items sold in the shop was proposed. The system decides the target item to be advertised based on the position of customers/clerks and the color of the clothes that the clerk is wearing. Then, the system displays the information about the item on a window-like see-through screen through which people passing by the window can see the information overlaid on the situation inside the shop. Since this system can produce a more dynamic display than mannequins in a shop window, it is expected to attract more interest from the passers-by and increase the number of customers visiting the shop.

In future work, the capability of the system to distinguish between shop clerks and customers should be improved since, in the current implementation, the system cannot achieve this. This capability is important, since in the popular item detection module, only customers should be counted as members of a crowd, and in the recommended item decision module, only a clerk should be tracked. In addition, the clothes recognition function does not allow the current system to recognize clothes with a complex design, since we applied a simple pattern matching method that uses RGB values. To improve the clothes recognition function, it is necessary to employ more sophisticated computer vision technologies. Finally, our ultimate goal is to use this system as a real shop window and test whether the system can attract more passer-by attention.

Acknowledgment

This study is partially supported by JSPS KAKENHI Grant Number 25280076.

References

[1] U. Hinrichs, H. Schmidt, and S. Carpendale, "EMDialog: bringing information visualization into the museum," IEEE Trans. Vis. Comput. Graph., vol. 14, no. 6, 2008, pp. 1181–8.

[2] E. Hornecker, "Interactions around a contextually embedded system," in Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '10, 2010, p. 169.

[3] P. Dalsgaard and K. Halskov, "Designing urban media façades," in Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10, 2010, p. 2277.

[4] P. Peltonen, et al, "It's Mine, Don't Touch!," Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08, 2008, p.1285.

[5] M. van Doorn, E. van Loenen, and A. P. de Vries, "Deconstructing ambient intelligence into ambient narratives: the intelligent shop window," Proceedings of the 1st international conference on Ambient media and systems, 2008, pp. 1–8.

[6] L. Mailund and K. Halskov, "Designing marketing experiences," in Proceedings of the 7th ACM Conference on Designing Interactive Systems - DIS '08, 2008, pp. 222–229.

[7] W. Reitberger, et al, "A persuasive interactive mannequin for shop windows," in Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09, 2009, p. 1.

[8] J. Müller, R. Walter, G. Bailly, M. Nischt, and F. Alt, "Looking glass: A field study on noticing interactivity of a shop window," in Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12, 2012, p. 297.

 ISBN: 978-1-61208-382-7

Human-Machine Cooperation in General Game Playing

Maciej Świechowski*, Kathryn Merrick†, Jacek Mańdziuk‡, Hussein Abbass†
*Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
Email: m.swiechowski@ibspan.waw.pl
†School of Engineering and Information Technology, University of New South Wales, Canberra, Australia
Email: (k.merrick, h.abbass)@adfa.edu.au
‡Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland
Email: j.mandziuk@mini.pw.edu.pl

***Abstract*—This paper presents a framework for cooperation between a human and a general game playing agent. Cooperation is defined as two entities causing each other to modify their behaviour to achieve some mutual advantage. Such human-computer cooperation has the potential to offer insights that can help us improve the performance of artificial agents, as well as improving the performance of humans during certain kinds of strategic interactions. This paper focuses specifically on game playing as a form of strategic interaction. By proposing a framework for cooperation between a human and a general game playing agent, our aim is to create a flexible system that may be applicable to cooperation in other kinds of problem solving and strategic interactions in the future. We evaluate the framework presented in this paper by means of a human study. We observe humans playing games with and without the cooperation of a general game playing agent. We present experimental results of the pilot study as well as proposed changes in the experiment. These changes aim to verify the hypothesis that human-machine cooperation within our framework can indeed lead to mutual advantage.**

Keywords–human-machine study; cooperation; General Game Playing; Monte Carlo Tree Search.

I. Introduction

General Game Playing (GGP) has been claimed as "The AI Grand Challenge", since it is seen as a step towards strong human-like intelligence [1]. The design and study of approaches that permit cooperation between humans and GGP agents is thus an important, complementary research stream. Such human-computer cooperation has the potential to offer insights that can help us improve the performance of artificial agents, as well as improving the performance of humans during certain kinds of strategic interactions. We borrow a concept of cooperation from [2] stating that it takes place when two systems cause each other to modify their behavior to achieve some mutual advantage. The type of strategic interaction we will consider is game playing. We will consider the type of machine cooperator as a GGP [3] agent as proposed by the Stanford Logic Group [4]. This is currently the most prominent embodiment of the multi-game playing idea, which aims to create systems capable of playing a variety of games (as opposed to agents that can only play single games). The specific type of GGP machine cooperator we will consider is a Monte Carlo Tree-Search (MCTS) based player. The MCTS is used as the main routine of the strongest state-of-the-art GGP players and is also widely applied to other games such as Go [5], Arimaa [6] as well as other areas of Artificial Intelligence (AI) [7]. We will conduct a human user study to validate our approach to human-machine cooperation. In this paper, we present two pilot studies we have performed. The aims of these pilot studies were to (1) verify our setup for cooperation and (2) provide preliminary verification of our research hypothesis. A large-scale experiment is the next step to undergo. Apart from providing the circumstances for the cooperation, we are also interested in measuring the effects of such cooperation, i.e., how it affects the average quality of play. Human-machine interaction has been a hot research area outside the scope of games, e.g., in the areas of aviation [8] or surgery [9]. In games, however, the task of creating machine players has been challenging enough on its own [5][10]. To our knowledge, there has been no related work concerning human-machine cooperation in GGP or in any other MCTS-based game playing. We believe that the way we approach the problem of cooperation can contribute to the area of general knowledge-free and learning-based methods in games [11], because we can examine the way humans learn from machines and provide a basis for automatic methods by which machines can learn games from humans.

The remainder of the paper is organized as follows: the next two sections contain brief descriptions of GGP, MCTS and our cooperation platform within the MCTS framework. In Sections IV and V, we formulate the research hypothesis and the experimental methodology, respectively. Section VI describes the two particular setups tested in the two pilot studies and Section VII discusses the results. The last section is devoted to conclusions and directions for future work.

II. General Game Playing

A. Basics

GGP is a trend in AI which involves creating computer systems, known as GGP agents, capable of playing a variety of games with a high level of competence. The range of games playable within the GGP framework is any finite deterministic game. Unlike specialized playing programs, GGP systems do not know rules of the games being played until they actually start. The concept of designing universal game playing agents is also known as multi-game playing or metagaming, but as stated in the introduction, we refer to the Stanford's definition of GGP [3] which is the most recent one. The official GGP Competition, which is *de facto* the World Championship Tournament, is also part of the GGP specification. The machine player used in this research is our entry in the latest installment of the competition (2014). Borrowing from the GGP terminology, we will use the term *play clock* for the time (in seconds) available to make a move by a player. To enable matches between our GGP program and humans, we

 ISBN: 978-1-61208-382-7

had to slightly loosen the official specification. For instance, GGP agents are normally penalized for not responding with a legal move in time by having the move chosen for them at random. In our scenario, human participants can think about moves as long as they want to without any penalty and the machine players always respond in time.

B. The Tree-Search Algorithms Used

MCTS is an algorithm for searching a game tree in a quasi-random fashion in order to obtain as accurate an assessment of game states as possible. In general, the assessment is computed statistically as the average score - **Q** - which is defined by the total score of simulations going through a state divided by the number of visits to that state. The total score is a sum of the outcomes of simulations. For all games considered in this article, the value of 1.0 denotes a win, 0.5 denotes a draw and 0.0 denotes a loss in a single simulation. The input to the method is the current game state. Then, the algorithm gradually searches the game tree starting from the current state in a series of iterations adding one node in each of them. An iteration consists of the following four steps:

1) **Selection.** Start from the root and go progressively down. In each node, choose the child node with the highest average score until reaching a leaf node.
2) **Expansion.** If a state contained in the leaf node is not terminal, choose an action which would fall out of the tree. Allocate a new child node associated with that action; simulation.
3) **Simulation.** Starting from a state associated with the newly expanded node, perform a full game simulation (i.e., to a terminal state).
4) **Backpropagation.** Fetch the result of the simulated game. Update statistics (average scores, numbers of visits) of all nodes on the path of simulation, starting from the newly expanded node up to the root node.

The algorithm can be stopped at any time. The final output of the search is the action with the highest average score **Q** for the player who is currently to make a move in a game. A significant improvement over the pure MCTS is the Upper Confidence Bounds Applied to Trees (UCT) algorithm [12]. The purpose of the algorithm is to maintain balance between the exploration and exploitation ratio in the selection step. Instead of sampling each action uniformly (as is the case of MCTS) or greedily, the following selection formula is applied:

$$a^* = arg \max_{a \in A(s)} \left\{ Q(s,a) + C\sqrt{\frac{ln\,[N(s)]}{N(s,a)}} \right\} \quad (1)$$

where ***s*** is the current state; ***a*** is an action in this state; ***A(s)*** is a set of actions available in state ***s***; ***Q(s,a)*** is an assessment of performing action ***a*** in state ***s***; ***N(s)*** is a number of previous visits to state ***s***; ***N(s,a)*** is a number of times an action ***a*** has been sampled in state ***s***; *C* is the exploration ratio constant.

III. Cooperation in the MCTS Framework

The machine cooperator used in this paper is an adapted MiNI-Player [13][14] - a GGP program equipped with additional features to enable cooperation. First and foremost, the machine provides statistics to help humans choose which move to play. During cooperative play, it is always a human who makes the final choice with or without taking advantage of the provided statistics. The second means of cooperation is by permitting interference with the MCTS. In this way, we propose an interactive process of building the game tree, while playing the game, involving both the machine and human. In the original MCTS, the same four-phase algorithm is repeated all the time during the *play clock*. For cooperative purposes we split this time into three equal intervals T1 + T2 + T3 = *play clock*. Between any two consecutive intervals (T1 and T2 or T2 and T3) humans can interact with the MCTS based on statistics presented to them. The statistics include: each action ***a*** available to the player to make a move with the ***Q(s,a)*** and ***N(s,a)*** values from (1). These values are scaled to the [0%, 100%] interval to be more readable by the participants. The final statistic is the actual number of simulations which ended with a win, draw and loss for the subject, respectively. The MCTS can be directed by the human in two ways: enabling/disabling actions available in the current state or toggling priorities of the actions on/off. If an action is disabled, the MCTS will ignore this action in the selection step, which means that no simulations will start with a disabled action. Changing the priority is equivalent to changing the value of the *C* parameter in (1) from 1 to 10. Participants are allowed to make any number of the aforementioned interventions at each step and once they are done, they click the simulate button to submit all of them in one batch and observe how the statistics have changed. By doing so, they can help the machine to focus on the most promising actions and avoid presumably wasteful computations. On the other hand, the feedback from the machine supports or questions the above-mentioned human player's choices. Our experimental design is justified based on two observations. First of all, in many well-established games, it has been found that the experts can intuitively discard unpromising actions and focus on the few best ones. Such behavior is manifested by human playing experience and intuition and is one of the aspects in which humans are better than machines. Provided that the human choice is correct, the process can converge faster to the optimal play. The introduction of action priority is a similar, but slightly weaker, modification to the MCTS algorithm. The second observation (or assumption) we made is that the cooperation has to be easy for participants to understand.

IV. Research Hypothesis

To focus the study of performance of human-machine cooperation we formulated the following research hypothesis: **a human cooperating with a machine GGP agent is a better player than human or machine agent individually**. We write this thesis in a shortened form of H + M > M and H + M > H, where H denotes a human player; M denotes a machine player and M + H denotes a hybrid player comprising a cooperating machine and human. We attempt to verify this hypothesis in a devoted experiment. The main research question is whether a mutually beneficial cooperation can originate and develop between human and machine players. In order to verify the above-listed hypotheses, we gathered samples from people playing without any machine assistance (H vs. M) and with such assistance (H+M vs. M). The first case involves a human simply playing a match against our GGP agent named MINI-Player [13] [14]. The second case involves a human playing against the same opponent but this time with assistance of a “friendly” GGP agent running in the background.

V. Pilot Studies

This paper reports on the results of two pilot studies that we have run to refine our experimental setup as well as to gather preliminary evidence regarding the research hypothesis. In this section, we present a technical setup and introduce one of the games used in the experiment. Because a well-played game is time consuming, we limited the number of games a single person can play to three. The experiment was performed separately for each human subject, so no information could be exchanged in the process, e.g., looking how other people play. The program participants used to play, and the opponent program were run on the same computer, both having access to two physical CPU cores. We set the *play clock* for the two machines (the cooperator and adversary) to 30 seconds in the first pilot study and 9 seconds in the second one. In order to avoid time-outs resulting from the human player, we discarded the concept of random moves if a player fails to respond in time. The matches were played only during weekdays anytime from the morning to the late afternoon. The age of participants varied from 21 to 30 with only one exception of 31 to 40. Most of them were PhD students of computer science. In the experiment, we used three games but one of them, named Tic-Tac-Chess, was discarded after the Pilot Study 1. Figures 1, 2 and 3 show screenshots of the program operated by participants for Inverted Pentago, Nine Board Tic-Tac-Toe and Tic-Tac-Chess respectively.

Inverted Pentago is a game played on a 6x6 board divided into four 3x3 sub-boards (or quadrants). Taking turns, the two players place a marble of their color (either red or blue) onto an unoccupied space on the board, and then rotate any one of the sub-boards by 90 degrees either clockwise or anti-clockwise. A player wins by making their opponent get five of their marbles in a vertical, horizontal or diagonal row (either before or after the sub-board rotation in their move). If all 36 spaces on the board are occupied without a row of five being formed then the game is a draw. Participants play as blue and are the second player to have a turn.

Nine Board Tic-Tac-Toe. In nine board tic-tac-toe, nine 3x3 tic-tac-toe boards are arranged in a 3x3 grid. Participants play as ’O’ and are the second player to have a turn. The first player may place a piece on any board; all moves afterwards are placed in the empty spaces on the board corresponding to the square of the previous move. For example if a piece was placed were in the upper-right square of a board, the next move would take place on the upper-right board. If a player cannot place a piece because the indicated board is full, the next piece may be placed on any board. Victory is attained by getting 3 in a row on any board.

Tic-Tac-Chess is a game played on a 7x7 board. Players start with one piece marked by a red or blue square in their respective starting location. Participants are the second player to have a turn. The starting locations are outside the movable area of the board which is defined by the inner 5x5 square. On their turn, each player may move a piece as though it were a Chess knight or capture with a piece as though it were a Chess king. Capturing is possible only with pieces belonging to the center 5x5 square. Pieces from the starting locations do not disappear when moved, so moving a piece from the starting location effectively spawns a new one on a destination square. The first player to get three pieces in a row, column, or diagonal in the center 3x3 square wins.

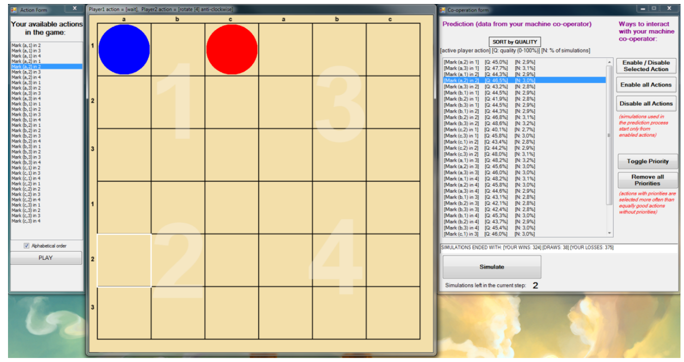


Figure 1. Screenshot of a program used to play Inverted Pentago (version with the cooperation).

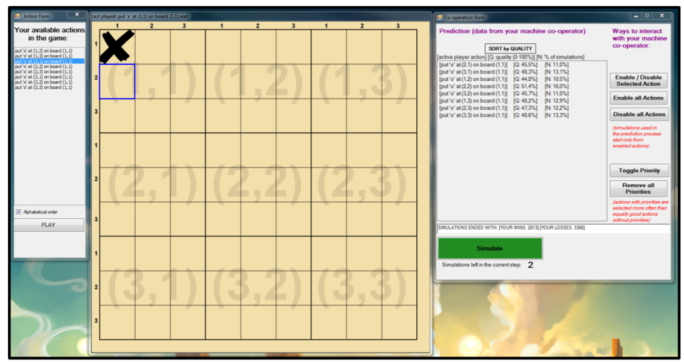


Figure 2. Screenshot of a program used to play Nine Board Tic-Tac-Toe (version with the cooperation).

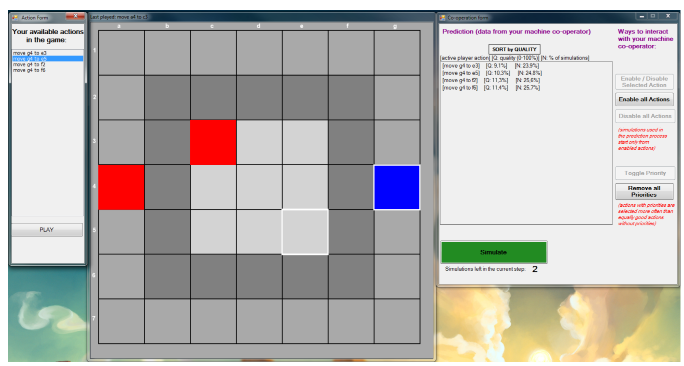


Figure 3. Screenshot of a program used to play Tic-Tac-Chess (version with the cooperation).

A. Pilot Study 1

We gathered 6 human participants for the first pilot study. They were divided into two groups of 3 people each. These two groups formed our two samples of data: playing with machine assistance (H+M) and without (H). During the experiment, we started each game with a short training session. We also gave participants a transcript explaining what they are asked to do and how the user-interface works. When participants were ready, they started playing a serious (i.e., not training) game and when they finished all three matches they were asked to complete a short questionnaire to obtain a profile of the subjects. The assignment of human players to games was based on the Latin Square Design with 3 games, 6 participants and two playing modes, i.e., with machine assistance being switched ON or OFF. Using this design, the minimum required number of participants for a full experiment is 12, but in the pilot study we stopped at 6 participants.

B. Pilot Study 2

At this point, we decided to revisit the experimental setup slightly and continue the experiment, called pilot study 2, to mitigate some problems that arose. Instead of asking people to play each game once, we asked them to play one game three times in order to enable learning by experience. The first match played includes a training session. The training session was extended to be a full match to let participants learn from their mistakes in endgames (late phases), which are often the most tricky to play. It is also often the case that people learn how to play better from the way they lost. We also excluded Tic-Tac-Chess from the set of games for giving too much advantage to the first player to have a turn. As a consequence, each subject lost their match very quickly in the same way leaving us with no relevant data to work on. Although there exist certain strategies to avoid a quick loss, it is unlikely to be seen by players unfamiliar with the game. Having only one type of game per participant, we modified the players' assignment in such way that we have all combinations of participants playing at least one of the three consecutive matches with the co-operation of the machine. In order to deal with the problem of long experiments, which was mainly caused by the simulation time needed to get meaningful results, we decided to write highly-optimized dedicated interpreters for rules of the chosen games. We were able to reduce the *play clock* just to 9 seconds.

VI. Results

We make the following observations based on numerical outcomes and human players' behavior during the experiments:

- The score between samples is even.
- All games appear to be very demanding for participants.
- There were no wins for Inverted Pentago and for the discarded game of Tic-Tac-Chess. There were 2 wins for Nine Board Tic-Tac-Toe, one with the cooperation and one without.
- The main reason for poor performance as specified by subjects in the questionnaire (and said after the experiment) was the lack of experience playing the given games. The rotations in Pentago were commonly mentioned as something being particularly difficult.
- Despite understanding the role of the program and the advice provided to them, the participants often seemed not to have desire to cooperate. If they had an assumption about which action was the best, they just opted to play it instead of investing time for more simulations.
- The participants seemed to enjoy playing the game but some stress was caused by the level of difficulty and the expectation to win.

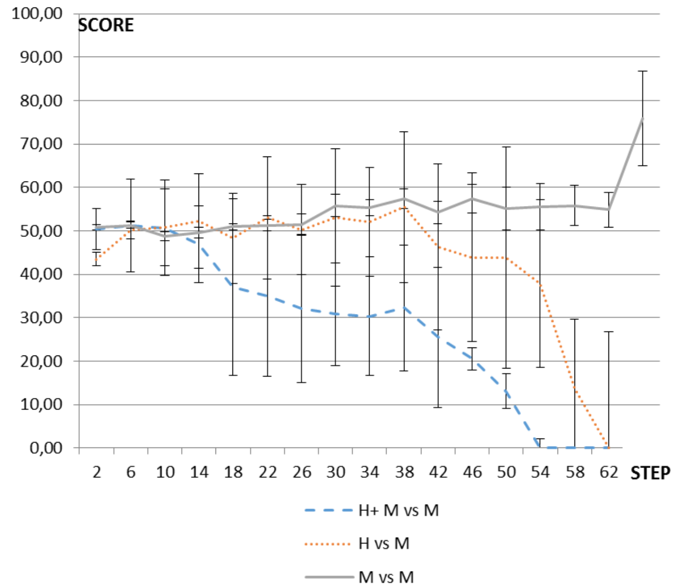


Figure 4. Graph showing the average scores obtained by the cooperating participants (H+M) and not cooperating participants (H) against the machine in Inverted Pentago.

Figure 4 shows the average scores (0 meaning loss and 100 meaning victory) obtained by the cooperating participants (H+M) and non-cooperating participants (H) against the machine in Inverted Pentago whereas Figure 5 shows the same graph for Nine Board Tic-Tac-Toe. Vertical error bars denote 95% confidence intervals. The X axis denotes game step (ply). The error bars overlap so the results cannot be used yet to formally verify the hypothesis. There were not enough participants in the pilot study to make any statistically significant claims. However, the trend so far is that the participants who did not cooperate played slightly better average games. This is reflected in the **H vs M** curve, starting from step 10, being above the **H + M vs M** one. However, both curves eventually meet at a common point which means that the average game results of both samples are even and equal to zero (which means a loss). The same properties are valid in the Nine Board Tic-Tac-Toe game. Because in the pilot studies, the participants rarely and quite chaotically used the cooperation possibilites, a conclusion that cooperation does not help would be an overstatement. The sample is too small, the participants would use the provided statistics when already behind in the game and because the cooperation options were shown only every second move, the machine was not able to help with a coherent line of actions.

Based on things we have learned during the pilot studies, these are the changes we want to make before moving to the final phase of the experiment:

- Each subject should play more than three times, preferably at least five. We have to make room for more learning possibilites, because it turns out that three games are not enough to learn how to play previously uknown games well (e.g., Inverted Pentago and Nine Board Tic-Tac-Toe). With more repeats we can also slightly reduce (though not eliminate) the effect of personal predispositions.

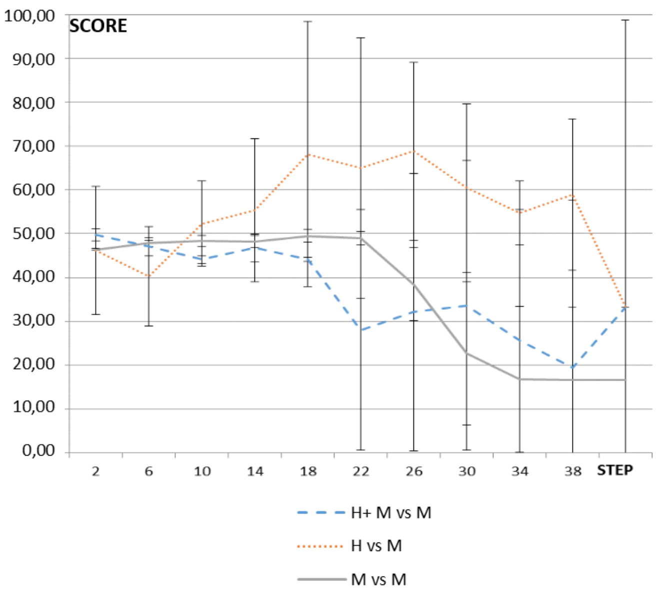


Figure 5. Graph showing the average scores obtained by the cooperating participants (H+M) and not cooperating participants (H) against the machine in Nine Board Tic-Tac-Toe.

- The cooperation options should definately be shown all the time for players playing with the help of a machine.
- We plan to remove actions' priorities and leave only enabling and disabling actions because the latter has more influence on the game tree and should be used more often. We have to make sure that all the participants understand why and when it is beneficial to disable actions.
- We will ask participants to play two games with the machine cooperation in the middle (e.g., the second and the third ones) to be able to observe, in the remaining games, the effects of learning from those games.

VII. Conclusions and Future Work

We analyzed the average outcomes of matches for the H + M vs. M and H vs. M samples of data as well as the average evaluation observed by the machine in every 4 steps of games. We computed 95% confidence intervals using the t-student test. It shows that the number of participants in the pilot study is not enough to make any significant claims regarding the hypothesis. Therefore, we plan to repeat the experiment for a larger sample of participants and with setup slightly modified.

We have presented a complex competitive environment in which human and machine can cooperate during strategic interactions. In general it appeared that subjects not having machine assistance fare slightly better, yet still worse than the machine opponent alone. The reason for this could, most likely, be attributed to the lack of continuous cooperation option (which was shown only at every other move). The other reasons include games' difficulty compounded by the lack of experience and possibly stressful activity of playing a game which is recorded. We believe that the way of introducing the cooperation into MCTS is a good idea, but the design of the experiment should be revisited.

An additional caveat is to maintain a proper balance of the experiment's difficulty. Games cannot be too easy for humans, because the machine cooperation would not be needed and, at the same time, cannot be too difficult to avoid a majority of games ending with a loss (which actually happened). We will restart the experiment with increasing chance to make the human participants learn the games. The participants also need to be clearly told that winning the match is not the exclusive goal of the experiment.

Acknowledgment

M. Świechowski would like to thank the Foundation for Polish Science under International Projects in Intelligent Computing (MPD) and The European Union within the Innovative Economy Operational Programme and European Regional Development Fund. The research was financed by the National Science Centre in Poland, grant number DEC-2012/07/B/ST6/01527. This work was performed while Maciej Świechowski and Jacek Mańdziuk were visiting UNSW Canberra. The ethics approval number granted from the university is A14-09

References

[1] J. Mańdziuk, "Towards Cognitively Plausible Game Playing Systems," IEEE Computational Intelligence Magazine, vol. 6, no. 2, 2011, pp. 38–51.

[2] C. P. Hoc J-M. and H. E., Eds., Expertise and Technology: Cognition & Human-computer Cooperation. Psychology Press, 2013.

[3] M. R. Genesereth, N. Love, and B. Pell, "General Game Playing: Overview of the AAAI Competition," AI Magazine, vol. 26, no. 2, 2005, pp. 62–72. [Online]. Available: http://games.stanford.edu/competition/misc/aaai.pdf

[4] "Stanford General Game Playing," 2014, URL: http://games.stanford.edu/ [accessed: 2014-12-05].

[5] S. Gelly, L. Kocsis, M. Schoenauer, M. Sebag, D. Silver, C. Szepesvári, and O. Teytaud, "The Grand Challenge of Computer Go: Monte Carlo Tree Search and Extensions," Commun. ACM, vol. 55, no. 3, Mar. 2012, pp. 106–113, DOI: 10.1145/2093548.2093574.

[6] O. Syed and A. Syed, Arimaa - A New Game Designed to be Difficult for Computers. Institute for Knowledge and Agent Technology, 2003, vol. 26, no. 2.

[7] K. Walędzik, J. Mańdziuk, and S. Zadrożny, "Proactive and Reactive Risk-Aware Project Scheduling," in 2nd IEEE Symposium on Computational Intelligence for Human-Like Intelligence (CHILI'2014), Orlando, FL. IEEE Press, 2014, pp. 94–101.

[8] B. Stevens and F. Lewis, Aircraft Control and Simulation. New York: Wiley, 1992.

[9] C. G. Eden, "Robotically assisted surgery," BJU International, vol. 95, no. 6, 2005, pp. 908–909.

[10] M. Buro, "The Evolution of Strong Othello Programs," in Entertainment Computing, ser. The International Federation for Information Processing, R. Nakatsu and J. Hoshino, Eds. Springer US, 2003, vol. 112, pp. 81–88.

[11] J. Mańdziuk, Knowledge-Free and Learning-Based Methods in Intelligenet Game Playing, ser. Studies in Computational Intelligence. Berlin, Heidelberg: Springer-Verlag, 2010, vol. 276.

[12] L. Kocsis and C. Szepesvári, "Bandit based Monte-Carlo planning," in Proceedings of the 17th European conference on Machine Learning, ser. ECML'06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 282–293.

[13] M. Świechowski and J. Mańdziuk, "Self-adaptation of playing strategies in general game playing," IEEE Transactions on Computational Intelligence and AI in Games, vol. 6, no. 4, Dec 2014, pp. 367–381.

[14] ——, "Fast interpreter for logical reasoning in general game playing," Journal of Logic and Computation, 2014, DOI: 10.1093/logcom/exu058.

Home Monitoring of Mental State With Computer Games

Solution Suggestion to the Mental Modern Pentathlon Scoring Problem

Pál Breuer; Péter Hanák; László Ketskeméty; Béla Pataki
Budapest University of Technology and Economics
Budapest, Hungary
e-mail: {breuer,hanak}@emt.bme.hu, kela@cs.bme.hu, pataki@.mit.bme.hu

Gábor Csukly
Department of Psychiatry and Psychotherapy
Semmelweis University
Budapest, Hungary
csukly.gabor@med.semmelweis-univ.hu

***Abstract*— As society is aging, an increasing number of elderly people is affected by cognitive problems. Early detection of mild cognitive impairment (MCI) is crucial for slowing down deterioration at an early stage. Improving detection would allow aging in place and thus more cost effective care. However, detection usually occurs too late. Clinical tests are expensive, not frequent enough, and give only a single snapshot of cognitive performance. Regular home monitoring of the changes in mental state would be important but clinical tests have not been developed for this purpose. In this paper, the use of computer games in measuring and maintaining mental wellness in a regular and voluntary way is proposed. Problems and potential solutions are presented, with special emphasis given on the sensor fusion problem caused by the various games present.**

Keywords-mental wellness; home health monitoring; serious games, Mild Cognitive Impairment (MCI), mixed data fusion.

I. INTRODUCTION

As society is aging (by 2060 more than 28% Europeans will be over 65 [1]), an increasing number of elderly people is affected by cognitive problems. With earlier detection of mild cognitive impairment (MCI) deterioration could be significantly slowed down at an early stage. Slow decay of mental abilities is a normal process, which affects already age group 40 of the population, and which increasingly progress with age. It is not easy to identify the stage at which the process becomes abnormal and the affected person requires serious attention, perhaps medical intervention. Cognitive tests are usually performed only if there are already some concerns in the family about someone's cognitive ability, but due to the natural denying effect (by the elder person, the family and the friends) detection typically comes too late.

Traditional, validated, paper-based tests constitute the gold standard but they have several drawbacks. To begin with, such tests require specialist centers and highly trained professionals. Therefore, there is a growing interest in the development of computerized cognitive assessment batteries [2][3][4]. But clinical tests, using either paper-based or computerized methods, are made quite infrequently, providing too sparse snapshots of the cognitive performance. More frequent and regular population screening would require exceedingly many professionals.

Regular home monitoring of changes in mental state offers a powerful alternative, even if it only allows relatively noisy and less targeted measurements. But, they have the advantage of frequent assessment and so the possibility of evaluating temporal trends. Current clinical tests are not suitable for this purpose. Therefore, new measurement methodology must be developed and validated specifically for this strategy. Given that need, recent years have seen a growing interest in the development of special computer games for cognitive monitoring or training purposes. Few such games have been developed; those aim to monitor and train a specific cognitive domain, e.g., verbal fluency [5], executive functions [6] or perceptual and motor functions [7]. A major challenge in this direction is that entertainment capability and measurement power pose contradictory requirements. There are three approaches to game development for older people:

- well-known, popular games (e.g., chess, tangram or tic-tac-toe [8], memory, freecell [9]),
- slightly modified special clinical tests (e.g., corsicube [9]) transformed into games,
- brand new games specially designed for this purpose [6].

Regular monitoring may be (1) controlled or (2) voluntary. Most elderly persons prefer to lead independent life as long as possible. Moreover, in the early monitoring period (before detecting any problem) they are mentally healthy. Therefore, controlled monitoring seems an undesirable option for them; it is expected to undermine independence and works only for a highly motivated minority. As increasingly more people (even the elderly) use computers, and many of them regularly play computer games, gaming activity could be exploited for measuring their performance in those games. In turn, that performance is related to their cognitive state, according to some experimental studies. The basic idea is the following: with regular use of computer games developed or modified specifically for elderly persons, we may be able to measure their mental changes and tendencies over time in an entertaining way, therefore, regularly and voluntarily.

The methods, problems, and possible solutions to those, that are presented in this paper, are based on a recent research project (M3W, Maintaining and Measuring Mental Wellness [9][10][11]). The final goal of the project is to develop a method for home monitoring mental state of elderly people, which is a very complex task. Therefore, only some subproblems and suggestions are presented in this paper. The considerations leading to the proposed architecture, the basic conceptual architecture of the system and some of the challenges are discussed. Among the numerous problems, this paper focuses on the special sensor fusion problem, which arises in the voluntary home monitoring scenario when several games (sensors) are present.

Due to the complexity of the project, some important problems, such as motivation of the players, are not addressed here in detail. However, some decisions were indeed influenced by motivational considerations, especially the use of several games, which gave rise to the sensor fusion problem, as noted earlier. Another hard problem just mentioned is game selection. The right balance must be found between entertainment capability and measurement power. Based on our pilot experiments, the game set is still evolving.

In that one year pilot study, more than 50 volunteers registered to take part and help evaluate the framework and the games developed at that time. Due to the voluntary nature of the project only about 20 of them played regularly for nearly one year. Of course, the parallel development of the program package was a drawback for the players. The average age of these regular players was 70.3 years, the standard deviation was 10.9 years. People played at home or in an elderly home. Eleven games had been developed and tested (two card games: FreeCell, Solitaire, one psychology-cal test: Corsi Test, two logical games: Graphs and Rabbits, three attention games: Fowler, Odd One Out, Pick One Out, one retention game: Memory Game, two language-skill games: Word Finder, Word Puzzle). The one year timespan has limited relevance on the timescale of mental aging, but some findings have already surfaced, which are clearly important for the long run as well. Parallel to the home monitoring pilot study, a clinical experiment on patients with mental problems (MCI, Alzheimer's disease, etc.) has also been running; those results are not discussed in this paper.

Section II describes the basic model of the suggested mental state evaluation system and describes the important problems. Section III gives the suggested basic detection method using a single computer game. Section IV addresses the special sensor fusion problem caused by the very different nature of games, and suggests a possible solution. Section V summarizes the findings and gives the directions for further work.

II. BASIC MODEL, PROBLEMS

The basic conceptual architecture of the proposed system is shown in Figure 1. The final goal is to provide appropriate long-term feedback to the user (or to the caregiver, family member, medical expert, etc.). Short-term feedback is for motivation to continue participating in the monitoring ("Well done!", "Play some more games!"). Long-term feedback is the result of the change detection estimation: whether a significant change of mental state has occurred or not.

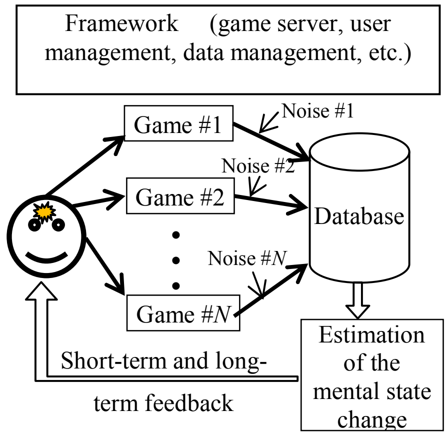


Figure 1. Basic conceptual model of the cognitive state estimation system

Beyond the general problems of such systems (e.g., data privacy concerns), this approach has its special challenges, some of these are given:

1) How to measure the ***cognitive performance*** using computer games?
2) How to cope with the sometimes heavy ***noise*** of the uncontrolled (home) measurement environment?
3) How to ***motivate*** people to take part in the long run?
4) How to compare performance shown in different games, which is basically a special ***sensor-fusion*** problem?

To ***measure the cognitive performance*** three principles are followed:

- To ensure the opportunity of measurements, proper serious games are selected, special ones are developed or clinical tests are modified taking into account the special requirements. Usually, games are modified to improve measurement capability; and tests are modified to be more entertaining. Most of them are logical puzzles, or they need the intensive use of the short-term memory (which is one of the best indicators of MCI), but other important parameters (attention, execution, language skills, etc.) are targeted as well. Two basic parameters are measured: the solving time of the puzzle and the good/bad steps taken during the solution. Currently only successful solutions are measured, for future work there are possibilities in the evaluation of the failed ones as well.
- Because the measurement of the mental state on an absolute scale is very hard, only the change in the person's performance is to be detected. For measuring a change, a reference is needed. There are two possibilities: the performance could be compared to a reference group; or it could be compared to a previously measured reference of the same person. Because the inter-personal comparison is affected by several parameters unknown in this voluntary, uncontrolled method (education, physical abilities, family conditions, profession, environment, etc.) the comparison in time to his/her own previous

performance was chosen. (However, since many people like to compare their own abilities to others' and to compete with others, such functionalities will be offered as well.)

According to our experiments, the ***noise*** can be modeled using two terms: (1) zero mean low level noise caused by the random differences between the consecutive puzzles and by minor environmental disturbances (2) major disturbances. This first term is eliminated by the averaging effect of the evaluation method (see section III): several game results are evaluated together. The second term is an impulse like noise caused by the physiological, environmental and social disturbances resulting in outliers (for example, the telephone is ringing; the person has to use the bathroom, a storm is arriving, neighbor is coming, etc.). This second problem is solved by a filtering step, the outliers are simply rejected; they are not used in further evaluation steps.

Early detection is the purpose; but the main problem is that nobody knows when the abnormal change will happen; maybe in some persons' life never. Therefore, the ***motivation*** must be managed probably for many years. It is a very complex problem itself; only some aspects are discussed here. Among several other aspects, one basic assumption is that although there is an extrinsic motivation that everybody wants to sustain mental abilities and an independent life of good quality, but generally it is not enough in the long-run. There must be intrinsic motivations too, e.g., entertaining ways of measurement, and short-term feedback (Figure 1) given to the user to encourage further playing (e.g., scoring or encouraging messages such as "Well done!" could generate motivation). Unfortunately, most people do not enjoy the same game for years. Therefore, in different time periods different games will be played by the same person. Not to destroy the level of motivation several games are offered (Figure 1); and the performance measured using different games should be somehow compared to each other (Figure 2). This implies a ***sensor fusion and estimation problem***, where the games are the sensors. It is similar to the modern pentathlon scoring problem, where performances in very different sports (fencing, show-jumping, running, swimming, shooting) have to be measured in one unified scoring scheme. In our case, the problem is even more complex because the same game could be played using different settings (e.g., different number of cards in the well-known memory game, see Figure 2); therefore, each setting creates a new game from the measurement point of view. All these games should be compared to each other. In Figure 2 only the results of a given player in the 3 most frequently played games are shown. (In the figure different games are marked by different colors; different settings of the same game are marked by different symbols.) The proposed solution for solving this problem is detailed in Section IV.

III. DETECTION OF THE MENTAL STATE CHANGE

For detection of the mental state change, the comparison in time to the player's own previous performance was chosen. First the evaluation method is considered when only one game (always with the same settings) is played.

Because the two-term noise is present, the effect of the impulse noise, the outliers should be eliminated first. For that purpose, the time between two consecutive elementary events during the solution (e.g., mouse clicks) is analyzed. Because the impulse noise is usually caused by an extreme interrupt, if the longest time between two such actions is too high in comparison to the average action time, then this game was probably seriously disturbed: it is taken as outlier and is rejected. In Figures 2-5, only the results in outlier-free and successful games are shown.

The other noise term, the small natural fluctuation must be coped with as well. For that reason, the change detection cannot be based on the performance measured in a single game; some sets of parameters should be compared. The goal is to detect the decline of performance, but in some periods improvements can occur as well. The assumption is that the decline is preceded by a period where no improvement is present; the situation is stable or deteriorating very slowly. Therefore, a reference set is selected, which is the group of consecutive games in which the person had stable performance (Figure 3).

It is reasonably assumed that the short-term fluctuations due to tiredness, puzzle-hardness, etc., are zero-mean, stable independent random variables. The puzzle hardness is a zero-mean, stable random variable, because the same game is used with the same parameters, and the current puzzle is selected randomly. The short-term change of cognitive power is again a zero mean random variable, because it models the effects of the random changes of the environment, tiredness and health. The very slow long-term change of the cognitive state is modeled differently. Therefore, if a change is detected in one of the integral characteristics (mean, median, standard deviation) or generally in the distribution of the composite random variable (mental-state plus game-noise), it is caused by the slowly changing component modeling the mental state.

Let the performance observation based on the game played in time t_k be $\pi(t_k)$, $k=1,2,\ldots,K$ (this could be the score, the number of steps, etc.). Decrease in the values indicates decreasing performance. Significant change in the time series cannot be stated while this seems to be a realization of an independent and identically distributed (i.i.d.) sample. Several statistical tests can be applied for testing the null hypothesis that the data is i.i.d. Such tests are the difference sign test, the turning point test and the rank test [12][13]. If the null hypothesis cannot be rejected, no significant change in the player's performance could be stated.

A less rigorous requirement is that we cannot justify a change, if the time series is weakly stationary; i.e., uncorrelated with constant expected value and variance. This null hypothesis can be tested with the Dickey-Fuller test [14]. If the time series seems to be non-stationary the change of the player's performance is detected.

 ISBN: 978-1-61208-382-7

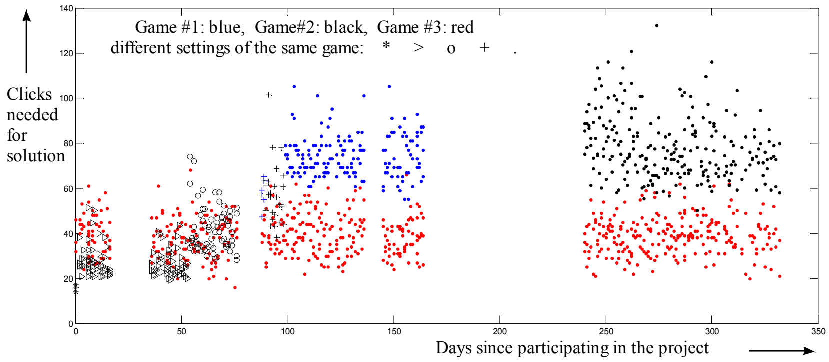


Figure 2. Typical example of a player's performance vs time. The time gaps are caused by travelling to realatives or by other reasons.

Using the Mann-Whitney U or the Kolmogorov-Smirnov two-sample tests, the comparison of the distribution of the reference subset with the distribution of the currently examined subset of the time series could be performed. If we detect a difference between the distributions of the two sub-samples; and the current part of the series has smaller average (of ranks, of scores, etc.), then the player shows performance degradation.

These statistical hypothesis tests were used to check the distribution of the composite random variables. The tests were implemented in Matlab and SPSS. The following findings were obtained:

- The resulting performance parameter is not normally distributed according to the Lilliefors test.
- The time gaps (several users produced 7…60 day gaps) did not change significantly the distribution of the random variable examined (see Table I).
- Several statistical tests were applied to compare the distribution of the reference period data to the current period data of users, who played some hundreds of games in the nearly one year period. (Two-sample Kolmogorov-Smirnov test, Mann-Whitney U test, Wilcoxon signed-rank test). The results confirm that both the stability and the change in the parameters are reliably estimated by the statistical tests. All these tests gave coherent results; later the performance of the different tests should be examined, and the best one should be selected.
- As an alternative to the two-sample statistical tests, a runs test on the sequence of observations was performed to prove the null hypothesis that the values came in random order, against the alternative that they did not. The runs test gave the same result: if there was no significant difference between the distributions of the reference and the current subsets the runs test did not rejected the randomness hypothesis, if there was difference between them, the runs test rejected the hypothesis.
- In some cases, when starting a new game a learning phase occurs, in which the results are improving. The reference is meaningful only when the performance has stabilized The stability could be defined the same way as the stationarity of the current performance. Evaluating these time series has proved that hypothesis testing detected the change of the cognitive performance as well.

In Figure 4, a time series measured during the learning phase is shown. The hypothesis tests accepted the same distribution null hypothesis (the first 30 games' data compared to the current set) for all the sets up to the 187th game; and rejected the null hypothesis for all the sets from the 260th game.

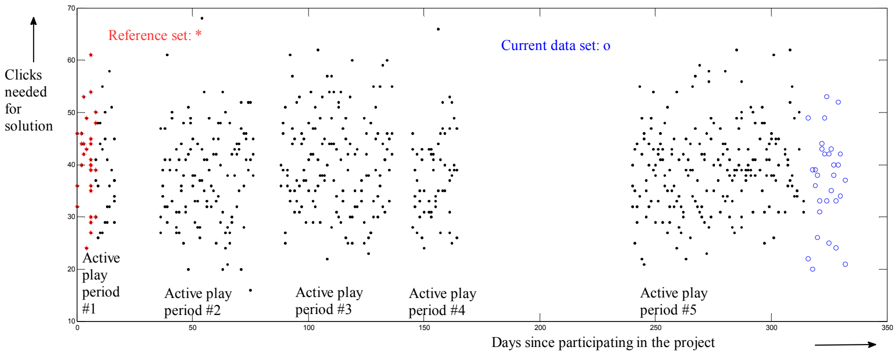


Figure 3. The current performance is always compared to the reference set

 ISBN: 978-1-61208-382-7

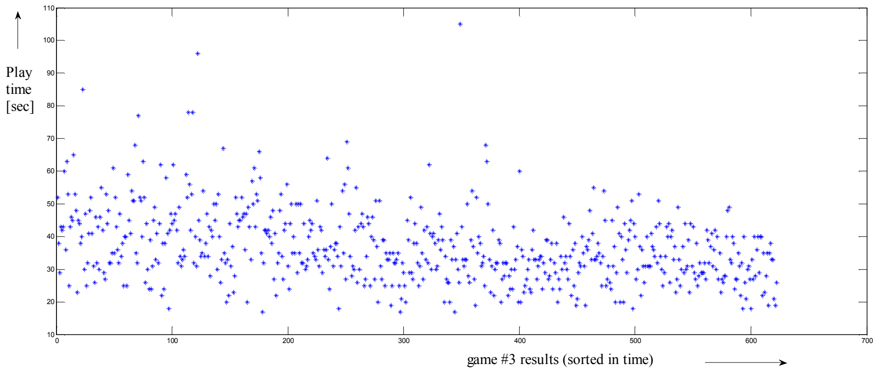


Figure 4. Nonstationary series of play times in the learning phase

IV. SOLVING THE SENSOR FUSION PROBLEM

Computer games are proposed for detecting change in mental state as soon as possible. For motivational purposes several different games should be offered (and different settings of the same game could be used). Because of the voluntary nature there is no guarantee that the same person will play with the same game in the long-run. In our pilot only a few voluntary participants played continuously the same game for this nearly one year period. (Results shown in Figure 4 belong to a participant who played about 2 games per day with the same game for nearly one year!) Most of them changed the game or at least changed the settings of a given game (harder or easier). Therefore, in different – overlapping and non-overlapping – time periods different sensors (games) are available to measure some parameters connected to mental state. Because the more data we have the more reliable the detection of the cognitive state; therefore, every effort is worth to keep all the data.

In this section, a possible solution of that sensor-fusion problem is proposed. The basic idea is that proper linear normalization of the performance measures results in parameters, which are compatible with the normalized parameters of other games. The normalization is based on the reference set of the current game. Let the performance observation using game m in time t_k be $\pi_m(t_k)$, the average of the performance measures of this game's reference set be denoted by $\mathrm{avg}(\pi_{m\mathrm{REF}})$, the standard deviation of this reference set is $\mathrm{std}(\pi_{m\mathrm{REF}})$. The normalization:

$$\pi_{mn}(t_k) \;=\; (\pi_m(t_k) - \mathrm{avg}(\pi_{m\mathrm{REF}}))/\mathrm{std}(\pi_{m\mathrm{REF}}),\; m=1,\ldots,N \quad (1)$$

After normalizing all the parameters of the different games the combined time series is constructed by simply sorting the data in time.

$$\{\pi_{COMB\mathrm{n}}(t_1),\ldots,\pi_{COMB\mathrm{n}}(t_k)\}=\{\pi_{m1\mathrm{n}}(t_1),\ldots,\pi_{mk\mathrm{n}}(t_k)\}$$

$$t_1 < t_2 < \ldots < t_k \quad (2)$$

The block diagram of the suggested idea is shown in Figure 5. The resulting combined time series derived from the data of Figure 2 is shown in Figure6

Using the time series of combined data gives very similar results as using the data of one game only. In Table I the null hypothesis of having the same distribution of the data subsets compared are shown in two ways. In both evaluations the reference set comes from the first 30 observations of the active play period 1, the comparison is made to the first 30 observations of the 2nd, 3rd, 4th, 5th active play periods, respectively. The difference is that in the first experiment only the Game#3 data are used, and in the second experiment the combined data are used.

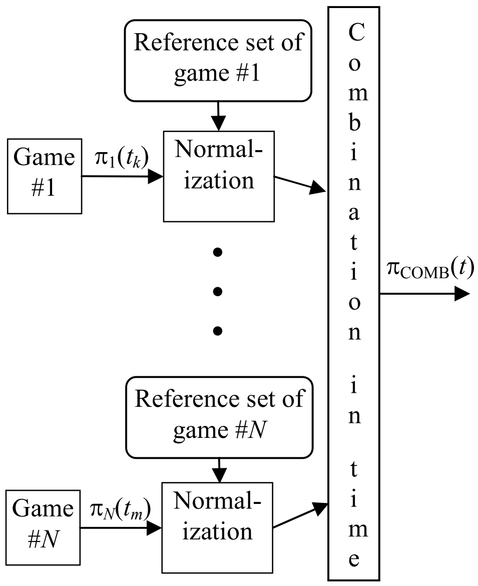


Figure 5. The normalized performance parameters of the different games are combined to form one composite time series

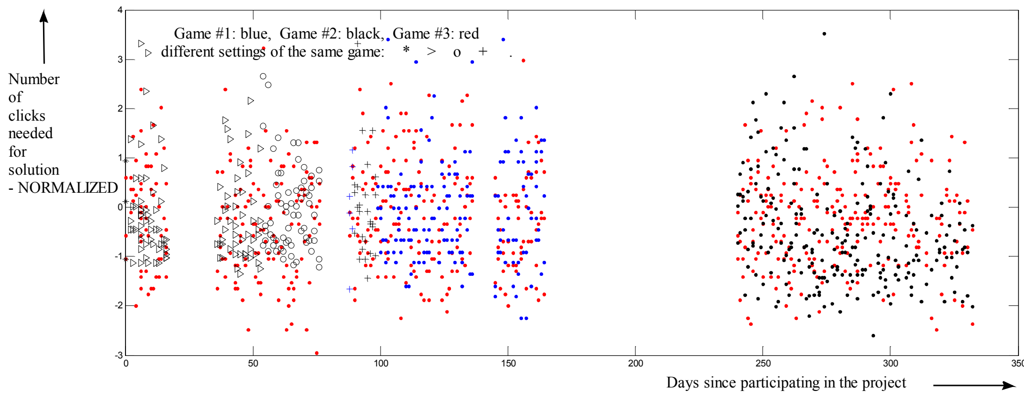


Figure 6. Normalized and combined data

In Table I, the acceptance or rejection (on the p=0.05 level) of the null hypothesis are shown.

TABLE I. RESULTS OF TWO-SAMPLE KOLMOGOROV-SMIRNOV TESTS: REFERENCE SET SHOWN IN FIGURE 2 (FIRST 30 OBSERVATIONS OF ACTIVE PLAY PERIOD 1) COMPARED TO THE FIRST 30 DATA OF EACH ACTIVE PLAY PERIOD

	Game #3 data only		***Combined data***	
Reference: active play period 1 compared with	***Null hypothesis accepted:0, rejected: 1***	***Probability value: p***	***Null hypothesis accepted:0, rejected: 1***	***Probability value: p***
Active play period #2	0	0.43	0	0.11
Active play period #3	0	0.76	0	0.20
Active play period #4	1	0.03	1	0.01
Active play period #5	0	0.54	0	0.06

Although in the first experiment only Game#3 data were used and in the second one combined data were used, they resulted in the same acceptance/rejection scheme although the pure one-game only data gave higher probability values.

V. CONCLUSION AND FUTURE WORK

Home monitoring of changes in mental state using computer games was proposed in a regular, voluntary scenario; some of the problems were analyzed and solutions were proposed. The system assumes voluntary participation; therefore, several different games were developed to sustain motivation in the long run. In the game battery there are both well-known, popular games and modified clinical tests, it is continuously evolving.

For cognitive performance change detection, the within-subject comparison is proposed. The reference set of performance results are to be compared to the current set of results using statistical hypothesis tests. The null hypothesis is that the two sets came from the same distribution. Until the null-hypothesis cannot be rejected, the stability of mental state could be assumed.

Because of the large number of diverse games, the problem of unifying the different data should be solved as well. This problem is basically a sensor fusion one. Proper linear normalization using the reference set of each game is proposed for producing a mixed time series that could be used for our detection purposes as well.

In the future

- a pilot has to be launched to validate the method using further clinical tests,
- the most appropriate games for both entertainment and measurement should be investigated,
- the feasibility of multiplayer games is to be analyzed,
- the potential in the evaluation of the failed games should be investigated,
- the best statistical hypothesis test is to be identified.

ACKNOWLEDGMENT

This research was performed in the Maintaining and Measuring Mental Wellness (M3W) project, supported by the AAL Joint Programme (ref. no. AAL-2009-2-109). The authors also gratefully acknowledge the contributions of their project partners in Greece, Luxembourg, Switzerland and Hungary.

REFERENCES

[1] Population structure and ageing. Eurostat, May 2014, http://epp.eurostat.ec.europa.eu/statistics_explained/index.php /Population_structure_and_ageing [retrieved: January, 2015]

[2] Cantab test battery, Cambridge Cognition, http://www.cambridgecognition.com/ [retrieved: January, 2015]

[3] MindStreams, http://www.mind-streams.com/ [retrieved: January, 2015]

[4] T. Dwolatzky, "The Mindstreams computerized assessment battery for cognitive impairment and dementia", PETRA'11 May 2011, pp. 501-504, ISBN: 978-1-4503-0772-7.

[5] H. Jimison, M. Pavel and T. Le, "Home-Based Cognitive Monitoring Using Embedded Measures of Verbal Fluency in a Computer Word Game" 30th Annual International IEEE EMBS Conference, Aug. 2008, pp. 3312-3315. doi: 10.1109/IEMBS.2008.4649913

[6] A. López-Martínez et al, "Game of gifts purchase: Computer-based training of executive functions for the elderly", IEEE 1st International Conf. on Serious Games and Applications

for Health (SeGAH), Nov. 2011, pp. 1-8, Print ISBN:978-1-4673-0433-7 DOI: 10.1109/SeGAH.2011.6165448

[7] K. Ogomori, M. Nagamachi, K. Ishihara, S. Ishihara and M. Kohchi, “Requirements for a Cognitive Training Game for Elderly or Disabled People”, Int. Conf. on Biometrics and Kansai Engineering, Sept. 2011, pp 150–154, E-ISBN 978-0-7695-4512-7, Print-ISBN 978-1-4577-1356-9, DOI: 10.1109/ICBAKE.2011.30

[8] V. Menza-Kubo and A. L. Morán, “UCSA: a design framework for usable cognitive sstem for worried-well”, Personal Ubiquitous Comput. Vol. 17, Issue 6, Aug. 2013, pp. 1135-1145. ISSN:1617-4909. DOI: 10.1007/s00779-012-0554-x

[9] Maintaining and Measuring Mental Wellness (M3W) AAL Joint Programme project (ref. no. AAL-2009-2-109) https://m3w-project.eu/ [retrieved: January, 2015]

[10] E. Sirály et al., “Differentiation between mild cognitive impairment and healthy elderly population using neuropsychological tests”, Neuropsychopharmacol Hung. (3), Sept. 2013, pp. 139-46.

[11] P. Hanák et al., “Maintaining and Measuring Mental Wellness”, Proc. of the XXVI. Neumann Kollokvium, Nov. 2013, pp. 107-110 (in Hungarian).

[12] P. Brockwell and R. Davis, Introduction to Time Series and Forecasting, Springer, New York, 1996.

[13] M. G. Kendall, and A. Stuart, “The Advanced Theory of Statistics”, Vol. 3, Griffin, London, 1976.

[14] D. A. Dickey, W. A. Fuller, “Distribution of the Estimators for Autoregressive Time Series with a Unit Root”. J. of the American Statistical Association 74 (366), Jun. 1979, pp. 427–431.

Exploring Facets of Playability: the Differences between PC and Tablet Gaming

Uttam Kokil
College of Humanities and Social Sciences
Kennesaw State University
Marietta, GA 30060, USA
ukokil@kennesaw.edu

José Luis González Sánchez
LIVE Research Group
University of Granada
E-18071 Granada, Spain
joseluisgs@ugr.es

***Abstract*–With the advent of mobile devices, game consoles and computers as gaming platforms, the gaming industry is growing at exponential rates. Players are now accessing the latest video game entertainment in more than one digital medium thus expanding a player's game time and making video games the number one leisure choice. If one of the goals of the player is to derive a quality of experience that highlights enjoyment, it is important to understand the relationship between player experience and pleasurable game play, not just in terms of game play, but also, with respect to the gaming platform. This study aims to examine the difference in player experience when playing a video game on a tablet versus a personal computer. By triangulating physiological data of emotional responses using galvanic skin response, heart rate, and subjective-feelings data of facets of player experience, this paper aims to determine whether player experience is affected by two different kinds of gaming input controls, a computer keyboard with mouse, and a tablet touchscreen. Data will be analyzed and reported in future work. This paper provides an overview of the literature survey and methodology.**

Keywords – *Tablet gaming; player experience; gaming platforms; physiological evaluation.*

I. INTRODUCTION

Game designers are constantly challenged to build more sophisticated interactive gaming environments to keep up with players' increasing demand. An interactive player experience provides the user with a hedonic pursuit of stress relief, cognitive challenges and enjoyment [1]. This interactivity allows the players to experience the narrative at their own pace and as such contributes to immersive features in the gaming environment. The gaming platform affects interactivity as the controller connects the player with the gameplay [2]. Facets of Playability are used to measure and evaluate the interaction experiences between the game and the player [21]. Player experience is a consequence of the interaction experiences. Flow state is an important factor for a player to attain optimal player experience. Player experience can be positive or negative, that correlates with the user's perceived enjoyment of the gaming experience. Lazarro's four Fun Keys are examples of how experience has become a new area of economic development in game development [3]. According to Ermi et al. [4], game play experience can be defined, as a mix of "player's sensations, thoughts, feelings, actions and meaning-making in a game play setting." This sets a ground for player experience by providing entertainment, escape, competition and challenge. As each aspect directly relates to the user, the type of game that a user chooses to interact with is an important factor in the area of player experience [5]. Marchland et al. [6] state that although consoles are considered the preferred gaming platform, mobile gaming is having a larger impact on the overall market putting the burden on mobile game developers to perform. PC games are still popular. They advocate the need for more research in the domain of gaming platforms. Csikszentmihalyi's [20] approach of Flow has helped game researchers develop new insights of player experience, as it will provide more clarity on areas of immersion and flow. The rest of this paper is organized as follows: Section II is dedicated to the literature reviews of related works. Section III describes the aims of the experiment. Section IV addresses the methodology approach. Section V has to do with the analysis of the results. The acknowledgement and conclusions wrap up the article.

II. LITERATURE REVIEWS

According to Ritterfeld et al. [7] "play is not a random activity selected to overcome boredom, but rather a rational choice." Individuals, therefore, make choices in how they relate to video game products based on the narrative, and interactivity. Players have an opportunity to choose their unique experiences by selecting a particular game to play for maximum pleasure [8]. The game designer creates a digital environment that allows players to immerse themselves within the gaming environment, provides opportunities for agency by allowing the gamers to choose their experiences, and transforms the players by allowing the sensations of becoming different people or objects [9]. These transformational experiences specifically do not provide a holistic framework for defining player experience (PX). According to Kidd [3], people are compelled to use technology for three reasons: utility, symbolism and experience. People want a gadget that serves a purpose, looks good, and engages the user's attention. His study examined which characteristics of "technology-mediated experience" would captivate the user. By observing the behavior of children and adults at an interactive exhibit and following up with focus groups, the study concluded that compelling technological experience requires the following dimensions: social engagement, sensation/drama, and self-expression/challenge. Social engagement carried the least weight, while self-expression and challenge were the most compelling. Participants wanted an experience that tested their mental/physical skills, allowed for creative expression, was tactile, and allowed for escape from the ordinary. The researcher concluded that a meaningful experience is grounded within the user and not within the technology. Finally, the experience of the user was individualized, and relied on both intrinsic and extrinsic cues for optimal satisfaction. Attributes from this study coincide with findings from Sweetser et al. [10] game flow model, describing how a player's enjoyment relates to game flow. Game flow criteria associated with core elements related to a positive player experience can be used to find issues and predict the popularity of the game; however, the model also "serves as a starting point for academics and game developers to understand enjoyment in games and to conduct further research into understanding, evaluating and designing enjoyable games." By including elements of concentration, challenges, player skill, clear goals, control, feedback, social interaction and immersion, game designers can predict a game's success. In order to validate these criteria, Sweetser et al. [10] evaluated two real time strategy games that represented a poorly rated and highly rated game. The game evaluation concluded that the higher rated games met more of the Game Flow Model criteria than the poorer rated video games. Challenge, as a core element of player experience, requires the game to be both intrinsically motivating, and goal orientated for a user to feel pleasure. Abuhamdeh et al [1] performed two separate studies to validate this claim. Study 1 observed the relationship between the perceived challenge, skills and the level of enjoyment, while Study 2 examined the strength of these relationships. The first study found that there was a relationship between challenge and level of enjoyment; however, a user's perception of his/her skill did not influence the level of enjoyment. Study 2 verified that challenge was a strong indicator of enjoyment for intrinsically motivated, goal directed activities. Trept et al. [11] agree that there is a strong relationship between enjoyment, and game challenge. Their study revealed that a user's subjective experience of enjoyment relies on both experiential and psychological aspects related to the player's sense of accomplishment, self-efficacy, and the challenges associated with game play. Malke et al. [12] examined human computer interaction through the assumptions that a user's experience influences their assessment of the system and components that interact with each other in distinct ways, which make up the user's experience. The Component User Experience model has the user completing a specific task within a specific context and time. These interactions can be influenced by either instrumental or non-instrumental systems that produce an emotional response within the user that directly influences the appraisal system. The results indicate the user's overall judgment of the experience. The perceptions related to instrumental systems relate directly to system ease of use, functionality, while non-instrumental systems include perceptions associated with the visual aesthetics, the look and feel of the gaming platform, haptic and symbolic quality.

Takatalo et al. [8] further the definition of system assessment with a holistic model that includes both the experiential and psychological aspect of enjoyable gameplay. They argue that player experience involves not only game flow, but also, presence and involvement. The idea behind the Presence-Involvement-Flow Framework (PIFF2) is that "players must invest time, effort and attention into a game in order to get any relevant experience from it." Takatalo et al. [20] explain that player experience can be measured using dependent variables such as *presence, involvement, cognitive evaluation, and emotional outcomes*.

Interest in the game, and the importance of the experience both work to establish a cognitive connection, which garners the meaning and relevance of that experience for the individual. Involvement therefore measures the quality of the relationship between the game and the gamer.

This involvement can manifest through the game narrative and the emotions of the user when playing the game. Tavinor [13] explains that there is a causal connection between fiction and emotional response. Players willingly enter a world of fantasy in which they interpret appropriate actions and reactions to the perceived stimuli presented on screen. Van Aart et al. [14] take a different approach to emotional interactivity by studying the areas of boredom and curiosity. By designing game play through emotional cues, players are better able to navigate the gaming environment. User's emotions are used to make decisions, enhance cognitive skills, and maintain memory. Boredom can be defined as both a lack and an overload of stimuli. This boredom can be intensive, collative, and affective in nature. Players are drawn toward emotional, exciting experiences and draw away from experiences associated with waiting. The duration, commonness, and user expectation of the waiting experience can deter a user's overall perspective of the game play. While boredom stagnate a user's cognitive experience, curiosity pushes one to explore, take risks, and motivate. A zone of curiosity is required to maintain alertness in the game play environment. Emotions are invoked through the images shown on the screen, the sound track, and the challenges presented to the user. As such, the element of involvement does not work alone, but in tandem with a feeling of presence.

Browne et al. [15] found that touch screens fell behind in speed and performance when comparing three different multi-touch game interfaces on an iPod Touch. This multi-touch interface "offers user interface capabilities beyond physical buttons such as accelerometers and touch screens capable of recognizing the movements of multiple fingers." Mobile games required a configurable touch interface that had pre-specified criteria such as diagonal direction touch gestures for game play that related directly to the virtual properties of the game narrative. Participants chose an accelerometer console most frequently in the experiment because it allowed for best performance.

Gleeson et al. [16] found the same results when they compared the effectiveness of the touch screen over the use of a mouse and keyboard as a tool to interact with information systems. The mouse performed better in terms of movement time for small-targeted areas. Where the target is medium or large sized, the touch screen interface performed at an equivalent rate. A mouse has a minimal error rate compared to touch screen and higher interactive accuracy rate. Moreover, it is believed that touch screens cause physical wrist and finger fatigue [16].

III. Aims

The aim of this study is to conduct a comparative study and examine *playability* in two game environments, using a PC and a touch screen tablet. A game interface acts as a bridge between the gaming system, and player's experience. This bridge provides a lens through which the player can assimilate the rules and pacing of the game narrative; hence, it becomes important to look at the differences between mouse and click, and touch experience.

IV. Methodology

The overall goal of this study was to determine the overall effect a gaming platform has on player experience. In order to provide both a subjective and objective approach to the research, Mandryk et al. [17] suggest using "subjective reporting through questionnaires, and interviews because they are generalizable." Mandryk explains that, subjective data collection that has both quantitative and qualitative results data provide a more robust experimental design; however, the subjective data provide only partial results as they lack certain patterns [18]. Physiological data collection is therefore required to validate the player experience.

A total of 14 participants were recruited from a Midwest university in the USA to play the strategy digital game *"Plants vs. Zombies"* on a PC (Windows) and OSX tablet (iPad4 retina) respectively for the purpose of data collection. Two instruments were used during the pre-test session such as *demographic survey*, *mood questionnaire*, while in the post-test session the following instruments were used to measure the dependent variables: *Self Assessment Manikin* to measure emotional responses [23], *Facets of Playability* to evaluate components of player experience [21] and *Game Enjoyment Questionnaire* (FUGA) to measure game enjoyment [22]. In addition to subjective-feelings questionnaires, we collected objective data to evaluate emotional reactions of valence using pulse rates while arousal was captured using electrodermal activity (EDA). The independent variables were *screen size*, *mode of interaction*, *screen resolution*, and *products*. First we ran a pilot study to verify the following items: instructions were clear and comprehensible; tools and game stimuli for capturing relevant data were in working conditions; selected questions of the validated questionnaires that were relevant to the research questions of this study. The sample frame chosen were between 18-35 years old. This accounts for 32% of digital game players in the USA [19]. The sampling and recruitment were conducted by network and convenience methods. The design of the study was within-subjects ANOVA test whereby the same participants took part in two different experimental conditions, playing the same game in a PC and a tablet environment.

When the participants arrived in the lab, the researchers went through the research protocols before they could sign the consent form. Biopac Systems electrodes were placed on each subject's second and middle finger of the non-dominant hand to record EDA, and onto the middle finger of the other hand to record pulse. Prior to starting the game, baseline pulse and EDA were recorded for the first 5 minutes and, thereafter, both physiological data were captured during gameplay. After 10 minutes of gameplay, the participant was instructed to fill a self-report questionnaire to record his/her emotional responses (SAM) and another questionnaire to self-report their challenges and skills at a given point. The participant also filled the PIFF2, GEQ, and Facets of Playability questionnaires at the end of game play. Playability is a property that characterizes player experience in games [21]. After that, each participant took a rest for 5 minutes before switching to the other game platform. The same procedure of data collection was followed while playing the game on the other platform. The disadvantages of this design were (i) order effects and (ii) practice effects (iii) fatigue effects. Order effects refer to the actual order the treatment is administered. In fact, participants were assigned randomly to the tablet and PC game to counterbalance the order effect. Similarly, to avoid any practice and fatigue effects, participants were instructed to take a break in between each treatment.

V. Analysis

A preliminary analysis of the *Facets of Playability* questionnaire was performed. Paired t-tests were conducted to compare the mean values of the five components that characterize player experience (Figure 1).

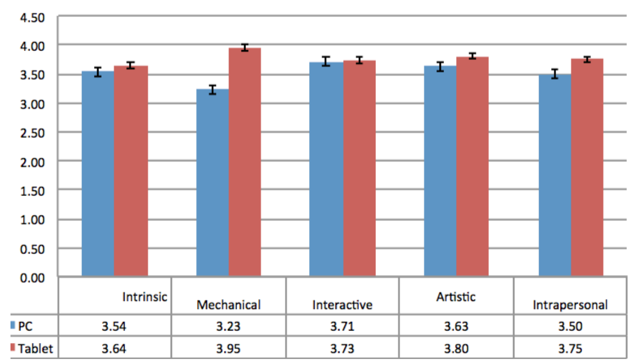


Figure 1. Player Experience results based on Facets of Playability

Figure 1 shows that the mean values for intrinsic, mechanical, interactive, artistic, and intrapersonal playability were greater when the players used a tablet device as compared to a PC. A hypothesis testing was conducted such that the null hypothesis Ho: $\mu_d = 0$ (difference of the means is equal to zero); the alternative hypothesis, Ho: $\mu_d \neq 0$ (difference of the means is not equal to zero). The results of the paired t-tests conclude that they were not statistically significant. We report the following probability-value for each component: intrinsic (p-value 0.770 > standard alpha level of 0.05); mechanical (p-value 0.168> alpha level 0.05); artistic (p-value 1.00 > alpha level 0.05); interactive (p-value

 ISBN: 978-1-61208-382-7

0.393 > alpha level 0.05); intrapersonal (p-value 0.371 > alpha level 0.05). Since the p-value is greater than our standard alpha level 0.05, we fail to reject the null hypothesis. The preliminary results reveal that there is a positive trend towards an optimal player experience using a tablet but based on statistical analysis, that trend fails to reach statistical significance. This implies that observed mean differences may still be reasonably attributed to chance rather than to the type of platform used. This is because we have fairly small effect sizes and few data points. We only had data for 14 individuals. In order to be able to detect at least a medium effect size of the type of platform used and how it affected players' experiences, we should have had data from at least 34 participants according to G-power analysis. The analysis of physiological data and other self-report questionnaires are currently underway.

VI. CONCLUSION

The results of this preliminary study suggest that Tablet games provide better user experiences in the facets of intrinsic, mechanical, interactive, artistic, and intrapersonal playability. We can note that Mechanical playability was more pronounced in the Tablet environment as compared to the PC. Considering practical significance, it is clear that the mode of interaction, i.e., the touch screen, brings a different kind of user experience that is not achieved with the other game platforms. Mobile games are the new consoles offering portable experiences. Users have the flexibility to play games on their tablets from virtually any location at any time.

REFERENCES

[1] S. Abuhamdeh and M. Csikszentmihalyi, "The importance of challenge for the enjoyment of intrinsically motivated, goal-directed activities," on Personality and Social Psychology Bulletin, vol. 38, (3), 2012.

[2] M. Wysocki, Ctrl-Alt-Play Essays on Control in Video Gaming. Jefferson, NC: McFarland Publisher, 2013.

[3] A. Kidd, "Technology experiences:What makes them compelling?" Hewlett-Packard Laboratories, Bristol, UK, 2002.

[4] L. Ermi and F. Mayra, "Fundamental Components of the Gameplay Experience: Analysing Immersion." DiGRA Changing Views - Worlds in Play, 2005.

[5] R. Koster, Theory of Fun for game Design. Sebastopol, CA: Paraglyph Press, 2005.

[6] A. Marchand and T. Hennig-Thurau, "Value Creation in the Video Game Industry:

Industry Economics, Consumer Benefits, and Research Opportunities," on The Journal of Interactive Marketing, *27,* pp.141-157, 2013.

[7] U. Ritterfeld and R. Weber, "Video Games for Entertainment and Education," in P. Vorderer & J. Bryant (Eds.), Playing Video Games-Motives, Responses, and Consequences. Mahwah, NJ: Lawrence Erlbaum, Inc., pp. 399-413, 2006.

[8] J. Takatalo, J. Jyrki, J. Kaistinen, and G. Nyman, "User Experience in Digital Games," in I. Pavlidis (Ed.), Human Computer Interactions. Croatia: In Tech. [Online] http://www.intechopen.com/books/human_computer_interaction/user_experience_in_digital_games

[9] C. Fencott, M. Lockyer, J. Clay, and P. Massey, Game Invaders: the theory and understanding of video games. Hoboken, NJ: Wiley, 2012.

[10] P. Sweetser and P. Wyeth, "GameFlow: A model for evaluating player enjoyment in games," in Computers in Entertainment - Theoretical and Practical Computer Applications in Entertainment, vol. 3 (3), 2005.

[11] S Trepte, L Reinecke, "The Pleasures of success: Game-related efficacy experiences as a mediator between player performance and game enjoyment," in Cyberpsychology, Behavior, and Social Networking, vol. 11 (9), 2011.

[12] S. Mahlke, and M. Thuring, "Studying antecedents of emotional experiences in interactive contexts," on B. Begole (Ed), The SIGCHI Conference on Human Factors in Computing Systems, New York, 2007, pp. 915 - 918.

[13] G. Tavinor, "Video games, fiction, emotion," in The 2nd Australasian conference on Interactive entertainment Sydey: Creativity and Cognition Studios Press, 2005, pp. 201-207.

[14] J. van Aart, B. Salem, C. Bartneck, J. Hu, and M. Rauterber, "Designing for experience: Arousing boredom to evoke pre-defined user behaviour." The sixth Design and Emotion Conference. Hong Kong, 2008.

[15] K. Browne, and C. Anand, "An empirical evaluation of user interfaces for a mobile video game," in Entertainment Computing, pp. 1-14, 2011.

[16] M. Gleeson, N. Stanger, and E. Ferguson, "Design strategies for GUI items with touch screen based information systems," ISSN 1172-6062, 2004. [Online]. Avaliable from: http://otago.ourarchive.ac.nz/bitstream/handle/10523/994/dp2004-02.pdf?sequence=3&isAllowed=y

[17] R. L. Mandryk and S. Atkins, "A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies," in International Journal of Human-Computer Studies, vol. 65 (4), April 2007, pp. 329-347.

[18] R. Mandryk, "Evaluating affective computing environments using physiological measures," in Proceedings of CHI'05 Workshop on Evaluating Affective Interfaces-Innovative Approaches, Portland, Oregon, pp. 1-4.

[19] "Game Player Data, Entertainment Software Association," [Online]. Avaliable from: Retrieved on Oct 26, 2014. http://www.theesa.com/facts/salesandgenre.asp

[20] J. Takatalo, "Psychologically-Based and Content-Oriented Experience in Entertainment Virtual Environments," (Doctoral dissertation), 2011, University of Helsinki, Finland.

[21] J. González Sánchez, F.L. Gutiérrez Vela, F. Montero Simarro, and N. Padilla-Zea, "Playability: analysing user experience in video games," Behaviour & Information Technology, vol. 31 (10), pp.1033–1054, Aug 2012.

[22] W. IJsselsteijn, K. Poels, and Y.A.W De Kort, "The Game Experience Questionnaire: Development of a self-report measure to assess player experiences of digital games," Eindhoven: TU Eindhoven. FUGA Deliverable D3.3. Technical Report, 2008.

[23] N.M. Bradley and P. Lang, "Measuring emotion: the self-assessment Manikin and the Semantic Differential," in Journal of Behavior Therapy and Experimental Psychiatry, vol. 1, pp. 49-59, 1994.

Physical Therapy Intervention Through Virtual Reality in Individuals With Balance Disability: a Case Study

Audi Mauro
Barrozo Amanda Lavagnini
Perin Bruna de Oliveira
Dept. Physical Therapy / University of Marilia
UNIMAR
Marília –Brasil
mauroaudi@unimar.br
aamandaalb@hotmail.com
bruna_net_eu@hotmail.com

Braccialli Lígia Maria Presumido
Sankako Andréia Naomi
dept. Special Education / University Estadual Paulista
UNESP
Marília- Brasil
bracci@marilia.unesp.br
asankako@yahoo.com.br

***Abstract*—Balance is an ability found in most of the human activities and it is essential under a functioning view. When balance is impaired due to any brain injury, several tools may be used for its treatment. One of them is virtual reality, a tool which allows individuals to make use of their senses and natural movements during virtual games in order to promote interaction on a virtual environment. This study was performed with a volunteer from the neurology section of Universidade de Marília (Unimar) physical therapy clinic under ethics parameters, and its purpose was to measure the balance of an individual with traumatic brain injury who was involved in a treatment with use of virtual-reality games. An Xbox kinect® console was utilized for this intervention, and the selected games required the use of many physical abilities, including balance. The tests were recorded on video, and pictures were analyzed by two researchers and three evaluators. Score data from each game were obtained during the process. The data analysis was quantitative in relation with game scores and qualitative in relation with the test pictures. The results obtained from the analysis of the balance tests were: (a) in romberg's test with eyes open and eyes closed there were not significant changes; (b) in sensitive romberg's test there was improvement in balance with support in both legs; (c) in dynamic balance test there was improvement in balance during the straight-line walk. In the statistical analysis of game data friedman's variance was found in three levels of significance (p): for each sequence in the seven days of attendance the results were p=0.0367, p=0.0281 e p=0.0136; it was considered as p<0.05 what affected significantly the performance of the volunteer. According to the results observed in the study, it was concluded that virtual reality as a therapeutic media provided improvement of the volunteer's physical balance.**

Keywords-physical therapy; balance; virtual reality.

I. INTRODUCTION

The balance or postural stability is the condition in which every force acting over the body is balanced so that the center of mass stays within the stability borders in the limits of the bases of support [1]. Balance is a sensorimotor function that ensures permanently the dynamic postural stability [2]. In the opinion of postural stability, also defined as balance, is the ability of having control over the center of mass in relation with the base of support [3].

Balance can be static, when a body stays stable in a given position, or dynamic, when it is able to advance through an intentional movement without losing balance [4].

In order to obtain postural control and then postural stability, a body needs to create and apply forces. However, the central nervous system (CNS) requires an exact image of where the body is in the space and if it is still or in movement. The nervous system performs that through the following systems: vestibular, visual, and somatosensory, primarily proprioceptive mechanism and cerebellum which manages constantly every motor activity of the body and compares movements intended by the motor cortex with the updated sensory information that receives and regulates the quality of motor movements [5].

That instability is the lack of ability for correcting the displacement of a body during its movement in space [6]. There are three media in daily life in which balance can be disturbed: by an external force applied to the own body through the movement of the base of support, or by internal forces applied during a self-initiated movement [7]. Any obstruction at controlling the segmental alignment and activating, coordinating and measuring muscular activity effectively can impair postural stability.

Balance is essential for independence in daily activities. Impairments in postural control that produce loss of stability has a profound impact in daily life for individuals with neurological pathology. The consequences of impairment on stability include loss of functional independence, increase of disability prevalence, and falls [3][7].

The ability of keeping a static position provides a lesser challenge for a patient's balance, whereas dynamic activities are more challenging since the center of mass is more displaced. Therefore, the program of rehabilitation for balance training must have exercises with progression of kinetic chain, exercises with eyes open and closed, progression of stable-to-instable bases, self-perturbation by external perturbation, single-to-multiple plans, and single-to-multiple movements [8].

There are also studies that demonstrate the benefit of other therapies in the treatment of balance disorders such as hydrotherapy [9], equine therapy [10], isostretching [11], pilates [12].

Studies on virtual reality (VR) took place in the 1960s, not successfully, then since 1980 they began to be strengthened and became even more popular in the beginning of 1990 with digital technology and the popularization of digital games [13].

The term VR was created to define virtual worlds developed with the utilization of high technology in order to convince users that they are in another reality. It enables the integration between user and virtual environment. So, there is an experience of immersion and interaction based on tridimensional graphic images generated by computer. The VR allows a user to utilize perceptions from our five senses [14].

VR can be defined as a computing simulation that enables to re-create real environments where a subject is able to interact with the game, to experience and simulate a real environment [15].

VR provides a unique media appropriate for the creation of several requirements for an effective intervention of rehabilitation, and can be provided in a functional, intentional and motivating context [16].

Virtual environments are created with the purpose of rehabiliting individuals with disabilities, and it aims to recover the motor ability and cognitive functions. It is a therapy provided for patients who have some brain injury, phobias, autism, traumatic brain injuries, brain paralysis, and in the prevention of falls and accidents with the elderly [17].

It is Schiavinato et al. [18] have reported an important effect in the rehabilitation provided by VR, i.e., a possibility where the patient interacts with a virtual environment, and this interaction provides an immediate feedback by the patient, because he has positive immediate responses of the efficacy of his/her movements. Then his/her brain is stimulated to be adjusted to the game, and makes the corrections necessary so that the patient has a good performance in the game. This enables competitivity, and the patient begins to make good efforts.

The Balance Rehabilitation Unit (BRU) is an equipment which allows handling balance disorder, primarily the vestibular ones. It utilizes 3D glasses that re-create situations that cause dizziness in the user and work on visual stimulation, perception of deepness, direction and movement speed, stimulating the maintenance of body balance [19].

The utilization of VR in rehabilitation, when compared with treatments using other technologies, has several advantagens such as the opportunity of experience of real situations in an illustrative environment that encourages the active and individualized participation of a subject who practice movements being performed later in the real world. It also provides an amusing environment that creates a high level of motivation for the acquisition of knowledge and learning [17].

The VR can be utilized in the treatment of vestibulopathy since the physical movements performed during the therapy reorganize the harmonious functioning between visual, vestibular and somatosensory systems, promoting visual stabilization during the movements of the head. This repetitive training associated with neuronal plasticity makes the recovery of body balance and vestibular compensation possible [20].

The purpose of this study was to measure balance of an individual with traumatic brain injury who underwent a treatment with virtual-reality games.

The present work is subdivided in sections. Section II contains the research methods used. Section III presents how the procedures were realized and which materials were utilized. Section IV presents the results and explains the way they were analyzed. In Section V, we present the discussion, comparing with the results from other authors. And, finally, Section VI presents the conclusion of this work.

II. METHODS

The study was issued and approved by University of Marilia (UNIMAR) Research Ethics Committee, and followed Brazilian and International Regulations and Guidelines, especially Brazilian Health Board Resolutions no. 196/96 and Supplementary, protocol no. 341.

A volunteer from neurology section at UNIMAR Physical Therapy Clinic, male, 30 years, who underwent a conventional physical therapeutic assessment performed by researchers in the last term of Physical Therapy undergraduate course, took part in the study. The volunteer had a clinical diagnosis of traumatic brain injury secondary to a car accident.

The accident occurred in January 31st, 2010, and the volunteer was hospitalized during 48 days, from which 31 days he was in an Intensive Therapy Unit (ITU), and 17 days in an outpatient room. While at ITU he had been into a coma and underwent a procedure with invasive mechanical ventilation. During the period of hospitalization he was put a cervical collar due to a fracture in the third cervical vertebra, underwent a facial drainage, had bronchoaspiration, and had two episodes of pneumonia. The patient had a hospital discharge in March 19th, 2010.

One month after the discharge the patient initiated a therapeutic treatment at UNIMAR Physical Therapy Clinic, where he followed attending up to later April 2011. In this period there was a significant progression in sequels. The patient has any former pathology, has no use of medication, no use of tobacco or alcohol. He has a moderate limitation in daily activities which require a higher level of body balance such as play soccer. The volunteer is a college student and he is temporarily retired from his professional activities.

There were not found force, skeletal or joint changes, muscular atrophy, tonus, profound or superficial sensitivity, or rough coordination at the physical examinations. Changes in static and dynamic balance, and dysdiadochokinesia in the right upper member were the only motor alterations.

The volunteer underwent magnetic resonance imaging examinations through which a small area of gliosis by brain contusion in the right temporal-occipital transition, and sequels from diffuse axonal injury in brain regions (corpus callosum, bilateral brain peduncle, and right lower brain peduncle) were diagnosed in the outcomes from recent examinations, previous to the beginning of the therapy.

A LG 29’’ TV, a Xbox 360° Kinect console and Kinect Adventures game, a Sonic 12.1 megapixels digital camera, a support tripod and simetrograph were utilized in the therapy.

III. PROCEDURES

The procedures of this case study were divided into stages consisting of: First Assessment, Therapeutic Intervention Protocol, and Final Assessment.

The volunteer was assessed initially through Romberg’s test as per O’Sullivan, which is used to evaluate static balance. In this test, the patient was instructed to stand erect and feet positioned together, arms hanging along the body. The patient had three minutes before the beginning of the test for postural adjustment. The test was performed in three stages: with eyes open, with eyes closed, and unipedal right stance, then left stance, hereby called Sensitive Romberg. A simetrograph was used in order to support the test.

A straight-line walk test was used to assess dynamic balance. This test was performed by the patient walking over a 5.7 m straight line in the floor so that the calcaneus of one foot was in front of the ankle of the other foot. He walked the line three times, back and forth; the first time was not considered for analysis.

All tests were recorded in video with a digital camera supported by a tripod. The camera was in a position 2.50 meters of length away from the simetrograph and at 79 cm of height in the Romberg’s test. In the walk test it was at 80 cm of length and 79 cm of height.

In relation to therapeutic intervention protocol, an Xbox Kinect console with Kinect Adventures game and a 29’’ TV were used for the intervention. This console captures movements performed by the user, reads them and then sends kinetic movement signs to the virtual reality; this console does not require joypads or any device to play. Games were not therapeutic and had background music. The distance between the console and the participant were always the same, adjusted by the console sensors. Two games were selected for the therapy: Reflex Ridge, where the virtual scenario was a platform running through trails, and the participant was required to jump, knee and avoid obstacles in the way; and Rally Ball, where he should make movements to catch balls using hands and legs along with side displacements as a goalkeeper. He also should destroy floating boxes with the same balls.

Both games have basic, intermediate and advanced levels. Each level had three different game, i.e. each game had a total of nine mini-games. At each attendance the volunteer played Reflex Ridge first, then Rally Ball. He was requested to play one basic-, one intermediate-, and one advanced-level mini-game from each DVD. Twenty one attendances were performed, with a sequence of three different games in each attendance. After three days the sequence was repeated, in a total of 7 attendances with each sequence. The sequences were: Mover, Dodger and Olympian; Collector, Wrangler and Shapeshifter; and Cruiser, Slingshotter and Speedster.

The game that provided more data for the collection was selected for data analysis: Reflex Ridge. Data were collectedon number of total obstacles, number of fails at trying to overcome obstacles, game time and score. The intervention lasted one month and two days, and four attendances per week were performed in this period, with therapy time of approximately 25 minutes. Initial and final blood pressures were measured in every attendance, since it was a therapeutic attendance. There were not intervals during each therapy, and the volunteer had only short rests due to the time spent to select a new game.

The volunteer underwent the same tests of first assessment one day after the end of attendances, performed in the same site and with the same equipment.

The data analysis was quantitative in relation with game scores and qualitative in relation with the test pictures. The results obtained from the analysis of the balance tests were: (a) in Romberg’s test with eyes open and eyes closed there were not significant changes; (b) in sensitive Romberg’s test there was improvement in balance with support in both legs; (c) in dynamic balance test there was improvement in balance during the straight-line walk.

IV. RESULTS

An imaging qualitative analysis was performed by recording for static and dynamic balance tests, as per table 1. Tests were recorded on video at assessment and reassessment. Then relevant parts were selected and edited for analysis. Both researchers assessed individually the edited parts and took notes considering parameters that aimed the purpose of the study. Following this analysis, they observed common findings and found some conclusions. In order to assure the common outcomes obtained, the video recordings were submitted to three evaluators who were not informed about the study, and the validation of the conclusions obtained by the researchers was accepted if most of the evaluators observed the same outcomes.

TABLE 1. IMAGING ANALYSIS OF BALANCE TESTS THROUGH

CATEGORY	SUBCATEGORY		OUTCOME
STATIC BALANCE	ROMBERG WITH EYES OPEN	ASSESSMENT	“mild deviation”
		REASSESSMENT	“mild deviation”
	ROMBERG WITH EYES CLOSED	ASSESSMENT	“mild deviation”
		REASSESSMENT	“mild deviation”
	SENSITIVE ROMBERG	ASSESSMENT	“left leg: regular balance right leg: poor balance”
		REASSESSMENT	“left leg: fairly normal balance right leg: good balance”
DYNAMIC BALANCE	STRAIGHT-LINE WALK	ASSESSMENT	“regular balance, high deviation”
		REASSESSMENT	“good balance”

In the statistical analysis of game data, Friedman's variance (Fr) [21] was found in three levels of significance (p): for each sequence in the seven days of attendance the results were p=0.0367, p=0.0281 e p=0.0136; it was considered as $p<0.05$ what affected significantly the performance of the volunteer, as per Table 2.

TABLE 2. PATIENT'S SCORE IN EACH SIMULATION GAME IN 7-DAY PERFORMANCE ASSESSMENTAND FRIEDMAN'S VARIANCE (FR)

Game	Level	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Statistical Outcome
Mover	Basic	90	131	109	144	141	134	133	Fr=13,428 p=0,0367
Dodger	Intermediate	124	204	214	256	237	284	280	
Olympian	Advanced	159	196	286	271	238	325	358	
Collector	Basic	218	217	246	243	224	239	259	Fr=14,142 p=0,0281
Wrangler	Intermediate	443	454	489	456	535	538	558	
Shapeshifter	Advanced	455	494	491	531	528	563	564	
Cruiser	Basic	176	202	183	219	217	234	253	Fr=16,035 p=0,0136
Slingshotter	Intermediate	204	187	228	241	259	266	266	
Speedster	Advanced	265	294	307	328	340	350	329	

The results showed a progressive increase in the score in the proposed games every day, noting improvement of the individual's performance.

V. DISCUSSION

The balance disability indicated in volunteer's outcomes is due mostly to a TBI affecting one of the areas responsible for balance control, the cerebellum. It can be observed through the diagnosis of the imaging magnetic resonance, previously addressed in methodology. However, as the cerebellum was not the only structure affected, according to the diagnosis, it was decided to adopt an overall treatment for balance disorders and TBI.

Physical therapy for patients with TBI shall be instructed for a functional performance of concrete tasks in daily life, as well as leisure activities. In this study games were utilized requiring from the player the execution of activities similar to daily life activities, leading him to carry them out in a pleasurable way [7].

One of the main purposes of rehabilitation by using VR is to provide tasks which require from patient repeated movements involving processes of different sensory modes, such as sight, touch, hearing, and proprioception. They provide quantitative and qualitative improvements in daily activities, and increase its functions, improving independent quality of life [16].

The games selected for the volunteer had many repeatable movements, such as jumping, avoiding obstacles, kneeing and making side evasions. At each therapy the volunteer himself observed he could perform a new movement, unable in the previous attendance. In addition, he better performed a given movement in the end of treatment due to repeatability. Repeatability has shown a strong aspect in this kind of therapy.

The adaptation with movements required in the game occurred quickly because the movements were similar to those required in the volunteer's daily life, besides being repeatable. The difference was that, in the game, the movements were reproduced in a virtual scenario, a new situation for the volunteer, and they required a high level of attention, concentration and effort. The adaptation with the console system, which does not require devices, was very good; it was also one of the factors that attracted the volunteer, since it was a different resource for someone who had been performed physical therapy one year ago and was used to traditional equipment.

The exploration of virtual environments by individuals with different disabilities provides new approaches that are impossible to be performed usually. Even the volunteer stated in an interview that he liked the therapy because, besides using the mind, he had to do movements with his own body [17].

The adaptation of games and its accessories by the patient in order to fulfill the purposes given in the routine of a rehabilitation process is a decisive factor for the success of this approach. By submitting the patient to a game as a part of his/her treatment, we assure a continuing involvement between the patient and his/her routine of rehabilitation [22].

The significant involvement of the volunteer with the therapy was also a positive factor for the improvement. Every time he could not perform a given movement or achieve a score, he turned that into an encouragement for the next session. And every time he could improve his results, it was a point of overcome. Many applications utilizing VR provide opportunities for individuals to take part in new and rewarding experiences [16].

The VR can promote the recovery of body balance through a program of repeatable and active physical exercises which involve the eyes, head, body, or physical maneuvers performed by the patient. It aims to stimulate the vestibular system and to enhance CNS neuroplasticity. The virtual reality games chosen for the therapy involved repeatable movements of the eyes, head, and body, improving balance [23].

A study of four cases of spinocerebellar ataxia showed improvement in motor coordination and postural balance with virtual reality use [26].

The study of nine children, aged 4 to 6 years, referred to physical therapy with developmental coordination disorder, participated in 10 game-based intervention sessions, it was concluded that children seemed to be motivated and to enjoy the interaction with the VR environment and VR games seemed to be beneficial in improving the children's motor function[27].

A study performed by Schiavinato et al. [18] that utilized Nintendo Wii as a VR tool with a 24-year female patient with early onset cerebellar ataxia, gait ataxia, and balance deficiency has shown results such as improved balance and higher independence in her daily life activities.

The study showed that the equilibrium, distribution plantar pressure and the teste Time Up and Go (TUG), measured the patient with cerebral palsy spastic diparesis the participant type this case study, were influenced positively the Nintendo Wii game systems with accessory Balance Board [28].

Another study performed by Doná, Santos & Kasse, with an 82-year female diagnosed with right-deficient vestibular syndrome, multisensory changes, sudden deafness, complaints on lack of balance, chronic vertigo, history of falls, who underwent a therapy with BRU has shown results such as improvement in functional body balance, quality of life, and functional ability [19].

 ISBN: 978-1-61208-382-7

The aim of this study was to investigate the role of VR using video game to the improvement of postural control in a chronic stroke survivor and the protocol resulted in higher amplitude of sway in x and y for both eyes open and eyes closed condition, and a higher area of sway in both conditions too [25]. That is consistent with the study performed.

It is Barcala et al. [24] performed a study in which balance was assessed in 12 hemiparetic patients who underwent a balance training with Wii Fit software. They were divided into two groups: a control group had only conventional physical therapy for one hour, and the other group had physical therapy for thirty minutes and thirty minutes of Wii Fit training. Results indicate that either the control group or the Wii Fit group had a better control of static and dynamic balance, and the researchers concluded that the use of devices had significant outcomes in balance rehabilitation for hemiparetic individuals, what may represent one more therapeutic resource.

According to the studies mentioned above, the utilization of VR brought good outcomes to the participants. This was also observed in this study since the volunteer had improvement in his body balance.

VI. CONCLUSIONS

According to the results observed in the study, we concluded that VR as a therapeutic media has made improvements in the volunteer's body balance possible.

The participant showed increased scores in all games proposed, improving their performance.

At our clinical practice, the use of VR has demonstrated itself successful.

Since it is still considered a new therapy in the rehabilitation, several studies can be exploited, in relation to motivational aspects, appropriate therapy time, data about the participant's functional performance, if the VR is most effective when applied alone or in combination with other conventional therapies.

Finally, a broad research field can be explored to the continuity of this work.

REFERENCES

[1] S.B. O'Sullivan, "Evaluation of motor function." In: S. B. O'Sullivan and T. J. Schimitz, "Physical Therapy: Evaluation and Treatment," 4rd ed., vol.1. Barueri: Manole, 2004, pp. 177-212.
[2] P. Perrin, and L. Francis, "Mechanisms of human balance: functional exploration, application to sports and rehabilitation," São Paulo: Andrei, pp. 43-44, 1998.
[3] A. Shumway-Cook, and M.H. Woollacott, "Motor control: theory and practical applications." 3rd ed., vol.1. Barueri: Manole, 2010, pp. 73-116.
[4] E.R. kandel, J.H. Schwartz, and T.M. Jessell, "Principles of Neuroscience," 4rd ed., vol. 2. Barueri: Manole, 2003, pp. 301-328.
[5] F.B. Horak, and C. Shupert, "Role of the vestibular system in postural control," In: S. J. Herdman, "Vestibular Rehabilitation," 2rd ed. Barueri: Manole, 2002. pp.25-51.
[6] R. Caixeta, "Postural instability and falls in the elderly," In: L.H.H. Hargreaves, "Geriatrics," Brasilia: Special Secretariat for Editorações and Publications, pp. 467-486, 2006.
[7] J.H. Carr, and R. Sheperd, "Neurological rehabilitation: optimizing motor performance," 2rd ed., vol.1. Barueri: Manole, 2008, pp. 65-67.
[8] M. Dutton, "Orthopaedic Physiotherapy: examination, evaluation and intervention," 2er ed. Vol.1. Porto Alegre: Artmed, 2010, pp 17-46.
[9] S.M. Resende, C.M. Rassi, and F.P. Viana, "Effects of hydrotherapy in balance and prevent falls in the elderly," Brazilian Journal of Physical Therapy, vol.12, n. 1, pp. 171-172, 2008.
[10] P.M. Tacani, and M. Marques, "Therapeutic Vaulting: balancing performance in individuals with neurological disorders," Brazilian Journal of Health Sciences, n. 14, pp. 68-122, 2007.
[11] V.V. Monte-Dof, P.A. Ferreira, M.S.de Carvalho, J.G. Rodrigues, C.C Martins, and D.H. Iunes, "Effect of isostretching technique in postural balance," Therapy and Research. São Paulo, v.16, n. 2, pp.137-42, 2009.
[12] B.G.S. Rodrigues, S.A. Cader, E.M. Oliveira, O.V.N.B. Torres, and E.H.M. Dantas, "Assessment of static balance in elderly post-training with pilates method," Brazilian Journal of Science and Movement, vol. 17, no. 4, p. 25-33, 2009.
[13] L.C. Meneguette, "Virtual Reality and experience in space: immersion, phenomenology, technology." [Paper presented at the College Pontifical Catholic University of São Paulo to obtain the title of Master of Science] São Paulo: Pontifical Catholic University of São Paulo, pp.54-73, 2010.
[14] E.F.S. Montero, and D.J. Zanchet, "Virtual Reality and Medicine," Surgical Brasileira Acta, vol.18. n. 5, pp.489-490, 2003.
[15] C.P. Perez-Salas, "Virtual Reality: Contribution to Real Un Evaluación la y el Treatment.- Personas con Intellectual Disability," Psychological Therapy, vol.26. n. 2, pp. 253-262, 2008.
[16] H. Sveistrup, "Motor rehabilitation using virtual reality," Journal of Neuroengineering and Rehabilation, n.1 (1), vol. 10. pp. 45-46, 2004.
[17] L. Cardoso, R.M.M. Costa, A. Piovesana, J. Carvalho, H. Ferreira, M. Lopes, A.C. Crispim, L. Penna, K. Araujo, L. Paladin, R. Sancovschi, R. Mouta, and G. Brandão, "Use of virtual environments in the rehabilitation of patients with brain injury from stroke and TBI," [Brazilian Congress of Health Informatics, p. 256, 2006].
[18] A.M. Schiavinato, C. Baldan, L. Melatto, and L.S. Lima, "Influence of Wii Fit balance in patients with cerebellar dysfunction: a case study," Journal of the Health Sciences Institute, Vol. 28, No. 1, pp. 50-52, 2010.
[19] F. Dona, F.B.C. Santos, and C.A. Kasse, "Rehabilitation of body balance by virtual reality in the elderly with chronic peripheral vestibular disease," Brazilian Journal of Medicine, vol. 67, pp. 15-23, 2010.
[20] A. Zanoni, F.F. Ganança, "Virtual reality in vestibular syndromes," Brazilian Journal of Medicine, vol. 67, suppl.1, pp.37-43, 2010.
[21] S. Siegel, and Jr N.J. Castellan, "Nonparametric statistics for the behavioral sciences," 2ª ed. Porto Alegre: Artmed; p. 448, 2006.
[22] R.S. Dias, I.L.A. Sampaio, and L.S. Taddeo, "Physical Therapy Wii X: The introduction of the play in the rehabilitation of patients in physical therapy process,"[Brazilian Symposium on Games and Digital Entertainment, p.153, 2009].
[23] P.R. Santos, A. Meek, C.F. Ganança, A.P.B.A. Pires, N.W. Okai, and T.S. Pichelli, "Vestibular rehabilitation with virtual reality in patients with vestibular dysfunction," ACTA ORL, vol.27, n. 4, p.148-52, 2009.
[24] L. Barcala, F. Colella, M.C. Araujo, A.S.I. Salgado, and C.S. Oliveira, "Analysis of balance in hemiparetic patients after training with the Wii Fit program," physiotherapy in Moviment, Vol 24, No. 2. pp. 337-343, 2011.

[25] S. L. Pavão, N. V. da Costa Sousa, C. M. Oliveira, P. C. G. Castro, M. C. M. dos Santos, "The virtual environment interface as in post-stroke rehabilitation: a case report". Fisioter. mov. [online]. 2013, vol.26, n.2 [cited 2014-01-29], pp.455-462. ISSN 0103-5150. http://dx.doi.org/10.1590/S0103-51502013000200022.
[26] B. S. Zeigelboim, S. D. de Souza, H. Mengelboim, H. A. G. Teive and P. B. N. Liberalesso, "Vestibular rehabilitation with virtual reality in spinocerebellar ataxia,"Audiol. Commun. Res. [online].2013, vol.18, n.2 [cited 2013-10-18], pp.143-147. ISSN 2317-6431. http://dx.doi.org/10.1590/S2317-64312013000200013.
[27] T. Ashkenazi, P. L. Weiss, D. Orian and Y. Laufer, "Low-cost virtual reality intervention program for children with developmental coordination disorder: a pilot feasibility study," Pediatr Phys Ther, vol.25.(4), pp.142-177, United States, 2013.
[28] L. J. P. da Fonseca, M. Brandalize, and D. Brandalize, "Nintendo Wii in rehabilitation of patients with cerebral palsy - a case report," Arq. Ciênc. UNIPAR Health, Umuarama, vol. 16. no. 1. pp. 39-43. January / April 2012.

G-IM: An Input Method of Chinese Characters for Character Amnesia Prevention

Kazushi Nishimoto
Research center for innovative lifestyle design
Japan Advanced Institute of Science and Technology
Ishikawa, Japan
knishi@jaist.ac.jp

Jianning Wei
The school of knowledge science
Japan Advanced Institute of Science and Technology
Ishikawa, Japan
wei_smiling@ezweb.ne.jp

***Abstract*— Character amnesia is a recent phenomenon in which native Chinese or Japanese speakers forget how to write Chinese characters (Kanji in Japanese), but maintain the ability to read them. It is generally believed that the constant use of computers and mobile phones equipped with pronunciation-based Chinese-character input systems is to blame. Therefore, particularly in China, several element-based input methods that require users to input radicals of the Chinese characters have been developed. However, these methods are not effective for learning how to write unfamiliar characters. This paper proposes a novel pronunciation-based input method called G-IM. Unlike conventional methods, G-IM sometimes outputs incorrect character shapes, which forces users to pay close attention to the character shapes and thus strengthens retention and recall. Through user studies, we confirmed that G-IM significantly strengthens the retention and recall of character shapes as compared to conventional input methods and writing by hand.**

Keywords-input method of Chinese characters; character amnesia; incorrect character shapes; pronunciation-based input method; (re)building retention and recall of Chinese characters.

I. INTRODUCTION

Character amnesia is a recent phenomenon manifested by the inability of native Chinese or Japanese speakers to recall how to write Chinese characters (Kanji in Japanese), although they know these characters well and can still read them [1]. According to a poll commissioned by China Youth On Line, published on July 12, 2013, 94.1% of the respondents reported experiencing problems writing characters and 26.8% reported always have such difficulties [2]. A similar phenomenon has been noted in Japan and is considered a significant issue [3].

It is generally believed that the constant use of computers and mobile phones equipped with alphabet-based Chinese-character input systems is the cause of character amnesia [3][4]. In China, "Pinyin typewriting" is the most popular method for inputting Chinese characters. Pinyin is a widely used representation method of Chinese pronunciations that allows Chinese-character pronunciations to be expressed with Latin characters. In Japan, "Kana-kanji conversion systems" and "Romaji-kanji conversion systems" are widely used, where "Kana" means Japanese phonetic symbols, "Romaji" refers to Romanized Japanese words, and "kanji" is the Chinese characters. Thus, both in China and in Japan, people usually input Chinese characters with personal computers (PCs) and mobile phones based on their phonetic features, not on their shape features. Therefore, it is not necessary to recall how to physically write these characters when inputting. As a result, they increasingly lose the ability to write Chinese characters.

It has been argued that it is no longer necessary to be able to hand write Chinese characters. Custer [5] asserted that some characters are used frequently enough that it's nearly impossible to forget them, and that only when people have occasion to write a less common word by hand, they will look it up on their phones. Custer [5] therefore concluded that character amnesia is actually not a serious concern.

We, however, cannot agree. We conducted a dictation examination of frequently used Chinese characters, the details of which are described below in the User Study section. The examinees (research participants) were 30 Chinese postgraduate students. Surprisingly, the average accuracy rate for these students was only 22.8%, with the best score reaching only 56.3% and the worst actually 0%. Thus, even this highly educated group forgot approximately 80% of frequently used Chinese characters, which is against Custer's assumption. People easily forget even frequently used characters. Character amnesia thus appears to be a compelling problem.

This paper proposes a novel Chinese-character input method called G-IM, an acronym of "Gestalt Imprinting Method," to address character amnesia. The remainder of this paper is organized as follows: We present an overview of several related works in Section II. We then describe the basic idea of G-IM in Section III and illustrate its system setup in Section IV. Finally, we present the results of our user studies in Section V and discuss the usefulness of G-IM in Section VI. Section VII concludes this paper.

II. RELATED WORKS

A straightforward and naive way to prevent character amnesia is to continue Chinese-character education. Both in China and Japan, Chinese-character education is provided from early elementary school, with pupils in China learning approximately 3000 characters and those in Japan approximately 2000 characters by the end of junior high school. Several online learning systems have been proposed [6][7][8]. However, results of a questionnaire we conducted for 135 Chinese people using an on-line questionnaire site revealed that only 2.96% respondents (namely, only 4 respondents) were interested in using such online learning systems. Thus, as mentioned in Section I, there are many adults in both China and Japan who cannot write Chinese characters, demonstrating that education alone is not enough to prevent character amnesia.

Various new approaches have recently been attempted. A prominent approach is the use of edutainment and gamification. A very popular new Chinese television program, “汉字英雄” (translation: “A Hero of Chinese Characters”), is a game show where elementary and junior high school students are asked to correctly transcribe the dictation of Chinese characters. A smart-phone application that provides the same dictation game is also gaining in popularity and has been downloaded over 800 thousand times. In Japan, kanji reading and writing quizzes are very popular on many television programs. These entertainment programs and game applications provide some remedy for character amnesia. However, at least in Japan, the quizzes provided in such TV programs often use very difficult characters that are seldom seen in everyday life. A method inextricably related to everyday writing (or inputting) of Chinese characters is required.

In this sense, developing methods for Chinese characters that provide functions not only for inputting, but also for learning, is a promising solution. In Japan, the pronunciation-based Chinese character input systems, i.e., the Kana-kanji and Romaji-kanji conversion systems, are almost the only current practical methods for inputting Chinese characters on most PCs, cell-phones, and smart phones. In contrast, in China, not only pronunciation-based methods, such as the Pinyin input method, but also element-based methods are used.

“Wubi” is an element-based method. With Wubi, users input a Chinese character by inputting its radicals [8]. For example, to input “柞” (“frame”), the user must input the following three radicals: “木,” “九,” and “十.” A combination of radicals almost uniquely relates to a specific Chinese character. Therefore, if the user masters Wubi, he/she can input Chinese characters faster than with the Pinyin method. However, it is difficult for people to master the Wubi method because there are 130 radicals; 5 times that of the alphabet. As a result, Wubi is not as widespread as the Pinyin method.

Another element-based Chinese-character input method is a stroke-based method that breaks down a Chinese character into strokes. There are five basic types of strokes:

1. Horizontal strokes: “一,”
2. Vertical strokes: “｜,”
3. Left-falling strokes: “丿,”
4. Right-falling strokes or dot strokes: “丶,” and
5. Turning strokes: “フ.”

These strokes are usually assigned to numeric keys one to five and a user sequentially inputs strokes that constitute a Chinese character in its handwriting order. For example, “康” (“healthy”) can be input by sequentially inputting the basic strokes as follows:

康：丶一丿フ一一｜丶一丿丶

In this case, eleven strokes are required, which is quite cumbersome. The six-digit stroke-based Chinese-character input method [9] is a refined stroke-based method. With this method, only the first three strokes and the last three strokes are required to be input. Therefore, in the case of “康,” only the following six strokes are necessary:

康：丶一丿一丿丶

Thus, the total number of strokes is reduced even when inputting complicated characters.

Various element-based and stroke-based methods have been proposed that aim at achieving flexibility and simplicity [10], investigate mapping of elemental strokes onto the keypad of mobile phones [11], and design a special keypad layout of a mobile phone’s touch screen for selecting radicals to achieve fast and easy input of Chinese characters [12].

When a person becomes adept at using such an element-based method, he/she can overcome character amnesia. However, most element-based methods require users to be familiar with each character’s shape, including handwriting order, in advance. Without this knowledge, users cannot use these methods. These methods provide no cure for character amnesia, as character amnesia is an obstacle to using these methods.

As a result of considering these conventional attempts, we concluded that pronunciation-based Chinese-character input methods require additional functions in order to solve the character amnesia problem. The pronunciation-based methods are most widely used both in China and Japan, and can be used without accurate knowledge of how to write Chinese characters in detail; even those suffering from character amnesia can use them. To the best of our knowledge, there have been no attempts to solve the character amnesia problem using a pronunciation-based Chinese-character input method.

III. BASIC IDEA OF G-IM

The basic idea of G-IM is very simple and straightforward: a function that compels users to verify Chinese-character shapes is added to a pronunciation-based Chinese-character input method. G-IM not only outputs correctly shaped characters, but also sometimes outputs incorrectly shaped ones. In other words, G-IM is an input method that sometimes miswrites, which sets it apart from conventional input methods. Figure 1 shows an example of

Figure 1. An example of a correctly shaped (left) and incorrectly shaped “歳” character (right). An extra horizontal stroke has been added to the example on the right.

correctly and incorrectly shaped characters: the one on the right is incorrect (it has an extra horizontal stroke). The incorrect characters output by G-IM are slightly different from existing correct ones.

When using a conventional pronunciation-based input method of Chinese characters, some homonyms are often incorrectly input. For example, the pronunciation of "歳入" ("annual income") is "sai-nyuu" in Japanese, which may be incorrectly transcribed as "再入" ("re-entrant"), whose pronunciation is the same as that of "歳入." However, when correcting "再入" to "歳入," the knowledge required concerns the combination of the characters (i.e., knowledge of idiom) rather than the shapes of the characters (i.e., the Gestalt of the characters). If someone noticed that he/she incorrectly input "再入," he/she would re-input "sai-nyuu" and correctly convert it to "歳入." In this correction process, he/she pays attention to the difference between "再" and "歳," but not to the detailed shape of "歳." Therefore, we employed incorrect characters that have shapes that differ slightly from the correct one in order to compel accurate attention to the detailed character shapes and to re-establish the correct Gestalt.

Since the conventional input methods never miswrite and always output correctly shaped characters, users exceedingly rely on these and never pay full attention to the detailed shapes. Eventually, these shapes are forgotten. However, because G-IM is not consistently accurate, the user has to always pay attention to the shape of each character and, if it is wrong, correct it, thus preventing and curing character amnesia.

IV. EXPERIMENTAL SYSTEM

It is not easy to modify the functions of existing Chinese-character input methods such as Microsoft Office IME™ [1] and JustSystems ATOK™ [2]. We therefore implemented a text editor, described later in this paper, instead of an input method to investigate the efficacy of the basic G-IM idea described in the previous section.

We first created a new font file consisting only of incorrectly shaped characters, like the one on the right in Figure 1, using the TTEdit™ font editor[3]. It is not necessary to create incorrect shapes for all Chinese characters. We created only the fonts necessary (shown in Figure 2) for use in our experiments. We then implemented a text editor equipped with a function that automatically replaces an input correct character shape from the existing input method with a similar, but incorrectly shaped, character with the same character code as the correct one as soon as one of the characters in Figure 2 are input. If a user tries to save the text file with incorrectly shaped characters, the save operation is rejected and a dialog box asking the user to correct all incorrectly shaped characters appears. Correction can be done by selecting the incorrect character, inputting its pronunciation again, and converting it. As a result, the incorrect character is replaced with the correct one. At this time, the quick change of the characters between correct and incorrect provides a visual animation-type effect and the difference in the shapes is emphatically demonstrated. Finally, after all characters have been corrected, the text file can be saved. By using this editor and font file, we can conduct user studies without implementing a completely new input method or modifying an existing one.

肇 戛 蹶 锲 颐 蘸 缭 赝
舆 耀 悍 葵 霓 鸩 慰 鹜
腻 鼎 粱 篡 葺 炙 瘙 祟
斡 嗦 墨 覆 孪 罄 罹 靡
纨 绔 戮 寥 迥 箴 率 媲
粕 瞰 凿 噱 霾 敷 瓮 帷
幄 醍 醐 寡 揠 捺

Figure 2. 54 Chinese characters prepared for the dictation pre-examination.

V. USER STUDY

We conducted user studies to investigate the efficacy of G-IM by comparing it to two Chinese-character input methods, specifically, handwriting (HW) and an existing input method (IM). We did not ask the research participants to use an element-based input method such as Wubi. Most of the participants had never used an element-based input method and hence would not be able to use it immediately. Furthermore, as will be shown later, most of the participants could not correctly write many of the Chinese characters that were used in the experimental examinations. When using the element-based input methods, the examinees could not input most of the test characters. Therefore, we did not use an element-based input method during our studies.

A. Procedure

The participants were 30 postgraduate Chinese students from the authors' institute. Character amnesia is a more significant problem for the Chinese than the Japanese because the Japanese can easily substitute Kana (Japanese phonetic symbols) for kanji. The Chinese do not have such a generally tolerated substitution method.

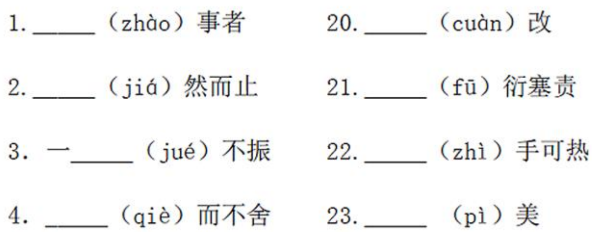
The user study consisted of three steps:

1. A pre-examination of Chinese-character dictation,
2. An exercise of inputting or writing Chinese characters using one of the input methods, and

---

[1] http://www.microsoft.com/ja-jp/office/2010/ ime/default. aspx
[2] http://www.atok.com/
[3] http://opentype.jp/ttedit.htm

 ISBN: 978-1-61208-382-7



1. _____ （zhào）事者
20. _____ （cuàn）改
2. _____ （jiá）然而止
21. _____ （fū）衍塞责
3. 一_____ （jué）不振
22. _____ （zhì）手可热
4. _____ （qiè）而不舍
23. _____ （pì）美

Figure 3. Examples of the problems in the dictation pre-examination.

3. A post-examination of Chinese-character dictation.

In Step 1, we prepared 54 characters selected from “100 frequently used Chinese characters that are often miswritten” [13] as the problems for the dictation pre-examination. Figure 2 shows the selected 54 characters and Figure 3 shows some of the problems in the pre-examination. The participants were required to fill in the blanks with the correct characters by hand referring to the adscript pronunciations. After the pre-examination, all the participants were not given the correct answers.

Based on the pre-examination results, we sorted the 54 characters by the miswritten-ratios and extracted 32 characters with higher miswritten-ratios. Figure 4 shows the extracted 32 characters and their miswritten-ratios. No participants were able to correctly write the top two characters. These 32 characters were used in the Step 2 tasks and in the post-examination of dictation in Step 3. In addition, based on the pre-examination results, we divided the participants into three groups, each of which included ten participants, so as to equalize the distribution of the pre-examination score of the three groups.

In Step 2, we assigned three different tasks to the three groups as follows:

- G-IM group: The participants were required to input sentences that included the 32 characters. Whenever one of the 32 characters was input, it was replaced by the corresponding incorrectly shaped character first and the participants then made the necessary corrections.
- IM group: These participants were required to input the same sentences as those provided to the G-IM group using a popular Pinyin input method. As described previously, with this method, the need for replacement of incorrectly shaped characters never arose.
- HW group: The participants of this group were required to write by hand the same sentences as those provided to the G-IM group.

For all groups, the sentences were read aloud. The participants listened then input the sentences or wrote them down. To avoid any unexpected effects of Step 1 from influencing the tasks of Step 2, we waited 15 days between Step 1 and Step 2.

In Step 3, occurring immediately after Step 2, we conducted the post-examination using the 32 characters shown in Figure 4. The form of the post-examination was similar to the pre-examination (Fig. 3): the examinees were required to fill in the blanks with correct characters by hand. Finally, we asked the participants of the G-IM and IM groups to answer a questionnaire on whether they had paid attention to the character shapes when inputting the sentences in Step 2.

Characters	Miswritten-ratio
鹜	100.0
霾	100.0
颐	97.0
赝	90.9
篡	90.9
斡	90.9
蘸	87.9
祟	87.9
霓	84.8
敷	84.8
腻	84.8
舆	84.8
噱	84.8
锲	81.8
悍	81.8
靡	81.8
蹶	81.8
粱	81.8
慰	78.8
媲	78.8
寡	75.8
罹	75.8
肇	72.7
迥	72.7
戛	69.7
箴	69.7
炙	69.7
鼎	69.7
葵	63.6
耀	60.6
凿	57.6
覆	57.6

Figure 4. 32 extracted characters with higher miswritten-ratios based on the pre-examination results.

TABLE I. RESULTS OF THE PRE-EXAMINATION AND POST-EXAMINATION FOR THE THREE GROUPS.

Group	Pre-exam.		Post-exam.	
	Average	STDV	Average	STDV
G-IM	7.3	5.0	20.4	6.2
IM	8.4	4.4	12.0	5.7
HW	7.7	4.2	10.3	6.3

TABLE II. AVERAGE REQUIRED TIME IN MINUTES TO INPUT SENTENCES IN STEP 2 AND THE POST-EXAMINATION FOR ALL GROUPS.

Group	Step 2 (min.)	Post-exam (min.)
G-IM	25	8
IM	13	11
HW	33	7

TABLE III. PERCENTAGE NUMBER OF PARTICIPANTS WHO PAID ATTENTION TO THE CHARACTER SHAPES WHEN INPUTTING SENTENCES IN STEP 2.

Group	When selecting character	After selecting character
G-IM	40 %	100 %
IM	67 %	17 %

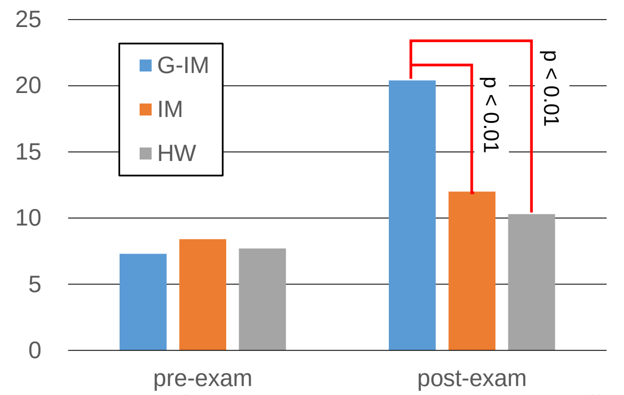


Figure 5. Results of post hoc test on the interaction: simple main effect of input methods at each examination.

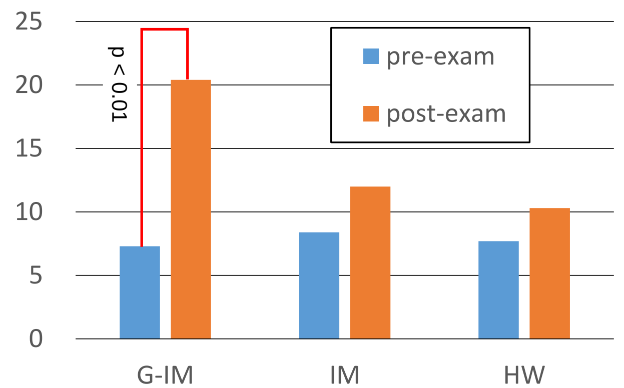


Figure 6. Results of post hoc test on the interaction: simple main effect of examinations at each input method.

B. Results

Table I presents the average scores and standard deviations (STDVs) for the pre-examination and post-examination of the three groups with a perfect score being 32. Therefore, the average score 7.3 is, for example, 22.8 points on a scale for which 100 is perfection. Table II shows the average time required to input the sentences (Step 2) and the average time for writing the answers down by hand in the post-examination for all groups. Table III displays the percentage number of participants who paid attention to the character shapes when inputting the sentences in Step 2.

To determine whether G-IM could improve the performance of the post-examination when compared to the other two methods, we performed a two-way ANOVA (three input methods and two examinations) on the results shown in Table I. The analysis results were:

- The main effect of the input methods was significant. ($F(2, 54) = 3.99$, $p < .05$)
- The main effect of the examinations was significant. ($F(1, 54) = 19.41$, $p < .01$)
- The interaction effect was significant. ($F(2, 54) = 5.25$, $p < .05$)

We performed additional post hoc tests. We first conducted Tukey's HSD (honestly significant difference) test on both main effects. The results were as follows:

- About the main effect of the input methods: A significant difference ($p < .05$) was found only between G-IM and HW.
- About the main effect of the examinations: A significant difference ($p < .01$) was found between the post-examination and the pre-examination.

Next, we performed a post hoc test on the interaction effect. Figure 5 shows the simple main effect of the input methods at each examination and Figure 6 shows the simple main effect of the examinations at each input method. Based on Figure 5, we can say that

- In the pre-examination, no significant difference was found among all three groups, but
- In the post-examination, significant differences ($p < .01$) could be found only between G-IM and the other two methods.

From Figure 6, we can also say that

- As for the G-IM group, a significant difference ($p < .01$) was found between the pre-examination and the post-examination, but

- As for the other two groups, no significant differences could be found between the pre-examination and the post-examination.

From these results of the post hoc test on the interaction effect shown in Figures 5 and 6, we can clearly conclude that G-IM is superior to both the existing Pinyin input method and writing by hand. Although the performances of the participants of all three groups were equalized based on the pre-examination results, they differed significantly in the post-examination and the G-IM group achieved significantly higher performances than the other groups (Fig. 5). The G-IM group members' performances were significantly improved after inputting sentences using the G-IM method (Fig. 6).

VI. DISCUSSION

In order to confirm whether G-IM can actually prevent and cure character amnesia, further long-term investigation is necessary. However, it became evident that the recall and retention of Chinese-character shapes was certainly strengthened after using G-IM, which fortifies long-term memory of them. Accordingly, G-IM has the potential to prevent and cure character amnesia.

Those who are resistant to information technology often argue that we should not use the PC for writing and that writing by hand will solve the character amnesia problem. However, our experimental results dispute this claim. Even when writing characters by hand, the post-examination scores did not significantly improve (see the results of HW in Fig. 6). This means that writing by hand did not strengthen character shape recall. This is not surprising. When writing by hand, incorrect memories are not rectified. If someone thinks that the correct shape of "歳" is the one shown on the right in Figure 1 and always writes it this way, there is no chance to correct it. In addition, if he/she does not know how to write "歳" at all, he/she of course cannot write it and will not remember how to write it correctly. In such cases, those who are anti-information and technology believe that we should always refer to a dictionary. However, this is too cumbersome, time consuming, and impractical. It is also impractical that an expert of Chinese characters always checks the shapes of the written characters and corrects the mistakes. As a result, most people do not have any opportunity to develop correct memory of the character shapes. It is necessary to "push" the correct answers, not "pull" answers, such as referring to a dictionary.

The existing pronunciation-based input methods "push" the correct shapes. Opportunities to build recall and retention are provided whenever these methods are used. In this sense, pronunciation-based input methods are potentially learning support systems. However, we do not actually learn the correct shapes when using the existing input methods, which is most likely because the ordinary input methods never miswrite. We do not doubt the characters that are output by these systems and thus completely rely on them. The results shown in Table III support this view. When selecting a correct character, many of the IM participants paid attention to the character shapes. In contrast, most did not pay attention to them after selecting them. This suggests that they pay attention to the shapes when selecting, for example, "歳," and not "再," but they do not pay attention to the details of the shape of "歳." Therefore, because the correct answers are always provided, recall and retention are not reinforced.

In summary, although writing by hand forces attention on the detailed shapes of Chinese characters, it cannot correct mistaken memories. It is effective for maintaining existing correct knowledge, but is not effective for correcting incorrect knowledge or for acquiring familiarity with unknown character shapes. In contrast, although the existing input methods have the potential to correct mistaken memories, they do not force users to pay attention to shape details. G-IM provides a balance; it forces users to pay attention to detailed character shapes, as well as "pushes" correct shapes to correct flawed memories. Therefore, G-IM can be an on-the-job training system for writing Chinese characters. It is effective for acquiring the shapes of unknown characters, as well as maintaining and correcting already-learned ones.

The only disadvantage of G-IM is the extra burden of correcting the incorrect characters, which does not occur when using the existing input methods. As Table II shows, the G-IM group spent more time inputting the sentences in Step 2 than the IM group. This is an unavoidable trade-off. When we explained the G-IM method to other people at, for example, an academic meeting, many responded that they did not want to use such a cumbersome system. However, we believe that this minor inconvenience is not fatal considering the benefits and pay-off, just as foregoing taking the car and walking to your destination in order to get some fresh air and exercise, when it is not too far away and you are not in too much of a hurry, while less convenient, is preferable and beneficial to your health.

Similarly, we should approach the use of G-IM as opposed to the ordinary input methods on a case-by-case basis. If time is an issue, a conventional input method should be used. However, when convenient, G-IM should be used to correct and (re)establish knowledge, recall, and retention of Chinese characters. G-IM does not require as much time as studying educational materials. In addition, with G-IM, people can efficiently learn the correct shapes of common and frequently used Chinese characters, while educational materials require the additional study of characters that are seldom used.

This design policy, which does not just aim at improving the performance of inputting Chinese characters, has something in common with the Further Benefit of a Kind of Inconvenience (FUBEN-EKI) concept proposed by Kawakami et al. [14][15]. They argued that convenient tools hamper skill acquisition and motivation. They propose that designing tools that incorporate or retain some inconveniences is important. G-IM incorporated the input methods of Chinese characters with some inconveniences: users are required to maintain constant attention to the shapes of the output characters, which provides a different benefit from inputting efficiency, that is, the correction and (re)establishment of Chinese character memories. Kawakami et al. did not demonstrate any concrete method of incorporating such inconveniences. G-IM embeds quizzes in the operations of inputting Chinese characters. Thus, embedding quizzes (or, more generally speaking, embedding some gamifying functions) in everyday activities can be a concrete design methodology based on the FUBEN-EKI concept.

Another gamification approach may alleviate the cumbersomeness of G-IM. Scoring correctly saved text files and sharing scores among G-IM users can lead to enjoyable competition among users. We would like to implement such additional functions to make G-IM much more useful and enjoyable in the near future.

VII. CONCLUSION

This paper described a novel pronunciation-based Chinese-character input method called G-IM. The primary feature of G-IM is that it sometimes outputs incorrectly shaped Chinese characters, which the conventional input methods never do. G-IM forces users to pay attention to detailed character shapes in order to build, correct, and strengthen the memory of Chinese characters. We conducted user studies that compared G-IM with a conventional pronunciation-based input method and writing by hand. As a result, we confirmed that G-IM significantly strengthened the participants' memories of Chinese characters compared to the other two methods. Accordingly, we can conclude that G-IM has the potential to solve character amnesia, which is a current and increasing problem in both China and Japan.

In the near future, we would like to incorporate some gamification functions to alleviate the cumbersomeness of G-IM. We also would like to conduct more long-term user studies to estimate G-IM's efficacy more accurately. Furthermore, we need to conduct experiments in the wild for evaluating actual usefulness of G-IM and for confirming compatibility of G-IM as an inputting method of Chinese characters and as a learning tool of Chinese characters. We are now developing a new version of G-IM that can be applied to everyday usage to conduct such experiments.

ACKNOWLEDGMENT

The authors sincerely thank all of the research participants who willingly cooperated to our experiments. This work was supported by JSPS KAKENHI Grant Number 26280126.

REFERENCES

[1] Character amnesia, Wikipedia, http://en.wikipedia.org/wiki/Character_amnesia [retrieved: Dec. 2014]

[2] http://article.cyol.com/news/content/2013-07/12/content_8708413.htm (in Chinese) [retrieved: Dec. 2014]

[3] H. Kaiho, "Why does memory slips of Chinese characters occur?," in Psychology of Cognition and Learning, Baifu-kan, 2007. (in Japanese)

[4] Wired youth forget how to write in China and Japan, http://www.independent.co.uk/life-style/gadgets-and-tech/wired-youth-forget-how-to-write-in-china-and-japan-2065228.html [retrieved: Dec. 2014]

[5] C. Custer, "Is "Character Amnesia" a Problem?", http://chinageeks.org/2010/07/is-character-amnesia-a-problem/ [retrieved: Dec. 2014]

[6] N. Inami, H. Tominaga, Y. Matsubara, and T. Yamasaki, "Network based training system for handwriting Japanese characters using interactive electric whiteboard," Technical Report of IEICE, ET2002-108, 2003, pp.79-84. (in Japanese)

[7] J. Tynystanova, and J. Miwa, "An Associative Kanji Learning Support System Using Animations with Story for Digital Natives," IEICE Technical Report, ET2009-15, 2009, pp. 7-12. (in Japanese)

[8] Q. Li, and W. Wu, "Remote Education Software for "Wubi" Typewriting: --for elective course of elementary school," Master thesis, KTH, School of Information and Communication Technology (ICT), 2012.

[9] L. M. Po, and C. K. Wong, "Six-Digit Stroke-based Chinese Input Method," Proc. IEEE International Conference on Systems, Man and Cybernetics, 2009, pp.818-823.

[10] S. C. Hsu, "A Flexible Chinese Character Input Scheme," Proceedings of the 4th annual ACM symposium on User interface software and technology (UIST '91), 1991, pp.195-200.

[11] M. Lin, and A. Sears, "Graphics matter: a case study of mobile phone keypad design for chinese input," CHI '05 Extended Abstracts on Human Factors in Computing Systems, 2005, pp. 1593-1596.

[12] J. Niu, L. Zhu, Q. Yan, Y. Liu, and K. Wang, "Stroke++: A hybrid Chinese input method for touch screen mobile phones," Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI '10), 2010, pp.381-382.

[13] http://zhidao.baidu.com/question/192981329.html [retrieved: Dec. 2014]

[14] H. Kawakami, "Inconvenience utilizing Design," Kagaku-Dojin Publishing Company INC, 2011. (in Japanese)

[15] H. Kawakami, and T. Hiraoka, "Incorporation of evolutionary computation for implementing the benefit of inconvenience," International Journal of Advancements in Computing Technology, Vol.4, No.22, 2012, pp.248-256.

HCI Education: Innovation, Creativity and Design Thinking

Alma L. Culén
University of Oslo
Oslo, Norway
E-mail: almira@ifi.uio.no

***Abstract*—Human-Computer Interaction (HCI) education needs re-thinking. In this paper, we explore how and what creativity and design thinking could contribute with, if included as a part of the HCI curriculum. The findings from courses where design thinking was included, indicate that design thinking contributed to increased focus on innovation and creativity, as well as prevented too early fixation on a single solution in the initial phases of HCI design processes, fostering increased flexibility and adaptability in learning processes. The creativity and adaptability may be the best long-term foci that HCI education can add to its curriculums and offer to students when preparing them for future work practices.**

Keywords-HCI; innovation; creativity; design thinking; education.

I. INTRODUCTION

It has been argued in favor of purposefully managed innovation through design and creativity in many different ways [1]. Design thinking is one of those options. Understanding design thinking is not straightforward. In [2, p. 13], three different ways (of understanding design thinking) are offered: as a cognitive style, a general theory of design, or as an organizational resource. The latter understanding lends itself well as an approach to innovation and real-life problem solving through human-centered design, employing empathy with users, rapid prototyping and abductive thinking as its main components. This understanding of design thinking has strongly impacted the innovation in business, education, health and other crucial domains [3]–[7]. Many examples of how businesses and organizations could benefit from incorporating design thinking into business and organizational processes were given [8], making design thinking into an efficient innovation engine emphasizing observation, collaboration, fast learning, visualization of ideas, rapid concept prototyping, synthesis and concurrent business analysis.

However, no approach solves all problems. Thus, just some years after the design thinking made a breakthrough in the world of business strategy and management, its limitations were brought forth in works such as [9], [10]. The point made by Nussbaum in [11], though, hits home best: *"From the beginning, the process of Design Thinking was a scaffolding for the real deliverable: creativity. However, in order to appeal to the business culture of process, it was denuded of the mess, the conflict, failure, emotions, and looping circularity that are part and parcel of the creative process. In a few companies, CEOs and managers accepted that mess along with the process and real innovation took place."* In short, the core of innovation is creativity, a messy and unstructured process. By framing design thinking in a particular way, the creativity became limited, leading, in turn, towards failure to innovate.

Concerns are voiced around the failure of design processes currently applied within the field of HCI to support more radical innovation [12]. In particular, HCI design processes are held to lead mainly to incremental innovation and small changes. Innovation, of any kind, is a much more complex process than design and invention of new products, systems, or interaction modes. It also implies their acceptance and use by people [13]. Upon careful consideration of design practices within HCI, one could argue that the invention is common. However, a very small percentage of those inventions (prototypes) ever become finished products and even smaller percentage gets to be adopted and used, see [14].

Preparing today's students of Human-Computer Interaction (HCI) for tomorrow's work practices is challenging. The technologies, interaction modes and interfaces all change fast. In addition, there are rises and falls of techniques in use, design processes, work practices, software and platforms in use.

Students, on one hand, need to learn appropriate theories and research methods, understand the state-of-the-art research, importance of scientific rigor and relevance. However, being a profoundly inter-disciplinary field, HCI does not offer any unifying core theories, so this goal is hard to achieve once and for all (in other words, new application domains require acquisition of new theoretical knowledge, what the state-of-the-art research is, etc.).

On the other hand, students need to be able to design new technologies and interfaces, using design processes and methods. This is also hard to achieve without any formal training in design, which is, in part, why design processes in HCI often depend heavily on engaging users and other stakeholders, thus sharing the responsibility with them for success or failure of a designed prototype. The latter is not seen as problematic, as prototypes are often not intended to become artifacts, but are tied to the research objectives.

Combining insights from our previous work [12], [14]–[16], this paper argues that teaching about innovation, and engaging students in creative innovation processes such as the design thinking (with acceptance of the messy parcel of creative processes [11]), offers one possible answer to what kind of knowledge and skills the students could be taught in

HCI. Adoption of this approach may be successful in a long run because, while on the road to becoming an innovator within a design team, one usually experiences creativity (one's own or that of others) and a need to be adaptable to series of new situations. Creativity and adaptability may offer a greater permanent value to human-computer interaction students than many other kinds of knowledge and skills commonly considered to be part of the HCI education. As reported in [15], all ten students in a graduate HCI course that made use of design thinking processes, perceived themselves as non-creative individuals at the beginning of the course. At the end of the course all, except for one student who felt neutral, stated that the design thinking affected them and that they see themselves as more creative and confident in their skills. A new survey was conducted at the end of a combined bachelor-master course in the fall of 2014. All design teams who participated in the class filled the survey (18 teams consisting of 3-4 students each). They all said that they thought that HCI design is a creative process, and provided qualitative statements related to their experience of individual and group creativity. Some of these are presented later, in the discussion section of this paper.

In summary, the question this paper tries to answer is: what kind of knowledge and skills should be passed onto new generations of HCI designers and researchers? While the whole solution remains elusive (many discussions around what HCI curriculum are already going on [17], [18]), our experience from the past two years of including design thinking and innovation in the curriculum shows that these benefit HCI students significantly.

The paper is structured as follows: the next section offers some thoughts as to why HCI education should include innovation and creative thinking. In Section III, the concrete case is presented of how these elements are introduced within a mixed bachelor-master HCI course. Discussion of the case is presented in Section IV, followed by the conclusion in Section V.

II. FOCUSING ON INNOVATION

The ACM SIGCHI Curricula for Human-Computer Interaction defines Human-computer interaction as "a discipline concerned with the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" [19]. Teaching HCI typically includes teaching of user-centered requirement analysis, design and prototyping, implementation, design of experiments and evaluation. HCI's interdisciplinarity brings in tensions between the breadth and the depth of teaching, diverse theories and practices, including the basic choice between contributing to science, or to design (of new interfaces, products, services or interaction modes). Despite tensions, HCI education is very much alive and doing well in practice, although, still without generally agreed upon curricula.

Innovation, on the other hand, is known to be hard to achieve in practice, while it is very easy to understand the need for it and the benefits it brings [20]. There are various ways to define innovation. Oslo Manual [13] defines it as: *"the implementation of a new or significantly improved product (good, or service) or process, a new marketing method, or a new organizational method in business practices, workplace organization, or external relations."*

It is difficult to teach students to be innovative, creative and inventive. It is not easy to make good frameworks for doing so. The processes related to innovation rely a lot on creativity, but also on both existing knowledge and on technical skills that are already present among the members of the design team and those whom they chose to include in the design processes. In particular, it is hard to define learning outcomes for such processes.

Within HCI, the creativity bit is usually, at least partially, bypassed by two things: framing of the process as a procedure that all can follow on one hand, and relying on understanding of users and their needs on the other hand. It is, thus, usual to develop understanding of the knowledge domain first, and then this understanding is put to test through practical work involving a prototype design, evaluation, and the re-design cycle. However, creative problem solving, a core activity of innovative design [3], [4] is, as mentioned, harder to frame.

Design thinking is but one facet of design. It employs, in part, similar steps to those often proposed in HCI: it frames its process in ways that have familiar overtones to those used in HCI, see Fig. 1.

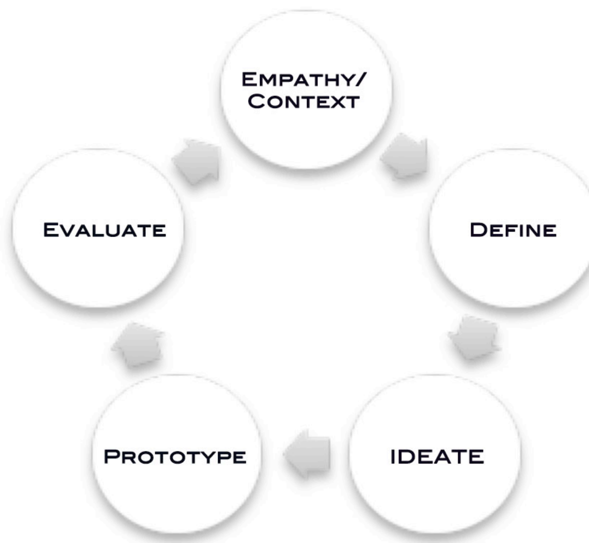


Figure 1. A process that may seem familiar to HCI students, as well as to those using design thinking.

Arguably, differences between design thinking and HCI must be sought by other means than comparing high-level design processes. One needs to consider differences in assumptions, scope and aim of the design process – concerning, for instance, the role of research, requirements specification, questioning assumptions, the consideration of organizational issues, the systematic exploration of design alternatives. Design thinking stands firmly on three main pillars: empathy with users and human centeredness, rapid prototyping to generate large number of alternatives in order to solve the right problem rather than a problem right (the creative part), and last, but not least, their synthesis leading to best viable and feasible solutions that incorporate desired values [6]. IDEO [21], a design and innovation consultancy, has made a 60 minutes version of the process shown in Fig. 1. Even though the process appear to be simple and short, its power rests in its capacity to initiate deeper engagement with the problem space, that may last over time in some other form.

As mentioned in the Introduction, HCI students need to master numerous and diverse types of knowledge and gain practical design experience. Why make teaching of HCI even more complicated by introducing a creative proposition through design thinking and innovation explicitly?

III. THE CASE: TEACHING HCI WITH A CREATIVE WREE

A. *Previous Classroom Experiences with Design Thinking*

During the fall semester of 2013 two student project teams from a combined bachelor-master course in interaction design were introduced to design thinking. These teams worked with service design in the context of the University Library. They were given additional design thinking workshops, using service design cards and customer journeys as tools [22].

In addition, during the fall semester of 2013, a small graduate course of ten students, mentioned in the Introduction, adopted the design thinking approach and the studio based teaching. There, three student groups were taught about design thinking and focussed explicitly on nurishing creativity [15], [16]. They also were requireld to read articles like [23]–[28], in order to gain deeper understanding and knowledge of research through design, and design-centered research. In addition, successful examples of applications of design thinking were discussed [6], as was the work on reflection over the design practice as well as possibilities for understanding daily living practices as a design material, e.g., [29], [30].

Experiences from both classes indicated strongly that cultivation of creative thinking and making has a potential in teaching of HCI.

B. *The Course Setup*

The teaching approach that we argue for here is carried out within the combined bachelor-master course in interaction design. The course in question teaches traditional HCI research methods [31], and has two prior HCI courses as prerequisites. In addition to teaching research methods through lectures and small group learning sessions, the course aims to address the real-world problems, by offering a semester-long project in cooperation with external, local organizations. Usually, the class leadership involves ten or more organizations, soliciting two distinct proposals from each [32]. Students then form design teams and select one of the proposals, based on a first-come first-serve basis. The project work is, thus, anchored in a real need of some local company or organizations. Sometimes this need is not clearly formulated, rather, the company wishes to renew its offerings and they ask for new, open, creative solutions. Students may experience such open requests as intimidating at the start, as indicated by the fact that problems with narrow scope and clear goal tend to be selected first, while explorative problems are chosen last.

The students in the course are further supported (or challenged) in their learning efforts as follows: they are free to make mixed master-bachelor student groups, but master students need to read, understand and be able to use in their project previously published research in the domain that their projects cover. All teams have a regular, hour-long design feedback sessions during the conceptual design and prototyping phases of the project. A senior researcher and a representative of a company for which the students are designing participate in these sessions. In addition, all groups make an in-class, mid-term, presentations of their design efforts. The presentations are open to anyone from interested organizations, other faculty members, professional designers and any other relevant parties, sometimes also younger students. All present could give constructive feedback to presenters.

The course ends with the best project design competition. An independent jury of three HCI and design professionals judges the contest. The criteria for the jury are novelty, clarity of presentation, a potential impact of the designed prototype (relevance), validation of the prototype with users and overall design. This exact setup has been run for three consecutive years and has included surveys at the end of the semester. The surveys were individual and optional previously, but this last year they were focused on cooperation with the industrial partners, innovation and creativity, and thus were conducted in teams (team members were filling the survey together, having enough time to discuss and agree (or, sometimes not) on a common point of view.

Although the course addresses real-life problems, which would be typically solved by multidisciplinary teams within professional circles, this was not always possible to achieve in the context of the course. In other words, despite the presence of the senior researcher and company representatives, teams were not truly multidisciplinary, although, some teams came close. For example, some students had background in psychology, some in graphic design, others in arts. In such cases, they were encouraged to understand the assemblages of skills and knowledge that they had within the group, organize work so that their skills could be well used, and focus on knowledge production forms from which the team could benefit the most.

C. *The Use of Creative Thinking and Innovation*

The teams were free to choose and follow an approach of their choice, as long as they complied with general course requirements, as described above. The challenge was how to support best creativity within each team. A lecture on creative thinking and design thinking was given at the start of the course, introducing concepts of assemblages of skills and practices. The idea that one can design a set of practices that support creativity was introduced. These were further practically demonstrated and re-enforced during design sessions.

In addition, all external opportunities were sought and used to motivate students. For example, every year, during the fall semester, the dean of the University offers his annual innovation challenge to all the students at the University of Oslo, whether they study science, politics, social sciences or entrepreneurship. The challenge runs through several selection processes, until the winner is chosen among the best projects that made the final round of selections. The student teams in the course were strongly encouraged to participate. Two teams took up the challenge. This has, in addition to the usual interaction design course work, involved making a financial proposal and a business plan for implementation of the proposed innovation. Both teams made it to the final round. Judged entirely independently during the final competition for the best project in the course, they won the first and the third place.

The two student teams consisted of four second year undergraduate students each, and were supervised by a PhD student whose research relates to elderly living in a smart house. Thus, both projects address design for and with elderly, see [33] and [34] (both projects were delivered in Norwegian, but one group also posted the abstract in English of the paper that they are writing for HCII 2015 conference [35]). The latter project, see Fig. 2, developed a high fidelity interactive prototype utilizing frequency based technology (iBeacons) that helps elderly with cognitive difficulties to navigate complex buildings indoor.

Figure 2. SmartWalker: design and testing. Photos from [34].

Clearly, the effective use of the smart-walker requires mastery of the technology, but enables self-management to an increased degree, for the users that it is designed for.

The second project [33], see Fig. 3, focuses on self-management and bodily mastery [36]. The solution is based on a motion sensor (Kinect), and tracks exercises needed for bodily mastery and maintenance of the physical ailments.

Even though these two groups have achieved very nice results, they were certainly not the only ones that pursued the goal to be innovative and creative. Different ways in which this focus on creative thinking and innovation affected the work of the project teams is discussed in the next section.

Figure 3: An exercise system that enables correction of movements during the exercise session. Photo from [33].

IV. Discusion

The contextual differences among briefs presented to students by organizations that participated in this educational endeavor were substantial. Some teams were required to find new application domains for existing technologies, others to design new applications involving new technologies, yet others had to use old applications and old technologies, but find new ways of working with them. For example, a team had to work with the latest technology such as Google glasses and their potential use in crises situations by police or paramedics. Another team had a complex web-based software used in the oil industry that required creative thinking around how to help users to customize it. The vast majority of teams benefited from being inspired by at least one of the three main components of design thinking: empathy with users, rapid prototyping, or abductive thinking. The use was never enforced, so teams could choose to use any component of design thinking, none or a combination of design thinking with other practices used in industry. Reflecting on possibilities and knowing why design process (for any given project) involved certain tools, techniques, and methods, was required.

Empathy is a multifaceted construct that includes emotional recognition, vicarious feeling, and perspective taking [37]. Empathy was 'new' for many HCI students. While students were used to conducting user studies, they seldom tried to take the place of a user themselves and develop empathy with users in that way (through role playing, for example, in the wild). This might, in part, be due to the perception that by including personal experiences, subjectivity in the study would increase. Another reason might be that the phenomenological perspective, a characteristic of the latest wave of HCI, is still lagging behind in education. Regardless of the reason for empathy's 'newness', once tried, the students understood its benefits and could apply it creatively when working with conceptual development of their solutions. For example, a team needed to make a product that could be used in the children's hospital waiting room. Being an emphatic observers in the particular hospital's waiting room brought insight that, whatever they were to make, it should be quiet, it should not pass germs around, it could engage others, but should be fun, and OK, to interact with it alone. The result of this team's design and research efforts, a Leap motion controlled water fountain with LED lights, was fun to play with, had a pleasing, very soft sound of water and was nice to look at, see Fig. 4 and [38].

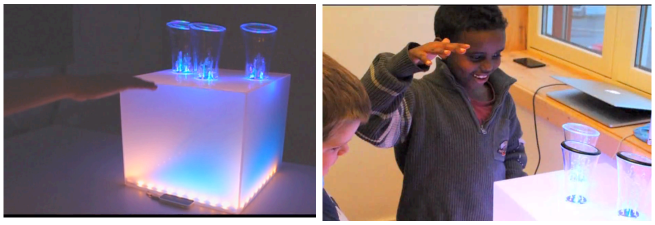


Figure 4. The water fountain project that uses LEAP motion to control the level of water. Photos from [38].

Rapid prototyping was frequently used. Students found that it enables easier communication of ideas, in particular across groups with different backgrounds and levels of knowledge. In one particular case, design meetings were attended by the student design team, a course instructor, a PhD student in interaction design, and a rather large, very interested group of library employees (between 5 and 8 persons at each meeting) with diverse backgrounds. Using pen and paper or tangible items such as service design cards for rapid ideation and construction of customer journeys was found to be helpful in such situations. Visualizing customer journeys using service design cards was valuable for creating a common understanding of certain services as they are today, and discussing points that offer opportunities for design interventions in order to improve those services. Unfolding one of those opportunities further was then undertaken using a workshop format, where all participants focused on producing as many rapid prototypes as possible, fostering good discussions around feasibility of solutions. The library experts could at once provide information on existing solutions and how the proposed new solutions could (or not) fit with the existing ones.

The last pillar of design thinking, abductive reasoning, is related to being able to synthesize solutions and optimize design, seeking to find the best option given the series of constraints. This is something that comes easier to people in design disciplines, rather than those using analytic way of thinking. Yet, some projects, among them the above mentioned [33], clearly show the ability to use synthesis.

At the end of the semester all teams filled a survey, providing 18 sets of answers. There were two questions related to creativity:

1) Do you think that the kind of work you did in this course is also creative?
2) What do you think about group creativity?

All teams answered the first one in affirmative. As for the second question, here are some of the answers (the answers were given in English, as presented, only the very last statement was translated from Norwegian): *“It really helps. Quite often you have some ideas, but you need help to be able to explain them. So in our group we really understood how each other was thinking, and we could really help each other describe and realize our ideas and creativity.”* Another team expresses it as follows: *“We have a group of different people with different ways of thinking, stirred together in a creative pot, it's awesome”!* The third considers that the *“group work increases creativity”*. The two most cautious expressions were the following two: *“We feel that the group works very well together, although this experience may vary”,* and *“Very good! Perhaps a bit too creative and ambitious”*.

V. CONCLUSION

The aim of this paper has been to inquire into the interplay between innovation, design thinking and creativity as educational channels that stand out as alternative or complementary to the ones traditionally used by HCI educators. The framework for learning about innovation, design thinking and creativity was introduced and explained. This setup has been repeated for the past two years and may be repeated by others. The concepts that have been helpful in cultivation of creativity were assemblages of skills and practices within multidisciplinary settings, empathy, rapid prototyping and abductive thinking. At the same time, care was taken not to reduce working with them as a specific procedure. Rather, tools, methods and techniques needed to be reflected over, and chosen in accordance with the problem at hand. Experimenting, or at least negotiating choices of research methods and techniques, was encouraged.

Further research is required regarding other frameworks and best practices for supporting creativity and innovation in HCI curriculums, including a comparative analysis of outcomes.

The achievements and learning outcomes in the here described course kept improving over the period of the last three years, as frameworks for supporting innovation and creativity got better and clearer described. The students' understanding of processes has also increased over time. The findings indicate that design thinking contributed to increased focus on innovation and creativity, as well as kept design processes wider and open for a longer period of time, fostering increased flexibility and adaptability in learning processes. The creativity and adaptability may be the best long-term goals that HCI education can add to its curriculums when preparing students for future work practices.

ACKNOWLEDGMENT

Many people deserve profound gratitude for helping with this work. Organization representatives who contributed with supervision and participation in design processes, PhD students and other professionals who helped student teams all have deep gratitude, as their work was entirely voluntary. Finally, thanks to all the students in the course, you have put in really good efforts and made it all worthwhile.

REFERENCES

[1] B. von Stamm, Managing Innovation, Design and Creativity, 2nd edition. Chichester, UK: Hoboken, NJ: Wiley, 2008.

[2] L. Kimbell, “Rethinking design thinking: Part I,” Design and Culture, vol. 3, no. 3, pp. 285–306, 2011.

[3] T. Brown and J. Wyatt, "Design Thinking for Social Innovation (SSIR, 2010)," Available from: http://www.ssireview.org/articles/entry/design_thinking_for_social_innovation. [Accessed: 20-Jan-2015].

[4] T. Brown, "Design Thinking," Harvard Business Review, vol. 86, no. 6, pp. 84, 2008.

[5] N. Cross, "Designerly Ways of Knowing: Design Discipline versus Design Science," Design Issues, vol. 17, no. 3, pp. 49–55, Jul. 2001.

[6] R. L. Martin, The Design of Business: Why Design Thinking is the Next Competitive Advantage. Harvard Business Press, 2009.

[7] A. L. Culén and M. Kriger, "Creating Competitive Advantage in IT-Intensive Organizations: A Design Thinking Perspective," in HCI in Business, F. F.-H. Nah, Ed. Springer International Publishing, 2014, pp. 492–503.

[8] T. Brown, Change by Design: How Design Thinking Can Transform Organizations and Inspire Innovation. New York, NY: Harper Collins Publishers, 2009.

[9] H. Collins, "Can Design Thinking Still Add Value?" Design Management Review, vol. 24, no. 2, pp. 35–39, 2013.

[10] K. McCullagh, "Stepping Up: Beyond Design Thinking," Design Management Review, vol. 24, no. 2, pp. 32–34, Jun. 2013.

[11] B. Nussbaum, "Design Thinking Is A Failed Experiment. So What's Next?" Co.Design. [Online]. Available from: http://www.fastcodesign.com/1663558/design-thinking-is-a-failed-experiment-so-whats-next. [Accessed: 20-Jan-2015].

[12] A. L. Culén and A. Følstad, "Innovation in HCI: What Can We Learn from Design Thinking?" in Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, 2014, pp. 849–852.

[13] Luxembourg, The Measurement of Scientific and Technological Activities Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data. OECD publishing, 2005.

[14] A. L. Culén, "Scaffolding Sustainability in the Academic HCID Practice," in Proceedings of the International Conference on Interfaces and Human Computer Interaction, 2014, pp. 45–54.

[15] S. Finken, A. L. Culen, and A. A. Gasparini, "Nurturing Creativity: Assemblages in HCI Design Practices," in Proceedings of DRS 2014, Umeå, 2014, pp. 1204–1217.

[16] A. L. Culén, H. N. Mainsah, and S. Finken, "Design Practice in Human Computer Interaction Design Education," in The Seventh International Conference on Advances in Computer-Human Interactions, 2014, pp. 300–306.

[17] E. F. Churchill, A. Bowser, and J. Preece, "Teaching and Learning Human-Computer Interaction: Past, Present, and Future," Interactions, vol. 20, no. 2, pp. 44–53, Mar. 2013.

[18] L. Bannon, "Reimagining HCI: Toward a More Human-Centered Perspective," Interactions, vol. 18, no. 4, pp. 50–57, Jul. 2011.

[19] T. T. Hewett, R. Baecker, S. Card, T. Carey, J. Gasen, M. Mantei, G. Perlman, G. Strong, and W. Verplank, "ACM SIGCHI Curricula for Human-Computer Interaction," ACM, New York, NY, USA, 1992.

[20] G. C. O'Connor, "Major Innovation as a Dynamic Capability: A Systems Approach," Journal of Product Innovation Management, vol. 25, no. 4, pp. 313–330, 2008.

[21] "IDEO | A Design and Innovation Consulting Firm." [Online]. Available from: http://www.ideo.com/. [Accessed: 20- Jan-2015].

[22] A. L. Culén and A. Gasparini, "Find a Book! Unpacking Customer Journeys at Academic Library," in The Seventh International Conference on Advances in Computer-Human Interactions, 2014, pp. 89–95.

[23] J. Zimmerman, E. Stolterman, and J. Forlizzi, "An Analysis and Critique of Research Through Design: Towards a Formalization of a Research Approach," in Proceedings of the 8th ACM Conference on Designing Interactive Systems, 2010, pp. 310–319.

[24] J. Zimmerman, J. Forlizzi, and S. Evenson, "Research Through Design as a Method for Interaction Design Research in HCI," in Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 2007, pp. 493–502.

[25] D. Fallman, "Design-Oriented Human-Computer Interaction," in Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 2003, pp. 225–232.

[26] E. Goodman, E. Stolterman, and R. Wakkary, "Understanding Interaction Design Practices," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2011, pp. 1061–1070.

[27] J. Löwgren and J. L. E. Stolterman, Thoughtful Interaction Design: A Design Perspective On Information Technology. MIT Press, 2004.

[28] B. K. Chakravarthy and J. Krishnamoorthi, Innovation By Design. Springer e-books, 2013.

[29] D. A. Schön, The reflective practitioner: How professionals think in action, vol. 5126. Basic books, 1983.

[30] L. Kuijer, A. de Jong, and D. van Eijk, "Practices As a Unit of Design: An Exploration of Theoretical Guidelines in a Study on Bathing," ACM Trans. Comput.-Hum. Interact., vol. 20, no. 4, pp. 21:1–21:22, Sep. 2008.

[31] D. J. Lazar, D. J. H. Feng, and D. H. Hochheiser, Research Methods in Human-Computer Interaction. John Wiley & Sons, 2010.

[32] Project areas. [Online]. Available from: http://www.uio.no/studier/emner/matnat/ifi/INF2260/h14/project-areas.pdf. [Accessed: 20- Jan-2015].

[33] "Motus - Kampen Omsorg+ - INF2260 - Høst 2014 - Universitetet i Oslo." [Online]. Available from: http://www.uio.no/studier/emner/matnat/ifi/INF2260/h14/presentations/kampen2/index.html. (in Norwegian). [Accessed: 20-Jan-2015].

[34] "SmartWalker - Indoor Navigation for Eldery - INF2260 - Høst 2014 - Universitetet i Oslo." [Online]. Available from: http://www.uio.no/studier/emner/matnat/ifi/INF2260/h14/presentations/kampen1/smartwalker---indoor-navigation-for-eldery.html (in Norwegian). [Accessed: 20-Jan-2015].

[35] HCII 2015 Abstract, SmartWalker. [Online]. Available from: http://www.uio.no/studier/emner/matnat/ifi/INF2260/h14/presentations/kampen1/abstracthcii2015.pdf [Accessed: 20-Jan-2015].

[36] A. L. Culén, S. Finken, and T. Bratteteig, "Design and Interaction in a Smart Gym: Cognitive and Bodily Mastering," in Human Factors in Computing and Informatics, A. Holzinger, M. Ziefle, M. Hitz, and M. Debevc, Eds. Springer Berlin Heidelberg, 2013, pp. 609–616.

[37] A. A. Gasparini, "Perspective and Use of Empathy in Design Thinking," in The Eighth International Conference on Advances in Computer-Human Interactions, ACHI 2015, ThinkMind, 2015.

[38] UTAIN – Interactive Fountain, [Online]. Available from: http://www.uio.no/studier/emner/matnat/ifi/INF2260/h14/presentations/LeapFountain/dokumenter/sluttrapport-utain.pdf (in Norwegian). [Accessed: 20-Jan-2015].

Web Based E-learning Tool for Visualization and Analysis of 3D Motion Capture Data

Andraž Krašček
eŠola d.o.o.
Slovenia
andraz@easistent.com

Kristina Stojmenova, Sašo Tomažič and Jaka Sodnik
University of Ljubljana
Faculty of Electrical Engineering
Slovenia
kristina.stojmenova@fe.uni-lj.si
saso.tomazic@fe.uni-lj.si
jaka.sodnik@fe.uni-lj.si

***Abstract*—In this paper, we propose an e-learning tool for visualization and manipulation of 3D data on a web platform. The data is streamed in real time from an optical motion capture system Qualisys consisting of eight infrared cameras and Qualisys Track Manager (QTM) software. A WebSocket protocol and WebGL application programming interface (API) are used to visualize and to interact with the data in a browser. The tool represents a web-based extension of QTM software providing also additional features and new possibilities to manipulate and analyze the data. We report also on a user study in which we evaluated the web based application and compared it with the original desktop-based application. The proposed application proved to be fast, effective and intuitive and can be used as an e-learning tool for demonstrating and teaching techniques for visualization and analysis of motion capture data.**

Keywords-motion capture; Qualisys; e-learning; 3D data; AIM model; WebGL; WebSocket.

I. INTRODUCTION

Optical tracking systems enable motion capture and recording of motion parameters of a selected object in space. They are commonly used to track motion in the field of biomechanics, industrial ergonomics, the moviemaking and entertainment industry, etc. Such systems consist of passive or active reflective markers that are placed on the points of interest of the monitored object, and infrared cameras that observe the motion of these markers in space. Active markers are light emitters, usually light-emitting diodes (LEDs), while passive markers are only light reflectors. Infrared cameras detect the light reflected from the markers and acquire their two-dimensional (2D) position in the recorded image. The system then combines 2D positions from all cameras and calculates the exact 3D position of markers in space. The tracking process begins with calibration, which provides information on exact positions and orientations of the infrared cameras. With higher number of cameras, a more accurate 3D position can be determined.

The goal of our research was to design and implement an e-learning tool for the visualization and analysis of motion capture data on multiple computers simultaneously. The analysis is of vital importance when dealing with motion and movements in sports. The tool we propose can for example be used for establishing the correlations between different segments of data (i.e., observing the amplitude of head movements in relation to knee angle when performing squats). Consequently, it needs to include numerous measurements, such as the length of selected bones, the angles between the bones, the velocity or acceleration of selected segments, etc.

By e-learning, we refer to a web-based system that makes information or knowledge available to students and teachers disregarding their geographic proximity and time restrictions [1]. The users can analyse and manipulate complex data sets remotely and use the tool to demonstrate different methods and procedures in real time. E-learning web applications are also affordable and easy to distribute to a large number of users by simply using a compatible web browser [2].

II. RELATED WORK

Visualization of 3D data in web applications has been addressed in numerous studies in different domains. In many cases, the rendering process was based on isosurface polygonization [3, 4, 5] enabled by various plug-ins in the browser. The most common issues related to this problem were specialized and dedicated programming languages, plug-in requirements, limited portability across browsers, devices and operating systems, and advanced rendering support [6]. Today, the rendering process has been significantly simplified by the technology called WebGL [7]. It is a JavaScript API based on OpenGL ES 2.0 for manipulation of 3D graphics in a web browser. It uses the OpenGL shading language, OpenGL for Embedded Systems (GLSL ES), and can be cleanly combined with other web content layered on top or underneath the 3D content. It is ideally suited for dynamic 3D web applications in the JavaScript programming language, and has been fully integrated in all leading web browsers.

Several researchers reported on using WebGL for monitoring and interaction of 3D graphics in a web browser. A lot of them exposed the benefits of using web-based applications in order to lose dependency of hardware.

Conote, Segura, Kabongo, Moreno, Posada, Ruiz discussed performance and scalability of the volume rendering by WebGL in different application domains [3]. In their work, they presented how implementation of a direct volume rendering system for the web articulates in efficient manner the capabilities of WebGL, making the formerly unusable accelerated graphic pipeline available.

 ISBN: 978-1-61208-382-7

WebGL with the ability to put hardware-accelerated 3D content in the browser represents a mean for creation of new web based applications that were previously exclusive for the desktop environment. C. Leung and A. Salga discussed how mid-level APIs can help develop web applications that do not only copy the desktop application but can contain unique 3D content as well [9].

3D monitoring of static objects has also been a subject of research. Museums and similar institutions show growing interest into showing their collections to a wider public trough the web. Schwartz, Ruiters, Weinmann and Klein have proposed a WebGL-based presentation framework, which does not only provide a 3G geometry, but a powerful material representation, capable of reproducing the full visual appeal of an object as well [10].

In this paper, we propose a web based e-learning tool for visualization and manipulation of 3D motion capture data in real time and on high number of computers simultaneously. It also supports observation and analysis of various motion parameters of observed objects and models, as well as active collaboration between different web users. Our goal was not only to develop an application that will provide access to data for observed models but also to provide a positive user experience and good usability of the system [11]. Based on the positive experience with WebGL API reported in the Related Work section, we selected this technology to develop the user interface in a browser.

A. *Our research contribution*

To our knowledge, this is the first example of the visualization and manipulation of motion capture data on a web platform supporting a high number of simultaneous users. The research hypothesis of our work is that by enabling the instructors and students to work with motion capture models through web-based clients, the teaching/learning experience will be greatly enhanced. The main field of use of the application is academia and education where it can be used as collaborative tool for teaching motion tracking system techniques and data processing methods.

In the rest of the paper, we describe the proposed system and the corresponding user interface. We also report on a user study performed to evaluate the usability and user experience of our web application in comparison to the original desktop motion capture software. The results of the experiment are presented and statistically analysed. The Discussion section summarizes the most important findings and proposes some ideas for future work.

III. System Architecture

A. *Qualysis Track Manager*

In our research, we use the professional motion capture system Qualisys [12]. The system consists of eight high-speed cameras, a set of passive markers and the proprietary tracking software called Qualisys Track Manager (QTM) [13]. This is a complex desktop application, which calculates the exact 3D position based on separate 2D images from all cameras. It also takes care of infrared (IR) cameras calibration, motion capture recording, creating and editing models, analyzing data on models, streaming captured motion, etc. It runs on a standard PC and exchanges data with the cameras trough a standardized Ethernet protocol.

QTM shows the 3D position of each marker in a Cartesian coordinate system as a coloured dot. Individual markers or a group of markers can be labelled and connected to a structure called "model". Each pair of markers with constant inter-distance can be connected with a line called bone (due to its rigid structure). When tracking the motion of humans the QTM bones correspond to the bones of human body.

The created model, which consists of a set of markers and bones, can be saved for future measurements as an Automatic Identification of Markers (AIM) model. However, the visualization and analysis of the stored AIM model can be done only on a single computer running the QTM software. Figure 1 shows a screenshot of QTM with an example of AIM model (the model represents an upper part of human body).

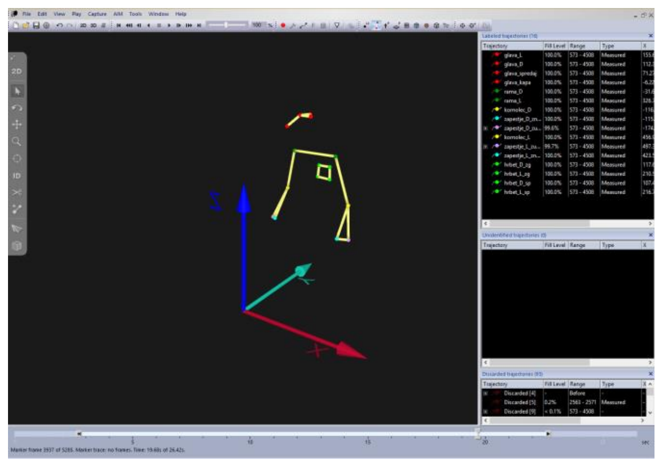


Figure 1. The visualization of an AIM model in QTM software.

QTM is a very complex tool, which supports a high variety of commands and features. The user interface is therefore rather complicated and not very intuitive. It is primarily intended for controlling the cameras and not for the analysis of the recorded data. The latter is rather limited and cannot be saved or exported in a way that it could be used in different application. Since it does not allow simultaneous work of a larger number of users it is inappropriate to use it as an academic tool.

B. *Web Based E-learning Tool*

The architecture of the web application is divided into three levels as shown in Figure 2 [14]. First, data on marker coordinates is streamed from QTM and stored to a special buffer on a Node.JS web server [15].

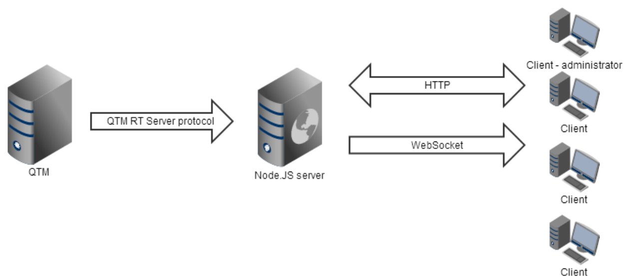


Figure 2. The web application system architecture [14].

Communication between the QTM and the web server is based on real-time (RT) protocol (Qualisys' proprietary protocol for streaming data in real time). Server acts as a hub and translates raw data to JavaScript Object Notation (JSON) and broadcasts it to all connected clients through a WebSocket protocol [16]. The latter allows full-duplex communication between a server and a client and its packet headers are smaller than HTTP's.

The client side application is divided into two main modules. WebSocket module is used for communication with the server while the WebGL module is responsible for rendering and interaction with 3D space through navigation panels. The WebSocket module reads the server data stream and updates the local storage when new marker data is received. It is built to be as lean as possible because both modules run in the same thread and the goal is not to block the WebGL module while it is rendering 3D data. WebGL module on the other hand is constantly rendering 3D space and exposing it in canvas HTML5 element. Constant rendering is needed for smooth user interaction with 3D space. When user moves, jaws, pitches or zooms in/out the 3D scene only the view matrix is recalculated and transformed. On each render loop, the markers' coordinates are read from local storage and the bone data is loaded from the model. Finally, based on the view matrix the scene-space is transformed to view-space.

1) User interaction with 3D space

The main goal of web application is the visualization of 3D space with markers and corresponding AIM models. Markers are coloured in colours defined in QTM application to help the user to differentiate between groups of markers. User can interact with 3D space using the mouse. While holding a mouse button and dragging a cursor user can rotate 3D space. If user drags the cursor horizontally, 3D space rotates around Z axis. If user drags the cursor vertically, 3D space moves around X or Y axis depending of Z rotation. User can also zoom in or zoom out by turning a mouse wheel. This command can also be initiated by using keyboard keys "page up" and "page down" in case user's mouse lacks the wheel. User can select a single marker by clicking it or an array of markers by holding "ctrl" key on keyboard while clicking several individual markers consequently. The click on an already selected marker will deselect it.

2) Creation of a model

A model can be thought as an undirected graph where markers represent nodes and bones between markers represent edges. In the web application, we tried to design a process of creating models as seamless as possible. To create a bone a user must select two individual markers in 3D space. The bone is initialized by clicking the *BONE* button on the right bottom section of the dashboard. Only one bone can exist between the two individual markers. If multiple markers are selected in a consequential order multiple bones will be created between these markers. When all the bones in the model are created a user can save the model by selecting the corresponding command in the models navigation panel and defining its name. Models can be saved, loaded to a set of markers in 3D space or deleted.

The model navigation panel is positioned in the upper right corner of the screen and it is divided in two subpanels. It allows a quick overview of all saved models and interaction with models. It features all possible actions a user can execute on the currently loaded model. The screenshot of the application is shown in Figure 3.

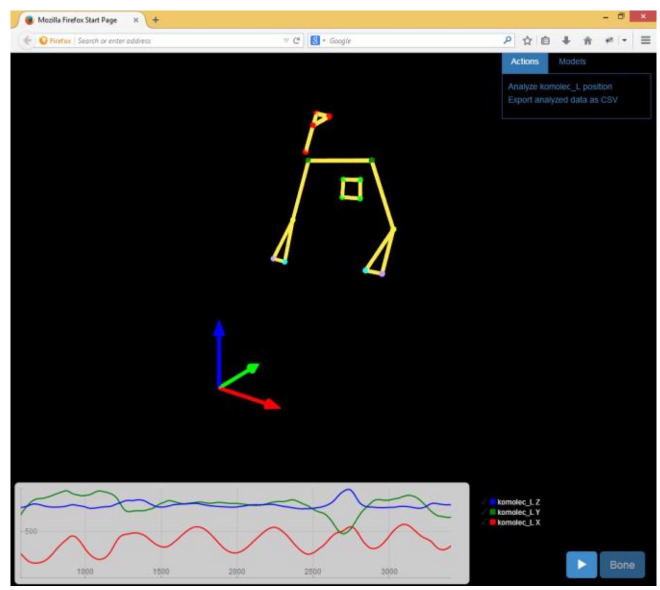


Figure 3. The web application with simple model and marker position analysis chart.

3) Manipulation with models and analysis

At each time, only currently available actions are visible in the panel depending on the current state of the model. The goal was to reduce the number of menus and settings in order to simplify the user interface and increase its intuitiveness. When, for example, only one marker is selected the only available action is the analysis of its position in space as a function of time. When two markers are selected the length between these two markers can be analyzed as well their individual positions in space. If two bones connected by a shared common marker are clicked and selected, it is possible to analyze an angle between those two bones as a function of time.

Data from all types of analysis are presented in an interactive chart. The X-axis shows time frames and the Y-axis shows the corresponding distance or angle (the unit is therefore expressed in millimetres or angular degrees). The exact value of the individual time frame can be extracted by dragging the timeline bar across time frames and locating the desired frame. Figure 4 shows two examples of analysis charts for an angle (in degrees) between two bones and for a position (in millimetres) of a selected marker in space respectively.

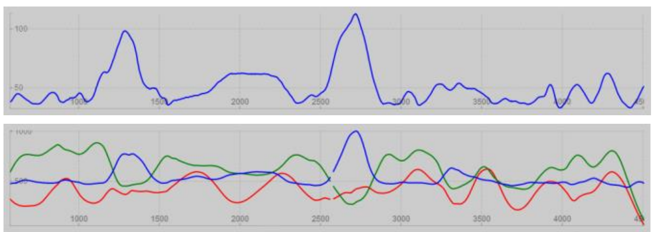


Figure 4. An example of two analysis charts showing the angle between two bones (upper figure) and position of a marker in space (lower figure) respectively.

All analysis' results can be exported by selecting the *EXPORT* action. The supported output formats are CSV (Comma separated Values) and JPG image. The CSV format can then be imported by majority of available third party software.

4) User collaboration and data synchronization

The most important feature of the proposed e-learning tool is the possibility of simultaneous work and collaboration of high number of users. A typical scenario is when a lecturer demonstrates various motion capture techniques and scenarios by operating QTM software and IR cameras in real time. Students participate in the experiment at the same location or remotely through an internet connection. They can use any stationary or mobile device with a web browser supporting WebGL API. They all work on the same stream of data (same set of markers) but the interaction with the content such as creation of models and corresponding analyses are individualized.

Additional feature of the application is a collaboration tool, which provides methods for synchronization of users' data. All users connected to the same session can synchronize their local AIM models. Each user can propose and send his or her model to other session members or load the model proposed by other members. In this way, students have an option to participate actively in the manipulation of data or to be just passive observers. In the second case, a model, which is a subject of analysis, is built and sent to their application by other user (e.g., the lecturer). However, they can still fully manipulate the view of the 3D scene and interact with the model.

IV. USER STUDY

In order to test the proposed e-learning tool and to evaluate its effectiveness and intuitiveness, we conducted a user study in which we compared the tool with the original QTM application. The two applications represented two independent variables of the experiment. The three dependent variables were:

- Task completion time or the time required to solve given tasks;
- Subjective evaluation of the applications assessed with User Experience Questionnaire (UEQ) questionnaire;
- General subjective remarks given by the participants.

A. Participants

A total of 15 students of multimedia, electrical engineering and computer science participated in the user study. The subjects ranged from 20 to 28 years of age (M=23.1 years, SD=2.4 years). All participants reported normal eye sight, except for one who was wearing glasses due to myopia. Eight participants had prior experience with the visualisation and manipulation of 3D data (applications such as Blender, Google SketchUp, SolidWorks, etc.). They were randomly distributed into two experiment conditions described bellow. None of them had any experience with motion capture systems whatsoever.

B. Experiment design

The experimental design was a between-subject, dividing participants into two groups to avoid sequence effects and confounds (due to the similarity of commands and controls in both application). The first group (8 test subjects) performed the set of selected tasks with the proposed web application while the second group (7 test subjects) performed same tasks with the QTM application. Prior to the experiment the participants were given a short explanation on how motion capture system works and about the purpose of the application they were about to use. They were also given approx. five minutes to get familiar with the application and its interface. Each participant conducted the experiment individually following the instructions of the experiment leader.

In the first part of the experiment, each participant was asked to perform 3 different tasks:

- Build a model by connecting markers with bones (T1);
- Save and load the created model (T2);
- Perform an analysis on the built model (T3); This task was broken down to 5 subtasks:
    - Analyze position of the right elbow;
    - Analyze length of the right upper arm;

- Analyze the angle of the right elbow;
- Read the exact value of the previously analyzed angle for the specific time frame;
- Save the analyzed data to a file in raw format.

In the second part of the experiment, the participants from both groups had to fill in a standardized UEQ and evaluate the used application [11]. The UEQ questionnaire assesses the participants' feelings, impressions and attitudes towards the tested application. The questionnaires were translated into Slovenian language – the native language of the participants.

In the third and the final part of the experiment, the participants were asked to give their opinion and general remarks about the tested application. Any comment or remark about the application at any stage of the experiment was also registered by the experiment leader.

All computers running QTM, NodeJS server and web browser were connected through a local network providing high bandwidth. As a consequence, the page response and loading times as well as latencies on the network were very low and did not affect evaluation procedure.

V. RESULTS

The three variables evaluated in the user study were the following:

- Task completion times;
- UEQ;
- Subjective comments about the applications and user interfaces.

A. *Task completion times*

The time required to complete the individual task was measured manually by the experiment leader. A timer was started just after the experiment leader would read the instructions for the selected task and stopped after when the participant was comfortable with the result achieved for that task. Figure 5 shows all task completion times for both applications. The between subject ANOVA and the post-hoc Bonferroni tests with a 0.05 limit on family wise error rate were used for the comparison of data (the normal distribution of data was confirmed with Shapiro-Wilk test).

The proposed web application seemed to be slower than the original QTM application for creating bones and building models (task T1). However, no statistically significant differences were found in this task ($F(1,12) = 3.733$, $p = 0.077$). There was also no significant difference for T2 ($F(1,12) = 2.037$, $p = 0.179$) and for T3.1 ($F(1,12) = 4.505$, $p = 0,055$). For all the other tasks the proposed web application outperformed the QTM application:

T3.2: $F(1,12) = 20.618$, $p = 0.001$;
T3.3: $F(1,12) = 15.826$, $p = 0.002$;
T3.4: $F(1,12) = 43.153$, $p < 0.001$;
T.3.5: ($F(1,12) = 14.182$, $p = 0.003$;

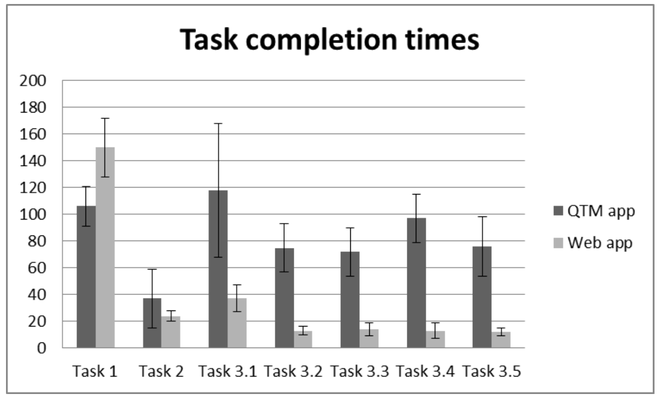


Figure 5. Average task completion times (in seconds) with confidence intervals.

The standard deviation is smaller and the corresponding confidence intervals are narrower for all tasks performed with the proposed web application. We believe this reflects intuitiveness, reliability and robustness of the proposed web interface.

B. *UEQ*

After completing the set of tasks, the participants were asked to complete the UEQ questionnaire. The questionnaire consists of 26 individual statements, which are graded with a seven stage Likert scale to reduce the well-known central tendency bias for such types of items [18]. The results of these grades are grouped to 6 different categories (attractiveness, perspicuity, efficiency, dependability, stimulation and novelty). Figure 6 shows average UEQ scores for individual categories.

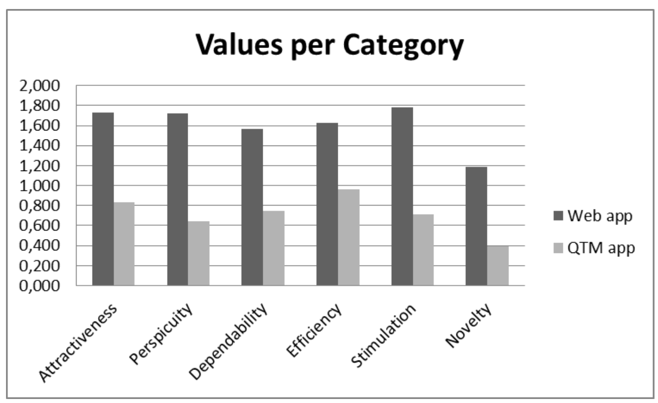


Figure 6. Average UEQ values for six categories.

The mean scores of the proposed web application were higher in all six categories. Again due to the normal distribution of data, the ANOVA and post-hoc Bonferroni tests with a 0.05 limit on family wise error rate were used for comparison. The statistically significant difference between the applications was found only in the categories attractiveness, perspicuity and dependability:

Attractiveness: $F(1,12) = 11.703$, $p = 0.005$;
Perspicuity: $F(1,12) = 11.456$, $p = 0.005$;
Dependability: $F(1,12) = 12.928$, $p = 0.004$;
Efficiency: $F(1,12) = 2.001$, $p = 0.183$;
Stimulation: $F(1,12) = 3.123$, $p = 0.103$;
Novelty: $F(1,12) = 0.795$, $p = 0.390$;

C. General remarks

After each experiment, we interviewed the subjects about their experience. We were primarily interested in some general comments and remarks about the proposed web application. In this section, we list some comments and suggestions for improvement expressed by several test subjects:

- markers were too small and hard to select;
- there should be more keyboard shortcuts for some commonly used actions;
- some descriptions of results of analyses are unclear;
- double-click could be used to list available commands and actions at any time;
- etc.

Some participants who had experiences with 3D applications and 3D interfaces reported to have problems using the current interface due to new interaction methods and metaphors introduced in the software.

VI. Discussion and Conclusion

The main goal of this user study was an evaluation of the proposed and developed web application to reveal its advantages and disadvantages in comparison with an existing QTM application. We aimed to prove that users' actions can be performed faster and more intuitively with such task-specific application compared to a general-use application. The final scope of functionalities and features for the analysis of the captured data is comparable in both applications as they both require similar complexity of analysis and manipulation of the recorded data. Since the proposed application runs on a web platform as an e-learning tool, it supports simultaneous use of high number of users as well as their collaboration and exchange of knowledge.

The user interface of a web application proved to be equally fast or faster for a majority of tasks. However, no significant difference was found in T1 in which users were asked to build bones and an AIM model for further analysis. This result was not expected since the new interface was improved with additional commands and tools enabling the creation of multiple bones at once. Since no specific instructions were given about these tools only a few users explored and effectively used these features.

For the T2 and T3 the web application was faster than the QTM application, which was an expected outcome. Our application was built primarily to enable various analyses on a model, such as, for example: detection of marker position, identification of length of a bone, an angle between two bones, etc. The user interface was therefore optimized to support these actions and to make them intuitive. On the other hand, in the original QTM application the analyses and the corresponding actions are just a small set of available features and several users have difficulties finding them among all other actions and menu commands.

The second evaluated parameter in the study was the subjective perception of both applications measured through UEQ test. The web application was judged to be significantly more attractive (attractiveness category), understandable and clear (perspicuity category) and dependable (dependability category). We believe these high scores reflect simple, clear and intuitive user interface, which adapts to user actions and changes its state and a set of available controls depending on what the user is currently doing. The lower score on the other hand were given in the category novelty. There was no significant difference found between our web application and the original QTM application. We believe these scores reflect the facts that majority of test subjects had no or very few experiences with motion tracking techniques and visualizations of 3D data. The main part of the user interface was a simple 3D visualization of a Cartesian coordinate system with a set of coloured points (markers) at different positions, which the user did not find very novel or exciting. We also believe a more significant difference could be found if a within subject test was performed enabling the users a direct comparison of both systems and user interfaces.

The collected set of subjective comments revealed some problems and ambiguities of the proposed interface as well as missing commands and potential extensions. Several comments referred also to visualization problems, which are related to the WebGL API and their improvement is out of our power. These comments will be used primarily to improve the proposed interface in terms of its efficiency and clearness and its upgrade with new features and commands.

Our future goal is to extend the set of available features for the analysis of motion data in the applications as well as the features related to remote collaboration and exchange of data. Another important module which is currently not available and should be implemented in the future is a common platform supporting predefined learning processes and tasks, the authentication of users, the creation and storage of the users' profiles, the monitoring of learning progress, etc.

We believe our application demonstrates the high usability of modern web technologies for the development of new powerful and rich services in the e-learning domain. Real time streaming of complex 3D data and their visualization in an interactive scene in a browser are an excellent use case for many other similar services. In the future, the impact of the proposed e-learning software on the learning performance and methodology should also be evaluated.

References

[1] P. Sun, R. J. Tsai, G. Finger, Y. Chen, and D. Yeh, “What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction,” Computers & Education, vol. 50, issue 4, May 2008, pp. 1183-1202.

[2] A. Anyuru, Professional WebGL Programming: Developing 3D Graphics for the Web, John Wiley & Sons, Apr 23, 2012.

[3] P.-A. Fayolle, B. Schmitt, Y. Goto, and A. Pasko, “Web-based constructive shape modeling using real distance functions,” IEICE - Trans. Inf. Syst., vol. E88-D, no. 5, May 2005, pp. 828–835.

[4] R. Cartwright, V. Adzhiev, A. A. Pasko, Y. Goto, and T. L. Kunii,“Web-based shape modeling with hyperfun,” IEEE Computer Graphics and Applications, vol. 25, no. 2, 2005, pp. 60–69.

[5] Q. Liu and A. Sourin, “Function-defined shape metamorphoses in visual cyberworlds,” Vis. Comput., vol. 22, no. 12, Nov. 2006, pp. 977–990.

[6] J. Behr, P. Eschler, Y. Jung, and M. Zollner, “X3dom: a dom-based html5/x3d integration model” in Proceedings of the 14th International Conference on 3D Web Technology, 2009, pp. 127–135.

[7] WebGL - OpenGL ES 2.0 for the Web. Available from: http://www.khronos.org/webgl/, 2014.10.03

[8] J. Conote, A. Segura, L. Kabongo, A. Moreno, J. Posada, and O. Ruiz, “Interactive visualization of volumetric data with WebGL in real-time” in 3D Web Technology, 2011 pp. 137-146.

[9] C. Leung and A. Salga, “Enabling WebGL”, World wide web, 2010, pp.1369-1370.

[10] C. Schwartz, R. Ruiters, M. Weinmann, and R. Klein, “WebGL-based streaming and presentation framework for bidirectional texture functions”, Virtual Reality, Archaeology and Cultural Heritage, 2011, pp. 113-120.

[11] J. Nielsen, “Usability 101: Introduction to Usability”. Available from: http://www.nngroup.com/articles/usability-101-introduction-to-usability/, 2014.10.10

[12] Qualysis Motion Capture Systems. Available from: http://www.qualisys.com/, 2014.10.10

[13] QTM – Qualisys Track Manager, User Manual, 2011.

[14] A. Kraščeк and J. Sodnik, “Qualisys Web Tracker – A web-based visualization tool for real-time data of an optical tracking system”, ICIST 2014 Proceedings, 2014, pp. 155-160.

[15] S. Tilkov and S. Vinoski. "Node. js: Using JavaScript to build high-performance network programs." IEEE Internet Computing, vol. 14, issue 6, 2010, pp. 80-83.

[16] The WebSocket Protocol. Available from: https://tools.ietf.org/html/rfc6455, 2014.10.06

[17] What is the UEQ? Available from: http://www.ueq-online.org/, 2014.10.06

[18] B. Laugwitz, T. Held, and M. Schrepp, “Construction and Evaluation of a User Experience Questionnaire”. Available in: Holzinger, A. (Ed.): USAB 2008, LNCS 5298, pp. 63-76.

Orientation Aids for Mobile Maps

Jussi P. P. Jokinen and Pertti Saariluoma
Department of Computer Science and Information Systems
University of Jyväskylä,
Jyväskylä, Finland
jussi.p.p.jokinen@jyu.fi, pertti.saariluoma@jyu.fi

***Abstract*—Using mobile maps to represent urban, work, or entertainment environments offers new possibilities to plan and carry out tasks. One potentially critical problem in mobile map usage is the misalignment between the user's frame of reference and the frame of reference of the map. In the experiment reported here, three different 'orientation aids' were tested in the context of restricted space, such as a large factory hall. The aim of the study was to find out how user interface design can help the user mentally align misaligned frames of reference for efficient mobile map use. The results of the experiment (N = 12) suggest using a 'you are here' marker and landmark highlighting, while canonical direction symbols proved to be less plausible. Further, a maximum number of seven targets on the map is suggested.**

Keywords- mental spatial orientation; mental rotation; mobile maps; orientation aids; reference-frame misalignment.

I. INTRODUCTION

It is common to consider people using the concepts of psychology. If the focus is on a developing child, the branch of psychology is called development psychology; if it is on car drivers, it is common to speak about traffic psychology. Therefore, when we consider people as users of technology, it makes sense to speak about user psychology [1][2][3]. The leading idea of modern user psychology is explanatory coherence, which means that researchers of human-technology interaction consider how consistent the outcomes of their research are with what we in general know about human mind, that is, cognitions, emotions, motive, and personality [2]. Here, the focus is on the possible explanatory roles of human visuo-spatial memory in analysing users working with maps.

People acquire knowledge from the environment by directly experiencing and interacting with it. However, in many tasks, information beyond direct spatial experience is required. A map is one widely used tool to provide such information. Inherent in the map usage, however, is the problem of combining environmental information from two sources, direct spatial experience and the map. A map provides an objective spatial representation of the environment with a world-oriented, that is, exocentric (or allocentric) frame of reference [4][5]. On the other hand, people themselves represent environment with respect to their own position, that is, egocentrically [4][5]. This difference in the frames of reference may lead to problems during a map use due to a phenomenon called reference-frame misalignment [4][6]. When the frames of reference are misaligned, a lot of evidence suggests that map-based decision times increase and become more error prone [4]-[9].

The use of small mobile displays in work and entertainment contexts has increased, and hence the problem of reference-frame misalignment has become relevant in mobile map design [7][10][11]. If such problems are to be solved efficiently, scientific knowledge concerning the domain of map use and reference-frame misalignment should be provided for the designers of mobile maps. As mentioned, the practice of combining useful psychological knowledge with practical solutions is called user psychology [2][3]. In our case, this means generating experiment-based insights for both cognitive and practical understanding of mobile map use using psychological theories, methodology, and concepts. This is the aim of the present article, which will operate in the domain of designing mobile interfaces for large enclosed spaces, such as factory halls. An example of a mobile map useful in such a context is a map-based controller for crane automation. In the design of map-based mobile interfaces for operating with complex systems, such as port or industry cranes, any errors from reference-frame misalignment need to be minimised due to safety and efficiency concerns.

In tasks requiring map-based and experience-based environmental information, task times and error rates seem to increase in a proportion to the misalignment between the user's and the map's reference-frames. This effect is due to the mental rotation required to combine information from the two misaligned sources [6][8][9][12][13]. The cognitive demand of mentally rotating maps may also be associated with the complexity of the stimuli especially if the environment is new to the user [6][14]. The amount and complexity of stimuli on the map is especially critical in the case of mobile maps, as the screens are usually small and the interface must be designed frugally [15]. Additionally, as the number of objects on a small display increase, the visual search time increases [16]. It seems that the limited amount of short-term capacity limits the efficiency of map use due to the increased demand in both visual search and mental rotation [9][17].

Different user interface modifications for maps have been proposed to reduce the cognitive demand caused by the misaligned reference-frames. Rotating the map automatically is one obvious proposal, but its effect on the efficiency of the map use is not clear [18], cf. [11]. Letting the user choose between exocentric and egocentric perspectives, depending on the use situation, would often be preferable [19], but this effect may be difficult to accomplish with small displays and would be solving a problem by including new design problems. Further, automating the map rotation completely may be confusing to the user. However, there exists a variety of less intrusive ways to aid the map user in both mental rotation and visual search required in efficient map use.

For the experiment reported here, three map aids were designed to help the user mentally rotate a stationary, non-rotating mobile map. These 'orientation aids' were used to investigate the problem of misalignment in mobile maps and provide suggestions for the design of mobile map interfaces. The amount of objects on the map (map complexity) and reference-frame misalignment were manipulated to find out how different orientation aids would help the map user in making inferences based on the combined information from a map and an environment.

The first aid, a 'You are here' marker, familiar to the users of GPS navigators and stationary shopping mall maps, shows the location of the user on the map [15][19][20]. Maps with 'you are here' marker are not always as useful as one might imagine, especially if the marker is not accurate [21]. However, locating the user accurately on the screen has become technologically easier. For example, the studies of 'you are here' markers conducted in virtual worlds show that they are relatively easy and effective map aid, e.g., [19][20]. 'You are here' markers seem to aid processing of spatial information by offering cues to one's location and its relation to the environment, but only if the location is salient and asymmetrical [22].

The second aid, landmark highlighting, makes a salient environmental object or feature visible on the map by making it distinct among the other map elements [23][24]. Conceptually and perceptually distinct locations (i.e. landmarks), have been shown to be effective aids in pedestrian navigation, for example [24]. Especially, if the map is complex, hierarchical, and has many objects, highlighting one of the more salient features provides for faster visual search [23]. Highlighted landmark also helps in mental rotation, as the user can discard other, less salient objects from the process and focus on visually salient features [24].

Third, direction symbols, such as the compass pointing to the north of the map, are often used to refer to the cardinal directions used in the map [4]. In a closed space, such as a factory hall or a shopping mall, it is possible to associate cardinal directions of the environment with the map. For example, associating the north wall of a factory with a certain symbol, and adding the same symbol to the map of the factory, may help the map user to align the two reference frames. Direction symbols associating the map directions with environmental directions may be especially effective when the environment is known to the map user [6][14].

Following hypotheses were formulated and tested in a laboratory experiment. The first two hypotheses were included to confirm that there indeed would be such effects as mental rotation and visual search associated with map use. Hypotheses H3–H5 were derived from the three orientation aids.

H1. As the misalignment between the referential frames of the map and the environment increases, map based decision times increase.

H2. Adding more objects increases map based decision times.

H3. Adding a 'you are here' marker to indicate the location of the user in the environment decreases map based decision times.

H4. Adding a landmark highlighting to represent a salient environmental feature decreases map based decision times.

H5. Adding direction symbols, referencing to the environmental directions decreases map based decision times.

The structure of the paper is as follows. Section 2 describes the laboratory experiment designed for testing the hypotheses above. The results of the experiment are reported in Section 3 and discussed in Section 4. The final Section 5 contextualises these results.

II. Method

Twelve (N = 12) participants were recruited for the first experiment. Their mean age was 23.8 years, SD = 3.3, and six of them were male and six female. All of the participants were right-handed, and had at least some experience with maps: when asked to estimate their own map using skills on a scale from 1 ('no skills') to 5 ('expert'), seven of the participants gave the scale midpoint, 'average skills', and five reported having 'advanced skills' (scale value 4). This self-rating was not found to correlate with performance in the experiment.

Two sets of pictures were made, one for environments, and one for maps representing those environments. The goal was to create an abstract closed space with salient objects, such as one could find in a large factory hall. The environment pictures were photographs of wooden cubes, which were placed on a white canvas in an otherwise empty space with grey walls (Fig. 1). Multiple environment pictures were taken with following modifications. First, the number of the wooden cubes was four, seven, or ten. Second, camera was placed either at the front of the canvas (0°), the corner of the canvas (45° counter-clockwise), or the right side of the canvas (90° counter-clockwise; shown in Fig. 1). As the third modification, in some environments, one of the white wooden cubes was changed to a red, otherwise identical, cube, as shown in Fig. 1. For a number of different environments, two symbols, a triangle and a circle, were added to the wall. All cubes were identified with a letter.

The map pictures were constructed to accurately represent the configuration of the wooden cubes in the environment as black 'target' icons (Fig. 2). When the environment contained a red cube or direction symbols, corresponding orientation aids were added to the map: a red target instead of a black one to highlight a landmark, and direction symbols at the sides of the map to match the symbols on a wall. In order to test 'you are here' markers, a red dot was added to the map to indicate where the environment picture was taken, that is, where the participant

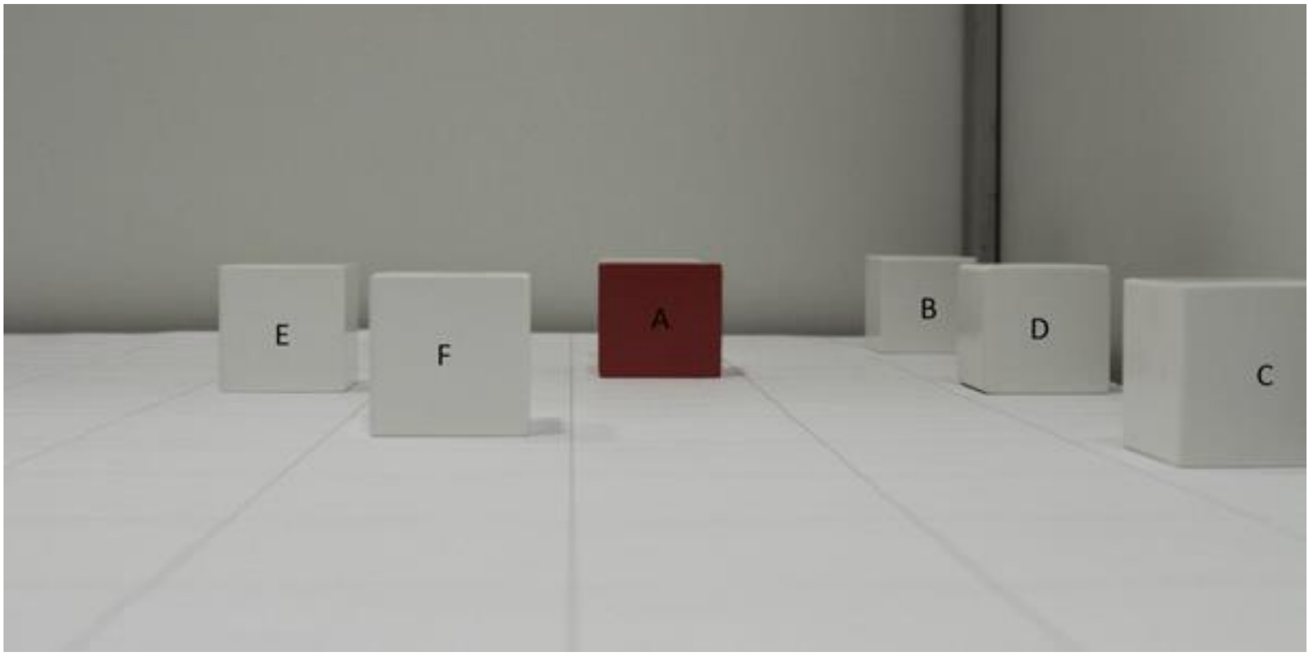


Figure 1. An example of the environment picture used in the experiment with landmark highlighting and 90° camera positioning. Corresponding map picture is seen in Fig. 2.

was 'standing' in the environment. Half of the maps were rotated 180° to increase the variation in misalignment between the environment and the map picture. This brought the total number of possible misalignments to five: 0°, 45°, 90°, 180°, and 135°. Due to environment rotations, some of the cubes were partly or almost fully occluded (e.g., Fig. 1, behind the A block). However, this did not benefit any single orientation aid, as the map templates were same for all tested aid.

One task consisted of an environment picture, a corresponding map picture, and a statement in the form of matching letter and number. (e.g., 'A = 5'). The task of the participant was to judge whether the statement was correct or incorrect by pressing either green (correct) or red (incorrect) button in front of them. Half of the statements were correct, and half incorrect. Each participant was presented 72 tasks (3 different number of cubes, 2 map rotations, 3 environment rotations and 4 orientation aids including a no aid -baseline). The order of the tasks was randomised for each participant. Before the tasks, the participants were given four easy tasks to practice the procedure.

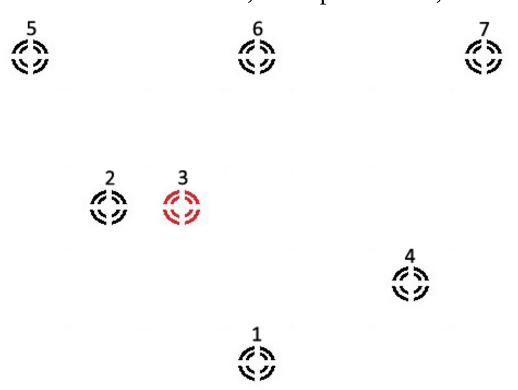


Figure 2. An example map picture, corresponding to the environment picture in Fig. 1. Red cube (target 3) is highlighted.

The participants were seated in a quiet laboratory room alone with the experimenter. In front of them, they had a 40-inch computer screen for the environment pictures, and to the left of it a 17-inch computer screen for the map pictures. The environment picture was scaled to take the whole 40 inch of the larger screen, but the map picture took less than half of the smaller screen and was therefore closer to a real mobile map size. A keyboard with one green and one red button was placed in front of the large screen. For half of the participants, the green button was on the left-hand side, and for the other half, on the right-hand side. The participants were instructed to have the index fingers of their both hands at the buttons. The participants completed a task by pressing either green or red button to indicate their agreement with the statement concerning the relation of the environment and the map. The reaction time (RT) was written into a log file, and the next task given after a brief pause. RTs have been often used in experimental studies of cognitive processing, also in studies of mental rotation and orientation [8][13]. Shorter RTs indicate that the processes associating stimuli with actions are less mentally demanding.

The data analysis was conducted using multilevel modelling ('Generalised linear mixed model' procedure in IBM SPSS 20), which is suitable for analysing nested longitudinal data, such as repeated RT measures within different experimental manipulations [25]. Contrasted, for example, with repeated measures analysis of variance, multilevel models are better suited to analyse nuanced effects, such as the size of individual intraclass correlation (i.e. how much RTs of a single participant correlate with each other) and learning during the tasks. The problem of individual differences in RTs (some participants tend to be faster than others in overall performance) has previously been solved with data normalisation [8], but multilevel models take this effect better into account.

In the multilevel model, RTs were predicted as gamma distribution with a log link, because the target distribution had only positive values and a positive skew (as is often the case with RT distributions; the gamma assumption was supported with a Q-Q plot) [25]. Because of the log link function, the coefficients of the resulting model can be interpreted as parameters in an exponential function. Fixed predictors in the model were the number of cubes in the task, the misalignment between the environment and map pictures, the included orientation aid (with baseline as a reference group), and the task number (to indicate learning as the experiment progresses).

III. RESULTS

On average, the participants made less than three incorrect responses in the 72 tasks. Incorrect responses were deleted from the dataset, which resulted in 832 individual RT responses from the 12 participants. The mean of the mean RTs of the participants was 11.8 seconds, SD = 7.7. The mean of the fastest participant was 7.4 seconds per task, and the slowest participant took on average 22.3 seconds to complete a task. This difference in individual abilities (some are generally faster than others) was expected and taken into account in the multilevel model. The intraclass correlation coefficient was .37, indicating that there was, on average, a moderate correlation between the RTs of a single participant.

All of the fixed effects included in the multilevel model were statistically significant as seen in Table 1, which displays the model coefficients. These coefficients indicate the effect as a comparison to the first term of the effect variable. For example, the negative coefficient of landmark aid means that compared to the 'no aid' reference group the orientation aid reduced the RTs. The predicted RTs can be calculated as an exponential function of the terms. The predicted average RT (in milliseconds) for the first task with four cubes, no misalignment, and no orientation aid, would be calculated using only the intercept (because the reference groups are zero), and would hence be $e^{9.4}$ = 12209.9 milliseconds (12.2 seconds). For 'you are here' aid, for example, the average RT would be $e^{9.4-0.358}$ = 8450 (8.5 seconds) with four cubes, and $e^{9.4-0.358+0.345}$ = 11932 (12.0 seconds) with ten cubes.

The results indicate that compared to no aid, all of the orientation aids decreased RTs statistically significantly, when controlling for other effects in the model. The effect of 'you are here' marker was the largest of the three, and the effect of the direction symbols the smallest. The change in RTs when adding a 'you are here marker' for example, in an otherwise similar task, is

$$\frac{e^{9.410-0.358}}{e^{9.410}} = 0.699073. \qquad (1)$$

In other words, while holding other effects fixed, adding a 'you are here' marker to a baseline map (no orientation aids) decreases RTs by 1–0.699 ≈ 30%. The corresponding number for landmark highlighting is 26%, and for direction symbols 16%. 'You are here' marker and landmark highlighting decrease map based judgment times more than the direction symbols.

Using the same formula, it is possible to calculate that an increase from four to seven cubes causes RTs to increase 4.8%, while increasing the number of cubes from four to ten causes an increase of 42%. Increasing misalignment to 45° increases RTs by only 7%, but a misalignment of 180° predicts 24% longer RTs when compared to 0° misalignment. These results indicate a clear linear effect of misalignment on map based judgments.

Mean normalised RTs by orientation aid, by the number of cubes, are also displayed in Fig. 3, but it should be noted that the multilevel model calculations offer more precise measure of the effect of orientation than the normalised RTs. However, the visual representation of the RTs restates the results of the multilevel model. What Fig. 3 adds to the multilevel model interpretation, is that it seems that with ten blocks, the helpful effect of symbols vanishes.

TABLE I. COEFFICIENTS OF THE FIXED EFFECTS IN PREDICTING RTS IN MAP BASED JUDGMENTS.

Fixed effect	Coefficient (s.e.)[a]
Intercept	9.410** (0.12)
Task number	–.002** (0.001)
Aid: landmark	–.297** (0.047)
Aid: symbols	–.173** (0.048)
Aid: 'you are here'	–.358** (0.047)
Aid: none	
Misalignment: 190	.219** (0.061)
Misalignment: 135	.206** (0.058)
Misalignment: 90	.101* (0.050)
Misalignment: 45	.068 (0.058)
Misalignment: 0	
Number of cubes: 10	.345** (0.041)
Number of cubes: 7	.074 (0.041)
Number of cubes: 4	

a. s.e. = standard error. *p < .05. **p < .01. N = 12, cases included = 832.

IV. DISCUSSION

The multilevel model supported all five hypotheses (H1–H5) of the study. The increase in RTs as the function of the increase of misalignment (H1) was observed to be relatively linear, suggesting linearity of the mental rotations required to align the map with the environment, a result which is in line with previous results concerning mental rotation and orientation [8][12]. It is possible that the linearity assumption can be replaced with more precise estimates, such as Fitt's law, which has been shown to hold for mental rotation [26], but this investigation would require different a kind of analysis. Further, it is possible that data on non-prototypical angles, would prove interesting, as it has been shown that there is a bias towards perceiving non-prototypical angles as prototypical [27]. Regardless, the effect size of the

 ISBN: 978-1-61208-382-7

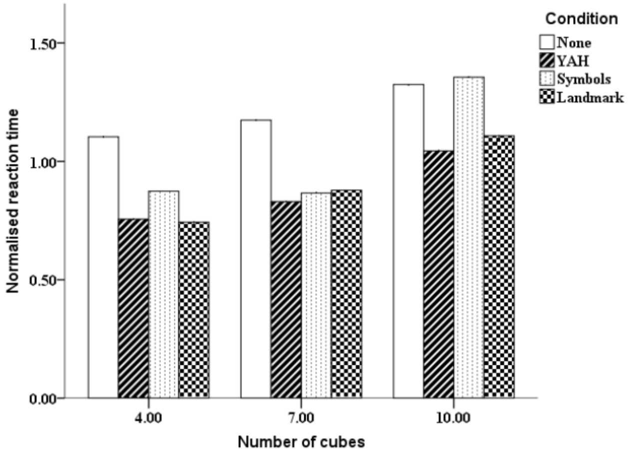


Figure 3. Mean normalised RTs between the orientation aids, by the number of cubes.

misalignment was smaller than the effect size of the number of cubes. This result is somewhat different from earlier similar experiments with reference-frame misaligned maps, e.g., [4]. One possibility is that the orientation was relatively easy, as both the environment and the map had clear, rectangular borders. Future experiments should consider this geometrical detail.

The effect size of the number of blocks (H2) was large, but it seems that the relationship between the number of blocks and RTs is not linear. In fact, almost no increase was observed in RTs when increasing the number of blocks from four to seven. It seems the tasks become more difficult only after seven items. This finding is in line with similar findings concerning search in interrupted visual search tasks [16], and may be at least partly explained by the capacity of the visual short-term memory [9][17]. When a map user is combining information from two sources (the map and the environment), information needs to be rehearsed in visual short-term memory in order to make successful alignment of the differing reference frames [9].

All three orientation aids were confirmed to decrease RTs, giving support to H3–H5. Hence, the laboratory experiment supports the use of the proposed solutions for decreasing judgment times, and possibly errors, when using mobile maps in restricted spaces. Comparing the effect sizes, it seems that landmark highlighting and a ‘you are here’ marker provided better support for map based judgments than direction symbols, especially with larger number of targets. The difference between ‘you are here’ marker and landmark highlighting, on the other hand, was small. Further experimentation should be planned to investigate interaction effects between the orientation aids and the number of items in the small display. Another suggested interaction study should focus on combining different orientation aids together.

The results reported here are also limited to a controlled laboratory setting, which is useful in studying cognitive mechanisms, but not for technology use in the real world. Hence, next steps for ascertaining ecological validity of the results would be to implement the orientation aids on an actual mobile map. Further, users with real-life goals should be utilised in testing the aids. In the laboratory experiment reported here, the participants did not have real-life goal. Both cognitive aspects, studied with experiments, and real-life, goal-oriented action, are necessary for successful designs.

While both ‘you are here’ marker and landmark highlighting have been proposed as orientation aids before, e.g., [7][19][20], the experiment reported here was first to compare these aids, and do so in the context of small displays. The results confirm the usefulness of the aids with varying number of environmental and map items, and varying degrees of reference-frame misalignment. The main results cohere with two important cognitive paradigms, visual rotation [12][13], and a number of important memory and stimulus set size effects such as visual short-term memory and Sternberg-paradigm [17][28][29]. This is in line with the user psychological research strategy of explaining or supporting laboratory user experiments and subsequent design by means of coherence between them and traditional findings of basic research [3].

V. CONCLUSION

In the design process of orientation aids for mobile maps, three orientation aids for improving map-based decisions were evaluated. A laboratory study suggested that having either a ‘you are here’ marker, a landmark highlighting, or canonical direction symbols as an orientation aid decreases RTs in map-based judgments. Another finding suggested that an increase in the number of items on the map quickly increases map based judgment times. Further, there seems to be an interaction effect between the orientation aid and the number of targets on the map: with ten targets, the direction symbols were not as efficient as the other two orientation aids. These findings can be used to design map-based mobile displays for efficient operation within relatively confined spaces, with relatively small number of important environment targets.

The central result of the study reported here is therefore that having either ‘you are here’ marker or landmark highlighting on the map is enough to allow for efficient map use, at least in a confined environment. These markers are not, however, a feasible option if their accuracy is questionable [21], and hence they are recommended only for situations, in which the accuracy can be guaranteed. Compared to ‘you are here’ marker, landmark highlighting offers technologically easier and stable way to aid in map based judgments, and should be considered as a viable alternative. On the other hand, landmark highlighting may not always be possible due to the dynamic nature of environments, such as factory halls. Therefore, while both orientation aids have been proven useful, the context of the use will need to be taken into account when choosing between orientation aids. Further, as individual differences between spatial cognition have been found [30], it is suggested that the choice of orientation aids may at some cases be left to the user.

Increasing the complexity of interaction in already complex environments poses problems for interface design. New features, such as orientation aids in mobile maps may make work tasks more efficient, but may also introduce additional usability problems, interferences, and sources for anxiety to the users. In order to facilitate the use of more complex automated features, new, smart interface solutions need to be conceptualised. The experiment presented in this article demonstrates how the understanding on human cognitive processes can give experimental insight into the evaluation of new design concepts.

ACKNOWLEDGEMENT

We thank research assistants Piia Perälä and Maija Kaibijainen for their help in constructing and conducting the experiment reported here. This research has been supported by the Finnish Funding Agency for Technology and Innovation (TEKES) and the Finnish Metals and Engineering Competence Cluster (FIMECC) programme UXUS (User experience and usability in complex systems).

REFERENCES

[1] T. P. Moran, “An applied psychology of the user,” ACM Computing Surveys (CSUR), vol. 13(1), pp. 1–11, 1981.

[2] P. Saariluoma, “Explanatory frameworks in interaction design,” In Future interaction design, A. Pirhonen, H. Isomäki, P. Roast, P. Saariluoma, Eds. London: Springer, 2004.

[3] P. Saariluoma and A. Oulasvirta, ”User psychology: Re-assessing the boundaries of a discipline,” Psychology, vol. 1(5), pp. 317–328, 2010.

[4] L. Gugerty and J. Brooks, “Reference-frame misalignment and cardinal direction judgments: Group differences and strategies,” Journal of Experimental Psychology: Applied, vol. 10(2), pp. 75–88, 2004.

[5] R. Klatzky, “Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections,” In Spatial cognition, C. Freksa, C. Habel, K. F. Wender, Eds. Berlin: Springer, 1998.

[6] A. J. Aretz and C. D. Wickens, “The mental rotation of map displays,” Human Performance, vol. 5(4), pp. 303–328, 1992.

[7] P. Fröhlich, R. Simon, and L. Baillie, “Mobile spatial interaction,” Personal and Ubiquitous Computing, vol. 13(4), pp. 251–253, 2009.

[8] D. L. Hintzman, C. S O’Dell, and D. R. Arndt, “Orientation in cognitive maps,” Cognitive Psychology, vol 13, pp. 149–206, 1981.

[9] L. Gugerty and W. Rodes, “A cognitive model of strategies for cardinal direction judgments,” Spatial Cognition and Computation, vol. 7(2), pp. 179–212, 2007.

[10] A. Oulasvirta, S. Estlander, and A. Nurminen, “Embodied interaction with a 3D versus 2D mobile map,” Personal and Ubiquitous Computing, vol. 14(4), pp. 303–320, 2009.

[11] W. Seager and D. S. Fraser, “Comparing physical, automatic and manual map rotation for pedestrian navigation,” Proceedings of CHI, pp. 767–776, 2007.

[12] L. Cooper and R. N. Shepard, “Chronometric studies of the rotation of mental images,” In Visual information processing, W. G. Chase, Ed. New York: Academic Press, 1973.

[13] R. N. Shepard and J. Metzler, “Mental rotation of three-dimensional objects,” Science, vol 171, pp. 701–703, 1971.

[14] C. E. Bethell-Fox and R. N. Shepard, “Mental rotation: Effects of stimulus complexity and familiarity,” Journal of Experimental Psychology: Human Perception and Performance, vol. 14(1), pp. 12–23, 1971.

[15] L. Kärkkäinen and J. Laarni, “Designing for small display screens,” Proceedings of NordicCHI, pp. 227–230, 2002.

[16] T. Kujala and P. Saariluoma, “Effects of menu structure and touch screen scrolling style on the variability of glance durations during in-vehicle visual search tasks,” Ergonomics, vol. 54(8), pp. 716–732, 2011.

[17] N. Cowan, “The magical number 4 in short-term memory: A consideration of mental strorage capacity,” Behavioral and Brain Sciences, vol. 24(1), pp. 87–175, 2000.

[18] F. Hermann, G. Bieber, and A. Duesterhoeft, “Egocentric maps on mobile devices,” In Proceedings of Workshop on Mobile Computing, 2003.

[19] R. P. Darken and H. Cevik, “Map usage in virtual environments: Orientation issues,” In Proceedings of IEEE Virtual Reality, pp. 133–140, 1999.

[20] R. P. Darken and B. Peterson, “Spatial orientation, wayfinding, and representation,” In Handbook of virtual environments design, implementation, and applications, K. M. Stanney, Ed. New Jersey: Lawrence Erblaum, 2002.

[21] M. Levine, I. Marchon, and G. Hanley, “The placement and misplacement of you-are-here maps,” Environment and Behavior, vol. 16(2), pp. 139–157, 1984.

[22] M. Levine, “You-are-here maps: Psychological considerations,” Environment and Behavior, vol. 16(2), pp. 139–157, 1982.

[23] S. C. Hirtle and J. Jonides, “Evidence of hierarchies in cognitive maps,” Memory & Cognition, vol. 13(3), pp. 208–217, 1985.

[24] A. May, T. Ross, S. Bayer, and M. Tarkiainen, “Pedestrian navigation aids: Information reqirements and design implications,” Personal and Ubiquitous Computing, vol. 7(6), pp. 331–338, 2003.

[25] G. Verbeke and M. Molenberghs, Linear mixed models for longitudinal data, New York: Springer, 2000.

[26] A. P. Georgopoulos and G. Pellizzer, “The mental and the neural: Psychological and neural studies of mental rotation and memory scanning,” Neuropsychologia, vol. 33(11), pp. 1531–1547, 1995.

[27] N. S. Newcombe and J. Huttenlocher, Making space: The development of spatial representation and reasoning, Cambridge, Mass.: The MIT Press, 2000.

[28] S. Stenberg, “Memory-scanning: Mental processes revealed by reaction-time experiments,” American Scientist, vol. 54, pp. 421–457, 1969.

[29] W. Schneider and R. M. Shiffrin, “Controlled and automatic human information processing: I. Detection, search, and attention,” Psychological Review, vol. 84(1), pp. 1–66, 1977.

[30] D. Waller, “Individual differences in spatial learning from computer-simulated environments,” Journal of Experimental Psychology: Applied, vol. 6(4), pp. 307–321, 2000.

Understanding Map Operations in Location-based Surveys

G. Batinov, M. Rusch, T. Meng,
K. Whitney, T. Patanasakpinyo, and L. Miller
Department of Computer Science
Iowa State University
Ames, IA 50011
e-mail: lmiller@iastate.edu

Sarah Nusser
Department of Statistics
Iowa State University
Ames, IA 50011

***Abstract*—Location-based surveys have been moving to handheld computing devices as the availability of such devices has become more common. The more limited screen size of the handheld devices has made the maps more difficult to use. The present work looks at the map operations of users to determine if they are having problems. Two studies have been analyzed to get an understanding of the types of patterns that might be used to identify users that are having trouble. The choice of the two studies was to find two studies that were quite different and use one of the studies to find patterns of map operations that would indicate that a user was having problems. The second study could then be used to test the relevance of the patterns in a different implementation of the same task. We have identified patterns of interest using the data from the first study and found that the same patterns were relevant in the second study.**

Keywords—location-based surveys, map operations.

I. INTRODUCTION

As computing devices have become common place, we are seeing more location-based surveys use handheld devices in the field. On many of these handheld devices screen space continues to be limited. As a result, maps in these surveys can be difficult to work with.

The present work focuses on understanding the types of difficulties that field staff have with map operations (e.g., zoom and pan) in such survey instruments. Ultimately, we are interested in whether it is feasible to identify map operation patterns that suggest that a map user is in trouble. To look at this question, we have evaluated the results of two studies that use the same survey task (address verification), but different implementations.

We couldn't find existing research results that directly apply to this problem. The closest work looked at map errors in the context of the sequence of map operations that were used to create a new map. Examples are Lodwick et al. [3] and Haining et al. [2]. More recently, work on map operations have used previous users' work to inform other users. For example, Wong et al. [7] looked at the impact of seeing previous users' map operation footprint in crowd

Shneiderman [6] looks at the notion of the Visual Information Seeking Mantra. The concept is related to the work discussed here in that Shneiderman's approach provides a framework for designing geographic software applications.

Roth [4] provides an overview of map-based primitives that provide the underpinnings of the map operations used in our studies.

The main contribution of this paper is that we were able to identify patterns in the data from Study 1 that suggested that the user was in trouble when he/she was using the map operations (zoom and pan) and verify that the same patterns could be used in the second study in spite of the differences in the way that the software was implemented. We also looked at the different treatments used in the two studies to extent this result over multiple variations of the software implementations. The fact that the two studies used different devices and were conducted in very different environments enhances the second study as a means of validating the patterns. Our tests show that the patterns could be found early enough in sequences of map operations that intervention has the potential to result in significant savings in terms of the number of map operations the user ultimately performed.

The remainder of the paper is divided as follows: Section II briefly reviews the two user studies used in the analysis. The results are presented in Section III. Section IV provides a discussion of the results. Finally, we look at conclusions and future work in Section V.

II. METHODS

A. Overview

The experimental task (address verification) involves comparing a housing unit configuration on the ground with the corresponding information in the map. Possible outcomes are: *1)* the ground situation is correctly reflected in the map requiring no further action; *2)* the map has an error of commission that requires a map spot to be removed; *3)* the map has an error of omission that requires a map spot to be inserted; and *4)* the map has an error in the housing unit location that requires the map spot to be relocated.

The term *scenario* is used to indicate the process of completing the verification of one address. The scenario

 ISBN: 978-1-61208-382-7

type was not significant as one would expect since the bulk of the map operations are used to get to the point that the user is able to view the addresses on the map in the target area.

The next two sub-sections briefly overview the relevant details of the two studies. The maps are based on the US Bureau of Census's Tigerline maps. The *map spots* on the maps are used as identifiers of the location of housing units. For example, the spot labeled 507 in Fig. 2 indicates the current map location of the target address – 507 Astaire Ct.

Beyond having the same survey task the implementations used in the two studies are different. Even within the two studies there are different treatments to be considered.

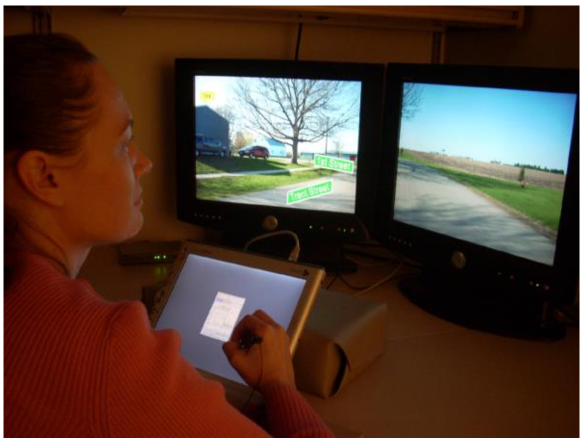


Figure 1: The computer set up used in the Study 1.

B. Study 1

Thirty-five participants were recruited from the community to perform 10 address verification scenarios. The map used in the software instrument covered a city of slightly over 40,000.

The experiment was designed to impose a rigid protocol on the participants. To successfully perform the task for each address, the following steps need to be executed in sequence: 1) find the address on the ground (i.e., in the photos presented to the subject), 2) locate the address on the software map, 3) answer a question posed by the software as to whether or not the address was on the map, 4) if so, answer a question posed by the software as to whether or not the address was in the correct location on the map, and 5) fix the map if an error was identified.

To focus the participants on the software instrument, the participants were seated at a table with two monitors showing the two sides of the street (Fig. 1). The application recorded the time it took participants to perform each step in the procedure, the number of attempts to match each address, the number of attempts to fix the map, the accuracy in fixing the map, and the number of times specific buttons or other software tools were used.

Two treatments were used in the experiment – guided (17 participants) and unguided (18 participants). The screen shown in Fig. 2 illustrates the guided treatment. The guided statements were general statements to indicate the next step in the protocol.

In addition to the rigid protocol another important property of Study 1 is that the map was reset after each completion of a scenario.

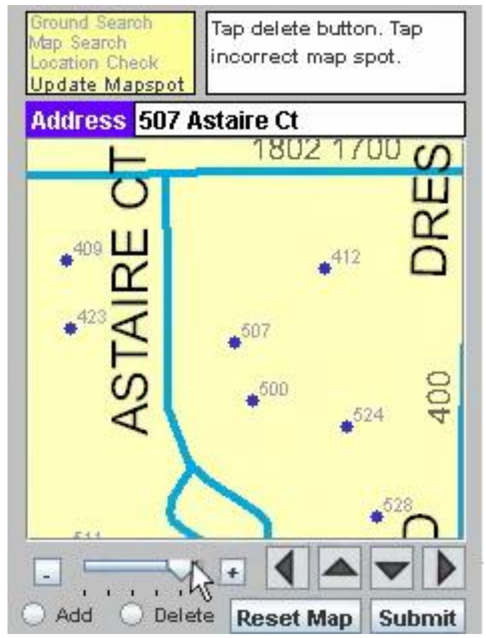


Figure 2. The guided interfaces of Study 1.

A more detailed look at the original user study can be found in Rusch, et al. [5].

C. Study 2

Thirty-one participants performed the address verification task for 6 addresses in the second study.
The second study required to physically navigate the address space. The map of the address space covered a 3X4 block neighborhood of Ames, Iowa. The participants in the study were divided into two treatments (field and virtual reality (VR)). The field group went into an Ames, Iowa neighborhood and had to navigate as well as perform the address verification on the software. The VR group performed the same task, with the exception that the navigation took place within Iowa State University's C6 (a fully immersive virtual reality environment).

A key difference between the two studies is that in Study 2, the participants could choose the scenario they wanted to work on in any order. They could also revisit any scenario at any point in their work. Fig. 3 shows a screen shot of the software showing the scenario menu. Another important difference is that completing a scenario in Study 2 did not automatically reset the map. Rather the map view remained the same until the participant performed another map operation.

A more detailed look at the original user study can be found in Batinov, et al. [1].

IV. RESULTS

A. Overview

Each operation performed by a participant in both studies was logged and time stamped. To look at the map operations, the log files have been parsed to generate the string of map operations. Examples of the parsed results for the two studies are given in Figs. 5 and 6, respectively. The legend for the map operations common to both studies is given in Fig. 4.

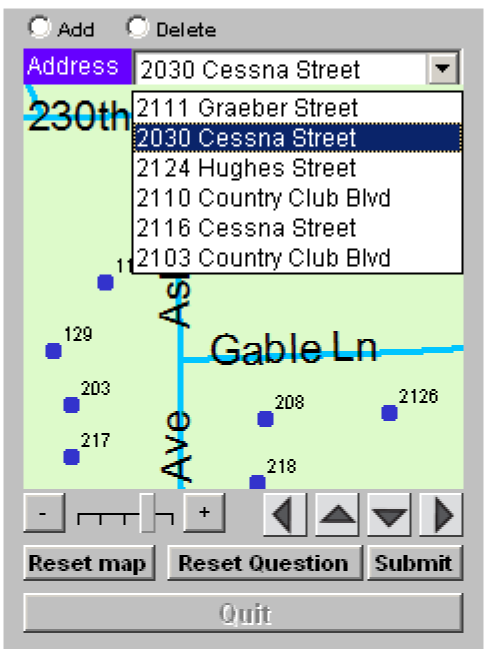


Figure 3. Edit screen with address list extended.

Before looking at the results of our investigation, we need to introduce some terminology that is important to the remainder of the paper. A *count pattern*, denoted n(list of unique map operations) is detected if the map operations in the list appears n times in the scenario sequence. For example, the count pattern 3(A) means that we are looking for the appearance of 3 up pans (A) that appear in the sequence. Note they do not have to be consecutive operations. Looking at line two in Fig. 5, we see that the line contains the count pattern 3(A). Note that it also contains 1(A), 2(A) and 4(A).

A *reversal count pattern* is a count pattern where the map operations in the list represent a reversal. For example, n(+-), n(AV) and n(<>) are reversal count patterns. Since they are only counts, n(AV)=n(VA) for each of the reversals.

The next sub-section looks at results for the first study.

\+ - zoom in by clicking + icon
B - one level zoom in using the scroll bar
C - two level zoom in using the scroll bar
b - one level zoom out using the scroll bar
c - two level zoom out using the scroll bar
\- - zoom out by clicking – icon
x - center zoom click
> - pan right
< - pan left
A - pan up
V - pan down
R - reset map
* - attempted to pan beyond the map borders

Figure 4. Map operation symbols common to the two data sets.

B. Study 1 Results

Fig. 5 shows the map operations for one of the 35 participants. Each line in the data represents the map operations for one scenario.

+x>+xV-R>++><>
+x+xV>-+xA>VAAVV+AVR++><<<VV<<--+>R<>+<+x
<+>VVV-R+xAAVVVVVVAAA+A
+<<R+xVAAAAAR<<<<<>>>>+xA+xA><>
+x+x><-+>-+x>><<-+xVV
+xV+x-+VA
+x+xAA>V<R+x+x>
+xV+xVA
+x+x
+x+x-+VA

Figure 5. Sample map operation data showing one line for each scenario for Study 1.

One obvious type of count pattern that tends to generate extraneous operations is a reversal. Table I shows the number of scenarios (out of 170) that contained at least two or three (n=2 and n=3) reversal count patterns for the guided treatment. Table II provides similar results for count patterns for the individual pan operations with n=2 and n=3.

TABLE I. Number of scenarios in the guided data set that contain the reversal count patterns.

Reversals	Count n=2	Average n=2	Count n=3	Average n=3
n(z+)n(z-)	19	0.112	10	0.059
n(>)n(<)	36	0.212	20	0.118
n(A) n(V)	55	0.324	36	0.212

TABLE II. Number of scenarios in the guided data set that contain pan count patterns of n= 2 and n=3.

Pan Count Patterns	Count n=2	Average n=2	Count n=3	Average n=3
n(>)	52	0.306	30	0.176
n(<)	52	0.306	31	0.182
n(A)	72	0.424	45	0.265
n(V)	65	0.382	49	0.288

Table III shows the number of map operations that participants used after one of the reversal count patterns was encountered at either the n=2 or n=3 levels in the guided treatment. Table IV shows the same results for the pan operators.

TABLE III. Number of the 17 participants in the guided data set that used the reversal count patterns.

Reversals	Count n=2	Average n=2	Count n=3	Average n=3
n(z+)n(z-)	13	0.765	9	0.529
n(>)n(<)	15	0.882	9	0.529
n(A) n(V)	17	1.000	13	0.765

TABLE IV. Number of the 17 participants in the guided data set that contain pan count patterns of n= 2 and n=3.

Pan Count Patterns	Count n=2	Average n=2	Count n=3	Average n=3
n(>)	17	1.000	13	0.765
n(<)	15	0.882	13	0.765
n(A)	17	1.000	14	0.824
n(V)	17	1.000	15	0.882

Table V shows the number of map operations that participants used after one of the pan or reversal count patterns were encountered at either n=2 or n=3 levels. The second value is the number of scenarios that contain one or more of the count patterns.

TABLE V. Number of map operations/impacted scenarios after encountering a reversal or a pan count pattern of size n in the guided data set.

n	pans	reversals	both
2	1130/104	815/72	1194/108
3	886/75	578/46	911/78

Tables VI-X show the same results for the unguided treatment (180 scenarios). Table XI shows the same results for the full data set (35 participants) after the optimal set of map operations have been removed from each scenario. The optimal set of map operations was determined by examining each scenario.

TABLE VI. Number of scenarios in the unguided data set that contain the reversal count patterns.

Reversals	Count n=2	Average n=2	Count n=3	Average n=3
n(z+)n(z-)	33	0.183	19	0.106
n(>)n(<)	53	0.294	33	0.183
n(A) n(V)	58	0.322	44	0.244

TABLE VII. Number of scenarios in the unguided data set that contain pan count patterns of n= 2 and n=3.

Pan Count Patterns	Count n=2	Average n=2	Count n=3	Average n=3
n(>)	66	0.367	45	0.250
n(<)	70	0.389	48	0.267
n(A)	82	0.456	56	0.311
n(V)	68	0.378	51	0.283

TABLE VIII. Number of the 18 participants in the unguided data set that used the reversal count patterns.

Reversals	Count n=2	Average n=2	Count n=3	Average n=3
n(z+)n(z-)	13	0.722	10	0.556
n(>)n(<)	16	0.889	12	0.667
n(A) n(V)	16	0.889	14	0.778

TABLE IX. Number of the 18 participants in the unguided data set that contain pan count patterns of n= 2 and n=3.

Pan Count Patterns	Count n=2	Average n=2	Count n=3	Average n=3
n(>)	17	0.994	17	0.994
n(<)	17	0.994	14	0.778
n(A)	16	0.889	14	0.778
n(V)	17	0.994	16	0.889

TABLE X. Number of map operations/impacted scenarios after encountering reversals or a pan count pattern of size n in the unguided data set.

n	pans	reversals	both
2	1537/111	1167/84	1568/112
3	1239/82	863/59	1261/84

C. Study 2

Fig. 6 shows the map operations for one of the 31 participants in Study 2. The map operations use the same symbols as were shown in Fig. 4. The other new symbols JZ, KZ, LZ, MZ, NZ and PZ indicate the selection of one of the six scenarios used in this study. As can be seen from Fig. 6, participants can open and work on a scenario at any time.

TABLE XI. Number of map operations/impacted scenarios after the optimal set of map operations have been removed.

n	pans	reversals	both
2	2560/199	1838/151	2634/205
3	2034/148	1352/101	2072/153

JZ--A<V><AVV><>Ax
PZ<V<AAVV>cBxBV<V<>AA><><BV><<>V<>
MZ
PZ
MZV<>V<>A<<VA>b>--++>>><<-+V**
NZ<>>><<<<>>><A<<V>>>>*<><A<<V+-
KZ>>><<><>><<<A
NZ<>>-->-++-++--++<V---
LZxA<<>>V
NZ<<A
LZ<>-->>AA<<VVxAVA
JZ>V>>V

Figure 6. Map operations showing one line for each open scenario for Study 2 for one participant.

From Fig. 6, it is clear that the notion of a scenario is not as consistent as it was in Study 1. Moreover, since it is not clear that operations at the end of a line are consistent with the open scenario, combining map operations from multiple lines for the same scenario is not meaningful. As a result, we use each line as a representation of a unit task. Table XII shows the number of lines containing reversal and pan count patterns for the full Study 2 dataset (both VR and field). Tables XIII-XV show the number of the number of map operations that exist beyond the count patterns for the full Study 2 dataset, the VR treatment and the field treatment, respectively. The full dataset contains 286 lines of map operations, while the two treatments (VR and field) consist of 157 and 158 lines, respectively.

Table XII. Line counts for reversal and pan count patterns for n=2 and n=3.

Operations	Count n=2	Average n=2	Count n=3	Average n=3
n(z+)n(z-)	59	0.2063	40	0.1399
n(>)n(<)	93	0.3252	63	0.2203
n(A) n(V)	53	0.1853	28	0.0979
n(>)	116	0.4056	74	0.2587
n(<)	116	0.4056	83	0.2902
n(A)	62	0.2168	36	0.1259
n(V)	72	0.2517	45	0.1573

TABLE XIII. Number of map operations/impacted lines after encountering a reversal or a pan count pattern for the complete dataset and n= 2 and n=3.

n	pans	reversals	both
2	1607/153	1350/124	1703/164
3	1135/103	857/87	1242/113

TABLE XIV. Number of map operations/impacted lines after encountering a reversal or a pan count pattern for the VR treatment dataset and n= 2 and n=3.

n	pans	reversals	both
2	997/88	834/67	1039/91
3	730/58	559/51	782/62

TABLE XV. Number of map operations/impacted lines after encountering a reversal or a pan count pattern for the field treatment dataset and n= 2 and n=3.

n	pans	reversals	both
2	610/65	516/57	664/73
3	405/45	298/36	460/51

D. Comparing Results

To compare the results from the two studies, we used the unpaired t-test with the null hypothesis that the two populations differ. The data drawn from the two studies for this test was the number of map operations that appeared after one of the count patterns was detected. Table XVI shows the values for the count patterns for n=2 and n=3. The value of p in both cases is not significant. As a result, we see the potential map operations saved from two very different implementations as being statistically equivalent.

Table XVI. T-test values comparing the potential savings from the two studies.

Study 1 vs Study 2	t	dff	p
n=2	1.5459	382	0.1230
n=3	1.3928	273	0.1648

We found very similar results when we compared the treatments (guided vs unguided and field vs VR) using the same approach.

V. DISCUSSION

The goal of this study was to use Study 1 to identify interesting count patterns and use the Study 2 data to see whether the same patterns are valid there as well. We have chosen to work with the raw data to provide a view of potential savings that intervening might bring to users of a survey instrument. Note that we are only looking at the potential savings, while realizing that the participant would still have to complete the task.

The expectation is that intervening would give them the opportunity to more efficiently complete the task. Pans in both studies have the side effect of causing some participants to wander. Also note that we are not looking to statistically compare results across studies or treatments. Rather we simply are looking to identify potential count patterns that exist in different implementations. The fact that the implementations of the software, the study environments, and the devices used are very different makes our approach of using the Study 2 data to validate our results more interesting.

A. Study 1

The Study 1 data has been evaluated across the two treatments as well as the complete dataset. From Tables I, II, VI, and VII, we find that the reversal and pan count patterns show up in both treatments. Table V illustrates that for all of the scenarios that contain a count pattern at either n=2 or n=3, there are more than 10.8 map operations per impacted scenario that could potentially be saved for the guided treatment. From Table X we see that there are even more map operations after the count patterns for the unguided treatment (at least 13.2 map operations per scenario impacted). Recognizing the count patterns and intervening gives the potential to significantly reduce the user frustration in map based surveys. This is especially useful in the handheld environment, where small screen size tends to complicate the use of maps.

Tables III, IV, VIII, and IX look at the number of participants that incur at least one of the count patterns. Here we see that over half of the participants in the guided treatment (0.529) and unguided treatment (0.556) have used at least one count pattern. The number of participants for most count patterns is closer to 1.0. Table XI shows the same results for the complete Study 1 dataset after the optimal query has been removed from each scenario string. The idea behind this data was to only consider the extraneous map operations in each line of map operations. Again, we find that the average number of additional map operations in the line to be over 12.1 per scenario.

From these results, we believe that the reversal and pan count patterns for n=2 and n=3 are reasonable choices for determining that a user is having difficulties using the map operations. In the next subsection we look at the impact of these count patterns on the Study 2 dataset.

B. Study 2

As noted earlier, the software implementation for Study 2 provided a more flexible protocol. Two important differences are that the participants could work on any scenario at any time and that the map was not reset at the completion of a scenario as it was in the first study. The first difference resulted in a breakdown in the way that scenario could be used. In the first study, a scenario was essentially the same as a line of data. In the second study a scenario was typically opened on more than one line. Since there is no way to relate operations on an open scenario to the scenario (the participant could be positioning the map for another scenario), the task unit was interpreted as a line of map operations.

The second difference is somewhat more important in the context of this study. Since the map was not reset after the completion of a scenario, most participants in the second study tended to use pans to move on to the next address location on the map. This provides an interesting point, as the optimal approach was to reset the map after completion of a scenario rather than use pans. A third difference is the size of the underlying map. The smaller map for Study 2 should mean less pans, but the pan count pattern numbers are still quite large. In addition to these three differences the two studies differed in the type of device used as well as the environment used for the study.

The result has been that we see the count patterns from Study 1 being useful in Study 2. Looking at Fig. 6, it is easy to see how this one participant tended to wander on the map when he/she was using pan operations. More important, the results in Tables XIII and XV show that the count patterns have been found early enough in the lines of map operations to potentially save participants from using extra pan operations and reversals.

VI. CONCLUSION AND FUTURE WORK

We were able to find an interesting set of map operation count patterns based on pans and reversals in two different implementations of the address verification task. Our next step is to use the count patterns in a new user study where we can intervene and study the actual impact on participants. Ultimately, our goal is to use the map operation count patterns found in this work to provide an adaptive approach to help users struggling with using maps on the mobile devices that agencies like the Bureau of Census are starting to use in the field for large tasks like address verification.

REFERENCES

[1] Batinov, Georgi, Kofi Whitney, Les Miller, Sarah Nusser and Bryan Stanfill. “Evaluating The Impact Of Spatial Ability In Virtual And Real World Environments,” The Sixth International Conference on Advances in Computer-Human Interactions. Nice, France. February 24-March 1, 2013. pp. 274-279.

[2] Haining, Robert and Giuseppe Arbia. “Error Propagation Through Map Operations.” Technometrics, Vol. 35, No. 3 (Aug., 1993), pp. 293-305.

[3] Lodwick, Weldon, William Monson & Larry Svoboda. “Attribute error and sensitivity analysis of map operations in geographical information systems: suitability analysis,” International Journal of Geographical Information Systems. Volume 4, Issue 4, 1990. pages 413-428.

[4]Roth, Robert. "Cartographic Interaction Primitives: Framework and Synthesis," Cartographic Journal. Volume 49, Issue 4 (November 2012), pp. 376-395.

[5] Rusch, M. L., S. M. Nusser, L. L. Miller, G. I. Batinov, and K. C. Whitney. “Spatial Ability and Map-Based Software Applications,” The Fifth International Conference on Advances in Computer-Human Interactions. Valencia, Spain. January 30-February 4, 2012, pp. 35-40.

[6] Shneiderman, B. "The eyes have it: a task by data type taxonomy for information visualizations". IEEE Symposium on Visual Languages. 1996. pp.336-343.

[7] Wong, Yuet Ling, Jieqiong Zhao & Niklas Elmqvist. “Evaluating Social Navigation Visualization in Online Geographic Maps,” Human Computer Interaction. 2014.

 ISBN: 978-1-61208-382-7

Using Crowdsourcing to Improve Accessibility of Geographic Maps on Mobile Devices

Tania Calle-Jimenez
Faculty of Systems Engineering
National Polytechnic School
Quito, Ecuador
tania.calle@epn.edu.ec

Sergio Luján-Mora
Department of Software and Computing Systems
University of Alicante
Alicante, Spain
sergio.lujan@ua.es

***Abstract*— The continuous growth of the use of technology and mobile applications means that more people have access to information published on the web, including geographic information. However, for visually impaired people interaction is difficult if maps are not accessible. For this reason, in this paper we analyze accessibility barriers of webpages with geographic content presented on mobile devices. With the purpose of showing an alternative to improve accessibility in these pages, this study proposes the use of a technique called crowdsourcing, i.e., a group of people that voluntarily access to webpages and provide information about physical accessibility and a general description in each map element (point, line or polygon). This description is written into the Scalable Vector Graphics Tiny (SVG Tiny) code. SVG Tiny is used to represent geographic maps with HTML. In this way, screen readers can interpret the descriptions to visually impaired people, thus making maps more accessible.**

Keywords- Web accessibility, map, crowdsourcing, geographic information, SVG Tiny, mobile devices.

I. INTRODUCTION

The geographic maps are very important in everyday life. They are present in several media such as television, magazines, Internet, and newspapers. The advancement of technology, the increasing number of mobile phone users, and the rapid growth of geographic information systems (GIS), has caused mobile devices to become an essential tool for accessing geo-services for social, professional or personal purposes.

We are witnessing a new era of geographic tools online such as Google Maps, Google Earth, NASA World Wind, OpenStreetMap, MapQuest and Microsoft MapPoint. These tools have a large number of users. For example, in the five largest European economies, 50% of users access online maps from their personal computers, and 35% of users access from their mobile phones on a daily basis [1].

However, not all people can access geographic information on the Web with their mobile devices. Due to the graphical nature of geographic information, some groups of users, such as visually impaired people may experience problems when accessing geographic information.

This study presents an alternative solution to improve accessibility of geographic maps. It uses the crowdsourcing technique and the specification Scalable Vector Graphics Tiny (SVG Tiny) for the implementation of geographic accessible maps.

The next sections are structured as follows. In Section 2, this work reviews the state of the art. In Section 3, it presents definitions of crowdsourcing, SVG Tiny, and web accessibility that guide the readers to understand the use of these concepts in the research. In Section 4, this study shows the accessibility barriers in geographic content. In Section 5, it shows a proposal to implement accessible geographic maps using crowdsourcing and SVG Tiny. Finally, in Section 6, it shows the conclusions of our research.

II. STATE OF ART

Finding a satisfactory alternative that allows visual impaired people to browse geographic information is a very active research field. There are several practical solutions for mobile devices developed by different authors: although some of them are already implemented, most of them are still prototypes. In the following paragraphs, these relevant works are described.

A. *Mobile GIS based on SVG*

Mobile GIS applications are becoming very popular from the last few years. However, the mobile devices used to execute these applications have serious constrains in three areas: screen size, memory and speed. Wu and Bin [2] present a mobile GIS application based on Mobile SVG (Scalable Vector Graphics) for hand held devices with limited resources.

Their case of study is a mobile application for tourism. This application uses mobile SVG as data carrier, display and parser of maps. It introduces constraints on content, such as attribute types, properties, and user agent behavior, due to low memory, low power and limited display. Wu and Bin [2] describes the application data format, presents the map, and points out the areas for future development.

These authors conclude that Mobile GIS can help people with disabilities to move around cities and other places, both outdoors and indoors. For example, blind people can use a mobile GIS to find the directions to arrive at a chosen place. Or a person with a motor impairment (e.g., a person in a wheelchair) can use a mobile GIS to find an accessible route in an airport or a railway station.

B. *Generic Multi-touch Presentation of Accessible Graphics*

Goncu and Marriott [3] present the design and evaluation of a new tool for accessing graphics. Graphics Viewer using Vibration, Interactive Touch, Audio and Speech (GraVVITAS) provides a generic approach for presenting 2-D content. It supports dynamic, interactive use of graphics and could be integrated with existing applications.

GraVVITAS is a multi-modal presentation device. Its core is a touch sensitive tablet PC that tracks the position of the reader's fingers, allowing natural navigation. Haptic feedback is provided by small vibrating motors of the kind used in mobile phones, attached to the fingers and controlled by the tablet PC. This allows the user to determine the position and geometric properties of graphic elements. The tool also provides audio feedback to help the user with navigation and to allow the user to query a graphic element in order to obtain non-geometric information about the element [3].

C. *Touching OpenStreetMap data in Mobile for Low Vision Users*

Kaklanis et al. [4] present an application that enables access to OpenStreetMap data for blind and low vision users using mobile devices. During the exploration, the user moves his finger on the touchscreen of the mobile device and receives feedback vibration when finger is on a road or a point of interest. Sonification and a text-to-speech module provide audio feedback about distance to the next crossroad and information of the road or point of interest [4].

D. *Crowdsourcing techniques for augmenting traditional accessibility maps*

Rice et al. [5] present a contemporary approach to collect and capture geospatial data using crowdsourcing. It reports, locates, and defines transitory obstacles in a built environment. These obstacles represent a significant hazard for visually impaired people when navigating through known and unknown spaces. Efforts like this that allow to quickly report, geolocate, and define transitory obstacles would present a major advance in cartographic support for visually impaired people. The contemporary techniques described in this paper include: gazetteer-based geoparsing, active harvesting of navigational points of interest, and ambient geographic information (AGI) present in social media. These techniques contribute to the characterization of transitory obstacles and facilitate their display in a crowdsourced accessibility system [5].

The papers presented in this section focus on hardware devices and touch and haptic vibration responses. Although it is true that these studies help to improve accessibility of geographic information, many people cannot access these devices.

Our study is a proposal to develop a technique at the software level (code) to describe details of the geographical maps that can be interpreted by screen readers.

III. DEFINITIONS

The following definitions are used in this work:

A. *Crowdsourcing*

Crowdsourcing is a phenomena of 21st century in GIS to generate online information from individual action voluntarily, i.e., a group of people that voluntarily access different webpages and save various kinds of information such as points of interest, addresses, ideas or content.

Crowdsourcing implies collecting large amounts of information and add it on the web through an interface. With this technique, geographic maps with a lot of descriptions of the places within can be obtained.

The mapping through crowdsourcing is usually done by means of a process called Volunteered Geographic Information (VGI) [6]. For this, there are several kinds of software and/or websites that gather information through an algorithm developed specifically for maps. OurMap [7] and WikiCrimes [8] are examples of proposals for data collection using volunteer users to report problems related to cities, crimes and transportation. The information can be loaded manually and/or automatically. For example, OpenStreetMap add new data manually.

Digital Globe Company sponsors the Tomnod mission that utilizes crowdsourcing to identify objects and places in satellite images. They created a web application where thousands of volunteers use satellite images to explore the Earth, solve real-world problems, and view images of the planet. When the Malaysia Airlines MH370 plane disappeared in the ocean, the Tomnod mission developed an application to gather recently collected imagery for any sign of Flight 370 that may have been recorded by a data collection sensor to help identify features, i.e., debris, raft, oil slick, and tag objects that could be useful to find the plane [9].

Mobile GIS Solutions for Crowdsourced Data and Real-time Database Editing is a mobile alert solution. It helps government agencies, utilities and transportation authorities by providing them with a reliable, cost-effective source of actionable information by allowing citizens in their communities to report incidents such as graffiti, illegal trash dumping, potholes, water leaks, broken street lights or signs [10].

B. *Scalable Vector Graphic Tiny*

Scalable Vector Graphic (SVG) has become popular for the development of webpages that contain images. SVG is an open standard defined by the World Wide Web Consortium (W3C) for the representation of vector maps on the web. SVG contains SVG Tiny, which is a profile specially adapted for mobile devices. Although it has many applications, SVG Tiny can be of great help for the design of vector maps presented in navigation systems and geographic information systems (GIS) for mobile devices. Some of the advantages of this format are:

- It is an open standard.
- It is very light because it is based on XML.

- It attaches metadata such as street names, geospatial information, geographic coordinates, RDF, and so on.
- It is scalable, so that it can zoom without deteriorating the quality of map.
- It is easily editable, since it based on XML.
- It can attach animations. This is useful for GPS navigation applications [11].

C. Web Accessibility

Web accessibility means that people with disabilities should be able to make full use of the web. Web accessibility is not interested in the specific conditions of people but on the impact these conditions have on their ability to use the web regardless of the technology used, such as personal computer, tablet, and mobile phone [12].

According to a report published by the United Nations in 2011 [13], more than 1000 million people suffer some form of disability. In this work, we focus on visually impaired. According to estimates by the World Health Organization (WHO), about 285 million people suffer from some form of visual impairment and 39 million people are blind, representing 0.7% of the world population [14].

IV. ACCESSIBILITY BARRIERS IN GEOGRAPHIC CONTENT

In this section we describe several accessibility barriers that mobile users face when accessing webpages with geographic content:

- Low contrast in content.
- Pages saturated of complex information and sometimes unintuitive.
- Contents not intended for use in mobile.
- Movement in the maps: distraction and concentration problems.
- Text represented as image: it means that the text that has an image is in image format.
- Problems with color: color blind people cannot distinguish certain shades of color.
- Mosaic maps: map consists of different images placed in an order so that they form a single image.
- Geographical maps without text.

To achieve web accessibility, we need to be aware of the fact that not all users access the web with the same devices and also not all users are capable of correctly perceiving some kinds of content published on the web. This general idea can be summarized in two basic design principles for web accessibility:

1. Create pages that offer content in different alternative formats.
2. Offer content in an understandable presentation to facilitate navigation through the website.

This study focuses on the second principle and proposes a mechanism to describe elements on geographic maps.

V. PROPOSAL TO IMPLEMENT ACCESSIBLE GEOGRAPHIC MAPS USING CROWDSOURCING AND SVG TINY CODE

At present, there are tools that help visually impaired people to manipulate mobile devices, such as screen readers, screen magnifiers and braille keyboards. These are assistive technologies that certainly help solve accessibility problems to access textual information, but do not help with geographic information such as maps, i.e., screen readers cannot read the maps in detail due to its complex design unless the code contains the appropriate tags.

Also, the crowdsourcing technique can help to input information about physical accessibility characteristics. For example: existence of ramps, elevators with braille signs, escalators, parking for disabled people, etc.

With SVG Tiny code, webpages with geographic information can be implemented and accessed regardless of the user's capabilities. SVG Tiny code includes elements which provide supplementary descriptive information about parent elements. Some specific examples are the <title> and <desc> elements, which can be a child of any graphic or container element in SVG Tiny code, and which contain textual descriptions of the parent element. The <title> tag is meant for a short text description of an element. If the text description is complex, the <desc> tag should be used because it is intended to provide arbitrarily long descriptions (nothing in the SVG Tiny code specification limits the length of these elements). These tags can be read via screen readers. The <desc> and <title> elements are not rendered as part of the SVG Tiny graphics. However, the <title> element can be displayed as a tooltip when the pointing device moves over particular elements. The container element <g> can be used to organize the content of the map at different levels (layers) that can be offered to the user on demand.

The SVG Tiny code can be interpreted by the screen reader for visually impaired users that need to access geographic maps.

The crowdsourcing technique can be used for adding information, i.e., the proposed application can allow volunteers anywhere in the world to access online the map and add physical accessibility information and general description for each of the elements of the map. This information is stored in the SVG Tiny code.

For example, the geographic map shown in Figure 1 is a schematic representation of a university campus with buildings, parking lots, streets and green spaces. The buildings are represented by lowercase letters; the parking spaces with uppercase letters and the green areas with numbers. The goal is to use the crowdsourcing technique to add physical accessibility information and general description to the map. Focusing on the Parking A, when the user positions the cursor on the map, it shows a pop-up window to enter the title and description associated to a specific element.

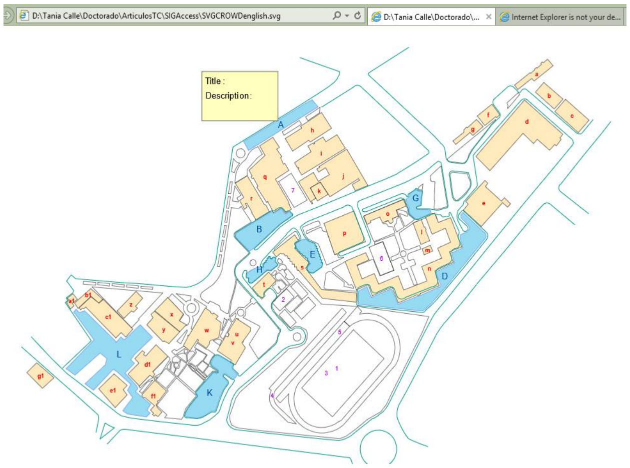


Figure 1. Map without physical accessibility information.

Figure 2 shows an extract of the SVG Tiny code of the geographic map corresponding to Parking A. In this code, there are not elements as description or title.

```
<?xml version="1.0" encoding="utf-8"
standalone="no"?>
<svg xmlns="http://www.w3.org/2000/svg"
version="1.2" baseProfile="tiny">
<g id="Layers">
<path clip-path="url(#SVG_CP_1)" fill="#97DBF2"
fill-rule="evenodd" stroke="none"
d="M856.53969,257.75181L871.41934,274.55127L721.90289,365.74838L697.18347,379.90793L685.66375,358.5486L837.82014,267.3515L856.53969,257.75181z"/>
        <path clip-path="url(#SVG_CP_1)" fill="none"
        stroke="#000000" stroke-width="0.47999" stroke-
        miterlimit="10" stroke-linecap="round" stroke-
        linejoin="round" d="
        M306.47273,710.13743L483.10855,725.01696L422.38999,787.65497L424.06995,789.33491L405.35039,809.01429L306.47273,710.13743
</g> </g>
```

Figure 2. SVG Tiny code without physical accessibility information

Once volunteer users add a title and a description, as shown in Figure 3, it is automatically entered in the SVG Tiny code of the geographic map.

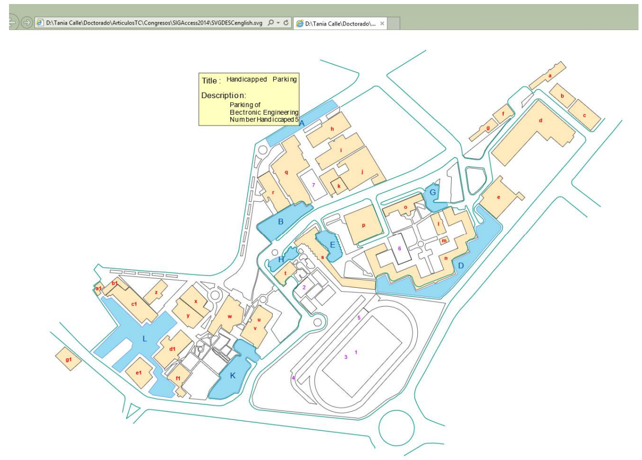


Figure 3. Map with physical accessibility information

In Figure 4, from the code perspective, the tags <title> and <desc> have the following information: "University Campus" and “Parking of Electronic Engineering”. Within this element, another container that describes the <g id = "Handicapped Parking"> containing <title> "Handicapped Parking ", and description <desc> "Parking of Electronic Engineering. Number of handicapped parking 5".

```
<?xml version="1.0" encoding="utf-8"
standalone="no"?>
<svg xmlns="http://www.w3.org/2000/svg"
version="1.2" baseProfile="tiny">
<g id="Layers">
<g id="Parking A">
<title> University Campus</title>
<desc> Parking of Electronic Engineering </desc>
<path clip-path="url(#SVG_CP_1)" fill="#97DBF2"
fill-rule="evenodd" stroke="none"
d="M856.53969,257.75181L871.41934,274.55127L721.90289,365.74838L697.18347,379.90793L685.66375,358.5486L837.82014,267.3515L856.53969,257.75181z"/>
        <g id=" Handicapped Parking ">
        <path clip-path="url(#SVG_CP_1)" fill="none"
        stroke="#000000" stroke-width="0.47999" stroke-
        miterlimit="10" stroke-linecap="round" stroke-
        linejoin="round" d="
        M306.47273,710.13743L483.10855,725.01696L422.38999,787.65497L424.06995,789.33491L405.35039,809.01429L306.47273,710.13743
        <title> Handicapped Parking </title>
        <desc> Parking of Electronic Engineering. Number of
        handicapped parking 5</desc>
        </path>
        </g>
</g> </g>
```

Figure 4. SVG Tiny code with physical accessibility information

Thus, the title and description tags of the SVG Tiny code can be used to provide physical accessibility information associated to the map: a screen reader can retrieve the information of the map and provide it to a visually impaired

 ISBN: 978-1-61208-382-7

user. If all the descriptions of the whole map are entered with the crowdsourcing technique using SVG Tiny code, then this would help visually impaired people know physical accessibility characteristics of places in online geographic maps.

VI. CONCLUSIONS

This study linked concepts to improve accessibility including languages and tools such as crowdsourcing, SVG Tiny code, and use of screen readers. This helped to propose a solution to overcome a barrier of accessibility providing a textual equivalent to visual content.

New applications in the geographic scope, such as OpenStreetMap and Google Maps make clear the need to use online geographic information opposite to the traditional use of printed maps. However, it is necessary to develop methods to provide more accessible information to the user in several levels such as form, scale and detail, especially for the user of mobile devices.

Most of the studies concerning web accessibility have little or no interest in the application of web accessibility guidelines in the development of geographic solutions, since geographic maps are not generally available and are strictly visual. This shows a clear weakness in the development of geographic maps.

Increased social interaction through the web creates behaviours among people, along with this, the need for mechanisms to obtain web users collaboration on a voluntary basis to solve real problems, such as transportation, pollution, safety, and accessibility. This is what the crowdsourcing technique is about.

This paper presented some elements of SVG Tiny code, such as <g>, <desc> and <title> tags, for describing geographic map elements, i.e., polygons, lines and points, that makes possible to interpret the code by a screen reader, so the visually impaired users can manipulate geographic maps.

Although SVG Tiny code was created to represent the images with dynamic content on mobile devices, the SVG Tiny characteristics also offer the possibility to include code that makes geographical maps accessible.

Most of the efforts to improve accessibility of geographic maps are focused on hardware solutions, such as those presented in the state of the art section of this paper. We believe that integrating hardware with software may provide better results in terms of improving accessibility of geographic maps.

In the future, we intend to combine the format SVG Tiny code, indoor navigation systems, augmented reality, crowdsourcing, and web accessibility guidelines for proposing new ways to access and display the geographical maps available on the web.

VII. ACKNOWLEDGEMENT

This work has been partially supported by the research group "Ingeniería Web, Aplicaciones y Desarrollos (IWAD)" of the Universtiy of Alicante.

We would like to appreciate the invaluable help of Sandra Sanchez-Gordon in the review of this work.

REFERENCES

[1] Google. "*What is the economic impact of Geo Services"*. 2013. Oxera. Available online: http://www.oxera.com/Latest-Thinking/Publications/Reports/2013/What-is-the-economic-impact-of-Geo-services.aspx. Access time: 29-01-2015.

[2] B. Wu and Bin X, "Mobile Phone GIS Based on Mobile SVG", *Proceedings Geoscience and Remote Sensing Symposium*, 2005, vol 8, pp.5527-5529, July 2005.

[3] C. Goncu and K. Marriott, "GraVVITAS: Generic Multi-touch Presentation of Accessible Graphics", 2011, Human Computer Interaction, Lisboa-Portugal, pp 30-48.

[4] N. Kaklanis, et al., "Touching OpenStreetMap Data in Mobile Context for the Visually Impaired". *Proceeding CHI 2013 Mobile Accessibility Workshop,* Paris France, April 28, 2013.

[5] M. Rice, et al., "Crodwsourcing Techniques for Aumenting Traditional Accessibility Maps with Transitory Obstracules Information". Cartography and Information Science, Jun 2013.

[6] A. Ballatore, M. Bertolotto and D.C. Wilson, "Geographic knowledge extraction and semantic similarity in OpenStreetMap. Knowledge and Information Systems". 2012, Springer-Verlag, London. DOI:10.1007/s10115-012-0571-0

[7] Gonzales et al., "Representação Aberta e Semântica de Anotações de Incidentes em Mapas Web". 2013, In Anais do IX Simpósio Brasileiro de Sistemas de Informação, p. 1-12.

[8] A. Graves, "Collaborative information on public safety: Potentials and challenges". 2010, *In Web Science Conf. 2010*, Raleigh, NC, USA.

[9] AnyGeo – GIS, Maps, Mobile and Social Location Technology, "CrowdSourcing Imagery to Help Find Malaysia Airlines Flight MH370", 2014, Available online: http://www.tomnod.com/campaign/mh370_indian_ocean/map/1g2x2cy4s. Access time: 29-01-2015.

[10] Direction Magazine, "Mobile GIS Solutions for Crowdsourced Data and Real-time Database Editing", 2014. Available online: http://www.directionsmag.com/webinars/register/mobile-gis-solutions-for-crowdsourced-data-and-real-time-database-editing/401402?utm_source=DM_homepage&utm_medium=web&utm_campaign=401402. Access time: 29-01-2015.

[11] W3C, *"Mobile SVG Profiles: SVG Tiny and SVG Basic", 2009* Available online: http://www.w3.org/WAI/intro/accessibility.php. Access time: 29-01-2015.

[12] W3C, "*Introduction to Web Accessibility"*. 2012. Available online: http://www.w3.org/WAI/intro/accessibility.php. Access time: 29-01-2015.

[13] Word Health Organization, 2013. *Word Report on Disability*. Available online:

http://www.who.int/disabilities/world_report/2011/en/index. Access time: 29-01-2015.

[14] International Agency for the Prevention of Blindness (IAPB). 2010. Technical Report. Available online: http://www.iapb.org/sites/iapb.org/files/State%20of%20the%20World%20Sight_2010.pdf .Access time: 29-01-2015.

[15] T. Calle-Jimenez, S. Sanchez-Gordon and S. Luján-Mora,."*Web accessibility evaluation of massive open online courses on Geographical Information Systems",*, 2014 , *Proc. IEEE*, Pages 680-686, DOI = 10.1109/EDUCON.2014.68261

 ISBN: 978-1-61208-382-7

X Sign Language (xSL) Forum: Considering Deafness as a Language Rather Than an Impairment

Zahen Malla Osman, Jérôme Dupire
Centre d'Etude et de Recherche en Informatique et Communications
Conservatoire National des Arts et Métiers
Paris, France
Email: {zahen.malla_osman, dupire}@cnam.fr

***Abstract*—In this paper, we describe a video based, text free, online forum called xSL Forum. The aim of this work is to provide a tool for signing deaf people, allowing them to communicate using sign language(s). In contrast to the widely used, real time, video chat system, xSL Forum does not require people to be present at the same moment, facing their webcam, in order to communicate. Beyond a digital library, xSL Forum can be used as an asynchronous communication tool and thus, can be useful for variety of other applications such as entertainment, education, or administration.**

Keywords-Deaf; Internet forum; online; video; communication; interaction; sign language.

I. INTRODUCTION

Hearing loss, as a disability, is often misunderstood by the hearing world. It is often an "invisible" disability: a hearing aid may be the only visible indication. Furthermore, many hearing people do not realize that deaf people often have difficulty learning spoken and written languages, as the ability to listen to the sounds of a language is an important part of learning how to read and write. Indeed, in France, for example, an estimated 80% of deaf people are illiterate [1].

In our daily lives, communication, when it is not oral, is mainly based on the written modality: newspapers, public displays, and television. All of which require the ability of reading (Figure 1). The captioning of movies, television shows and news broadcasts is an important advance, but is not useful for all deaf people. Professional environments have the same characteristics, nor is the World Wide Web an exception to this rule: text dominates on the internet and people cannot reasonably be expected to navigate it efficiently without mastering a written language.

A useful means of increasing the accessibility of textual content for deaf people is to propose a signed alternative in the form of a video. As far as we know, this solution is neither automatic nor systematic. If, technically, this is quite achievable, we still find very few sites that deploy this solution. Where it does existing, it is mostly found on specialized sites, initially intended for deaf people. The main obstacle to this alternative modality is the production cost.

So much so that even the dedicated websites or the sites which include deaf people in their audience offer this signed alternative only for their menu options, and switch to text mode (read and/or write) for all or part of their content, including dynamic content (Figure 2). When not in a native sign language form, the contents of the site are translated into a sign language according to the capacity of the structure and the constraints of the production schedule and budget.

Traditionally, websites have these common characteristics: (1) a high dynamicity of contents, and (2) an extensive use of text, which poses a problem for its appropriation by deaf users. These characteristics are also shared by online forums.

This paper is organized as follows: Section II presents motivations behind this work. Section III presents related work. Section IV describes the x Sign Language (xSL) Forum, including the software and hardware architecture, the modes of navigation and interaction, and the evaluations. Section V provides an outlook for future work. Section VI summarizes and concludes our paper.

II. MOTIVATIONS

The online forum has become a major communication tool in the landscape of the modern World Wide Web. It offers many advantages; such as allowing the participation of a wide community in a single conversation, whether for sharing information or seeking a solution; communicating through a forum allows people to instantly broadcast to all users.

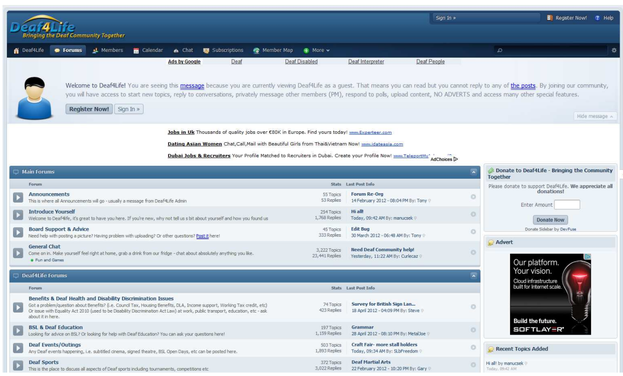


Figure 1. Example of a general forum intended for the deaf community [2]

Asynchronous communication dispenses with the need to be available simultaneously in order to communicate, by passing the constraints of time zone difference in communications between international communities. Finally, the online forum is, by virtue of the persistence of its contents, a perennial source of data.

Although the model of the online forum has proved its value and purely textual access, it is still very difficult to deaf users to benefit from it.

We decided to design a tool that will give access to deaf people to such an online forum, without the difficulties posed by their lack of knowledge of written languages.

III. RELATED WORK

Over the previous decade, the technology sector has proposed ideas towards facilitating communication between deaf and hearing people.

Video chat systems, such as the free software ooVoo, Skype or Live Messenger, allow signing deaf users to communicate remotely, one-to-one or one-to-many, using their webcams. This system compensates for phone use in oral synchronous mode but has the same weaknesses, namely: (1) the need for the participants to be available at the same time; (2) the need to own the equipment, i.e., a computer equipped with a webcam; and (3) a lack of a permanent record of the information exchanged during the communication.

From this starting point, many research projects studied how real-time video systems may be implemented, evaluated them in different contexts and looked at how it such a system could become a useful tool for deaf people.

The Mak-Messenger project [4] has been deployed in the field of education to help deaf students learn sign language. This application allows for the broadcasting of messages in the form of signs between users using an interface similar to a conventional chat. This technology can be used to address the educational needs of deaf people.

The Mobile American Sign Language (ASL) project [5] studied the limits of wireless video communications using mobile phones. This was done to assess the extent of the constraints imposed by mobile devices in terms of size and video quality.

The Learning Management System (LMS) [6] uses video in Greek Sign Language to translate any text in a learning environment. This system is designed specifically for deaf people who want to improve the mastery of a language. It offers a bilingual interface (video and text) and a real-time video chat system.

All of these projects use video for synchronous communication (in real time) between users and do not allow any communication in delayed time.

Researchers at the University of Washington created a project for enabling American Sign Language to flourish in Science, Technology, Engineering, and Management (ASL-STEM forum) [7]. This project provides a space for exchange, for the deaf to refine their comprehension of concepts used in science courses at university and provide a translation into American SL.

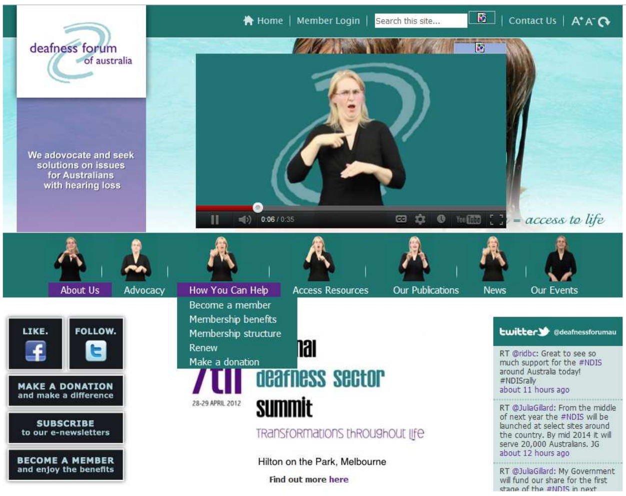


Figure 2. Example of a website with alternative videos [3]

This forum offers a mixed interface, text, and video; which is consistent with the context of use (higher education).

The ASL STEM Forum is primarily intended for the use of tertiary level students, which implies that the users' mastery of reading and writing is high.

This is the only project that includes asynchronous video communication between users. However, it is not designed to be general purpose, nor is it adaptable to another field of interest, like gardening or sports. Moreover, it is designed for a group that has already mastered a written language and wants to discuss the relationship between that and a sign language.

xSL Forum is designed to fill this gap, allowing signing users to discuss a wide range of topics in their native language (sign language), without the need to engage in reading or writing.

IV. XSL FORUM

The main motivation of this work is to offer the whole signing community a communication tool that is usable, accessible, efficient, and up-to-date; one that provides the same easy-of-use and functionality as text-based, online forums. All its content has to be accessible for signing people, especially if they do not know a written language.

"xSL" stands for "x Sign Language". The "x" means that the tool is not linked to a particular SL (Sign Language). It can be used with French SL (FSL), American SL (ASL), British SL (BSL), etc. Thus, all signing people, without any consideration for nationality or location, can use it without any modification for its functionalities.

Most computer applications nowadays are localized: they benefit from translation and some possible adjustments depending on the culture and the language mastered by potential users in a particular part of the world.

Our approach is to consider developing an application originally designed for signing people.

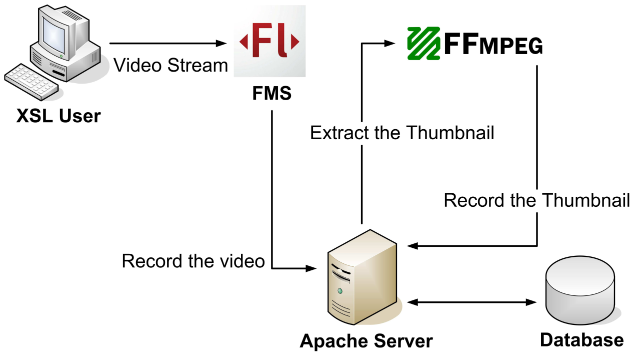


Figure 3. Recording a video (posting a message)

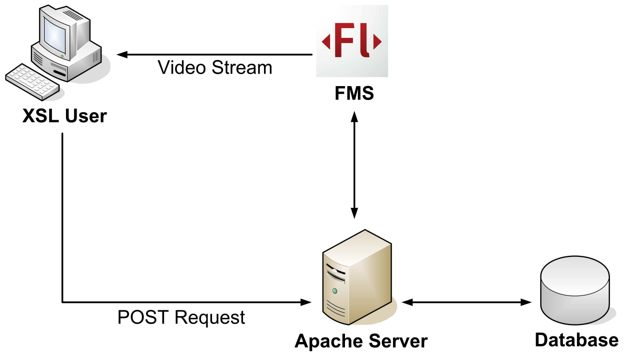


Figure 4. Playing a video (reading a message)

No written text is presented on the pages of the forum and we do not ask the user to enter any kind of text when using the xSL forum.

Forum access is possible even if the signing person does not know how to read or write. Knowledge of sign language (French or otherwise) is the only requirement to use this tool.

The xSL forum is an asynchronous tool. Users simply record a message in video form (a question, for example), and then come back later to see the answer(s). Beyond the constraints mentioned earlier, this feature can become a major advantage in certain contexts, e.g., to compensate for a large time difference between two interlocutors (a French one and a Japanese one).

xSL Forum can also be seen as a digital library with a capacity for long-term memory. Therefore, we can consider its use in specific contexts, such as: education (online courses), or administration (explanation of procedures, individual experiences, frequently asked questions, etc.).

A. *Software and Hardware Architecture*

The hardware architecture is standard for the World Wide Web landscape. It is a server (64bits PC with an Intel Xeon, 2GB of RAM and a 500 GB hard drive), on which are installed: an Apache server with PHP and a MySQL database (Figures 3 and 4). The software architecture consists of three principal parts:

- Flash Media Server
  The video streams are operated by a software component installed on the server side, Adobe Flash Media Server (FMS) [8]. This software supports streaming and recording video streams within web pages on the forum. On the client side, a small module coded in Flash Action Script 3 ensures the communication with FMS. Although Flash is not free and make us dependant on Adobe, we chose to use this technology to the extent that, at the time of development, it was the only solution available that allowed us to test the client behavior, independently of the browser used. Furthermore, many websites are using this technology and many users have already installed the Flash plug-in on their browser. It is a form of standardization in the management of video streams.

  FMS is not free in its standard version, but a free version for developers is available. It is merely limited in number of simultaneous connections from clients. Although the inclusion of FMS in the global architecture was quite complex, this choice allows the administrator/owner of the forum to keep a hand on the video content, in order to manage, for instance, privacy and restricted access to it. It seems that this would not have been fully achievable if we had decided to use another technology such as the YouTube API.

- Flux BB
  The software architecture of the website is derived from FluxBB [9]. It is an open source forum (GNU) developed in PHP, that we modified and added to. Common features have been preserved but all interfaces (GUIs) have been adapted to the specificities of video. The text input interface used for posting messages was replaced by a video recording interface (Figure 5).

- FFmpeg
  This is a library for manipulating video [10], which offers, for example, the ability to convert videos from one format to another. We use FFmpeg to dynamically extract images from videos that are posted on the forum. These images are used as thumbnails, which will be displayed as visual tips to recognize the messages of different users. We decided to extract the 10th picture of a movie and use it as a thumbnail to avoid the issues encountered if an earlier picture was chosen (blank or black ones, due to the streaming process).

B. *Navigation and Interaction*

One of the challenges of this development task was to find solutions for keeping interaction fluid, including exploration of the site contents and ease the manipulation and playback of videos. Interaction is exclusively via the mouse; the keyboard is never required. This feature will make its use on mobile platforms much less restrictive (no need of any keyboard) and more appropriate.

 ISBN: 978-1-61208-382-7

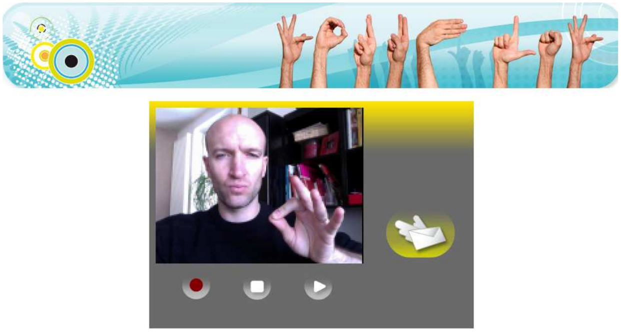


Figure 5. Video record interface

The hierarchy of the xSL Forum consists of three levels: categories, discussions and messages. The home page shows the categories hosted within the xSL Forum. By clicking on the thumbnail of one of these categories, users can access the various discussions (Figure 6). Similarly, by clicking on a thumbnail for a discussion, users can view the various messages exchanged in this discussion (Figure 7).

The pages on xSL forum are designed in such a way as to preserve simplicity while being visually pleasing and intuitive at the same time. Several design choices were implemented to meet these criteria.

First, we systematically combine two images to describe a given piece of content. For example, on a category page (Figure 6), the "category" tab in the top-left consists of an icon and a thumbnail, which reflects the title of the category. This combination allows us to offer two formats, one static and one dynamic, to take into account varying levels of navigation expertise. When users mouse over the thumbnail, the corresponding explanation video is played.

We also implemented a color code to complete the specificity of each level of the hierarchy of the forum: categories are green, topics are orange, and discussion messages are yellow. This encoding can also be retrieved via the tabs presented in the top left of the interface, which effectively completes the double (icon/thumbnail) presentation described above. These tabs allow the user to locate him/her-self through the existing topics and categories.

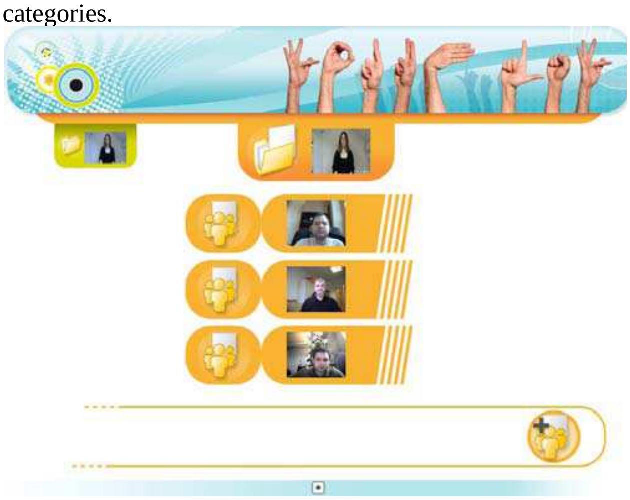


Figure 6. Different discussions in a category

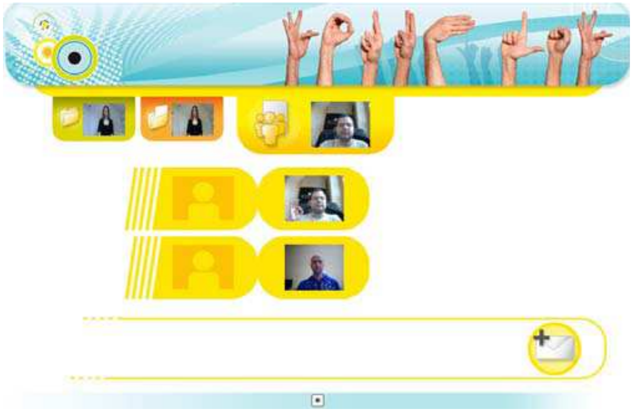


Figure 7. A particular discussion

- Playing videos (reading messages)
  A video (from a topic or a message) is played, when the cursor is placed over its thumbnail. To achieve this, our tool calls a script that executes a JavaScript code into a PHP file, using AJAX asynchronous mechanisms. This code implements a PHP/Flash playback that connects with the FMS to obtain the requested video and broadcast it (Figure 4).
  The requested video then appears next to the thumbnail in a window size of 200×160 pixels (Figure 8). This size was considered the most efficient as a result of an experiment to evaluate different sizes of videos on a web page by deaf signing users. The study was conducted with seven deaf users. They were asked to access a webpage with a 15" laptop and choose which sizes of video frames should be considered wide enough to understand the signed message.
  The available sizes ranged from 75×60 to 200×160 pixels (Figure 9). The comprehension of the message was good or acceptable on the three widest video frames, ranging from 125×100 to 200×160 pixels. However, all users reported that the widest one (200×160 pixels) was the most comfortable. If the user moves the cursor off the thumbnail, the video will stop and the window of the player will disappear (Figure 10).

Figure 8. Cursor over a thumbnail: the video appears (played)

 ISBN: 978-1-61208-382-7

Figure 9. Different sizes of videos were used to evaluate the understanding and comfort of viewing.

Figure 10. Cursor out of thumbnails

- Recording videos (posting messages)
  Our current version allows users to create threads or leave messages, while the administrator controls the creation of new categories. If a user wishes to create a discussion or leave a message, he/she has to click on the icon on the bottom right of the screen (Figures 6 and 7). A new screen appears that implements a video recording interface (made with Flash Action Script 3) (Figure 5), that includes four features: record, stop, play, and validate. The recorder can record Flash videos using the webcam of the user's computer, in conjunction with the FMS server-side module. Once the video is validated, a PHP script is called. FFmpeg will extract a thumbnail, and then the server saves the extracted image and video to the database (Figure 3).

C. Evaluations

Beyond the experiments outlined above, we offered the use of the xSL forum to several deaf people.

Our first observation was that this type of tool (an online discussion forum) is not very well known by among the deaf community. This is consistent with our discussion in the introduction, where we assessed the typical form of online (textual) forums as having a low level of accessibility for this population. Therefore, their use and operation remain unclear.

However, after explaining the operation and purposes, all those contacted expressed great enthusiasm, and considered the many contexts of use of the xSL Forum, its use in the educational context was the most frequently mentioned.

We are about to start a formal evaluation of the xSL Forum. The aim is to evaluate the human-computer interfaces and ergonomics of the navigation and interaction. For this, we are preparing an experiment that consists of showing part of a movie to a set of deaf people and allowing them to discuss it remotely using the xSL Forum. We will prepare a variety of topics for this purpose. After a period of time has passed, questionnaires will be given to the users in order to elicit their feelings, criticisms, and suggestions regarding the tool.

V. Future Work

A new version of the xSL forum is under development. We are working on the following improvements:

Users should be able to access their account securely, so we are looking into authentication tools based on face recognition, such as [11]. The integration of these tools will help us meet our initial constraints, i.e. not to have to handle text.

We are also replacing FMS with the RED5 media server [12], a royalty free audio/video streaming engine. We are also considering the integration of video using HTML5 on the client side, but its implementation in the browser is still limited. This will provide a free forum independent from proprietary technologies; the ultimate being to provide a tool that is open source and royalty free.

We are trying to integrate help videos with interactive icons, to be read in the same way throughout the forum (with the mouse over action).

Video playback will also be improved by providing a "full screen" mode with control buttons (pause, play, timeline). These controls do not exist in the current version and yet they are indispensable, especially in the case of long videos.

We are planning to evaluate this tool with deaf communities. We also intend to test it with people who practice Cued Speech, as it can be used in the same way as sign language in the xSL Forum.

We are also planning to develop and deploy a version of the website for mobile platforms (phones and tablets). A specific application, IOS and Android, will ideally complete the whole set. These kinds of platforms suffers from the need for a real or virtual keyboard. When the keyboard is a physical one, it makes the device bigger; and when it is virtual, it poses serious constraints and ergonomics issues for the user. The xSL Forum is a great laboratory for experimenting with non-verbal interaction, based on visual cues, images, videos and gesture recognition. Initially linked to the deafness context, all the experiments and results could be tested and applied in a more general context, i.e. toward all users (hearing or deaf).

Another important discussion has been raised: the way search engines deal with video indexation and the use case offered by the xSL Forum.

Finally, the aim of this project is for the xSL Forum to be used to bring together different communities, including deaf, hearing-impaired, and hearing people. We, therefore, wish to collaborate with teams working on issues of recognition and automatic translation from sign language into written language and vice versa. This kind of collaboration would lead to a true bilingual access to xSL Forum and total accessibility.

VI. CONCLUSION

As far as we are aware, an online forum using videos as its main modality of communication for signing people -deaf or hearing- did not exist before the xSL Forum. It allows users to simply exchange experiences and knowledge in various fields, according to previously unexploited temporalities in video mode (asynchronous communication and message archiving).

Beyond the important contribution to the deaf community, xSL Forum allows us to devise and propose non-textual interactions for situations in which text is not or is no longer be available. This forum can be used in contexts in which the modality or text is not very usable (mobile devices, embedded, no keyboard, etc.).

ACKNOWLEDGMENT

We would like to thank Laure Giry, Katharine Neil, Lubna Odeh, and all the people who accepted to give feedback on our study.

REFERENCES

[1] D. Gillot, “The rights of deafs”, rapport to the French Prime Minister, 1998 (in French). See http://clas-lyon.fr/spin.php?article116 [retrieved: January, 2015].

[2] Deaf4Life forum: http://deaf4life.co.uk/ [retrieved: January, 2015].

[3] Deafness forum of Australia: http://deafnessforum.org.au/ [retrieved: January, 2015].

[4] J. Ohene-Djan, R. Zimmer, J. Bassett-Corss, A. Mould, and B. Cosh, “Mak-Messenger and Finger-Chat, Communications Technonlogies to Assist in the Teaching of Signed Languages to the Deaf and Hearing”, Proc. of the IEEE International Conference on Advanced Learning Technologies, 2004, pp. 744-746, doi:10.1109/ICALT.2004.1357643.

[5] A. Cavender, R. E. Ladner, and E. A. Riskin, “MobileASL : Intelligibility of sign language video as constrained by mobile phone technology”, Proc. of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (Assets ‘06), 2006, pp. 71-78, doi:10.1145/1168987.1169001.

[6] A. S. Drigas, J. Vrettaros, and D. Kouremenos, “An e-learning management system for the deaf people”, Proc. of the 4th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering Data Bases (AIKED ‘05), 2005, pp. 28:1-28:5.

[7] A. C. Cavender, D. S. Otero, J. P. Bigham, and R. E. Ladner, “Asl-stem forum: enabling sign language to grow through online collaboration”, Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI ‘10), 2010, pp. 2075-2078, doi:10.1145/1753326.1753642.

[8] Flash Media Server: http://www.adobe.com/products/flash-media-server-family.html [retrieved: January, 2015].

[9] Flux BB: http://fluxbb.org/ [retrieved: January, 2015].

[10] FFmpeg: http://ffmpeg.org/ [retrieved: January, 2015].

[11] Pam Face Authentication: https://code.google.com/p/pam-face-authentication/ [retrieved: January, 2015].

[12] RED5 Media Server: http://red5.org/ [retrieved: January, 2015].

Are Current Usabilty Methods Viable for Maritime Operation Systems?

Yushan Pan
Maritime Human Factors Lab &
Department of Informatics,
University of Oslo,
N-0316. Oslo, Norway
yushanp@ifi.uio.no

Sisse Finken
Department of Informatics,
University of Oslo,
N-0316, Oslo, Norway
Finken@ifi.uio.no

Sashidharan Komandur
Maritime Human Factors Lab,
Ålesund, N-6025, Norway
sash.kom@hials.no

Abstract*—Usability is strongly linked to loss of life in many technical and incident reports. Maritime operation systems are sociomaterial systems in which many operators work cooperatively on ship bridges and decks. However, current usability methods focus more on individual interaction. Hence, applying such methods to maritime operation systems leads to several problems. Moreover, a few evaluation methods are hard to duplicate from other research fields owing to various reasons. In this paper, we indicate that maritime operation systems should consider cooperative work for providing a complete picture of interaction issues. In addition, evaluation for maritime operation systems needs deeper understanding of the relationships between human beings and systems. We discuss several usability methods that have been extracted from other close field (e.g., aviation systems, fishing systems, maritime navigation systems, and nuclear power plants) and apply insights from such fields to our case – deep-water anchor handling operation. We assert that usability in maritime domain should be expended as interaction in ecosystems such as the maritime operation system. We suggest that interaction study in maritime operation systems can offer a path to draw and measure a complete picture of maritime operation rather than purely focusing on individual usability issues.

Keywords-Interactions; usability; maritime operations; sociomaterial systems.

I. INTRODUCTION

In recent decades, technological advancement has reshaped the patterns of maritime offshore vessel operations. Engineers design many types of offshore vessels for multiple operational requirements and environments. Ship bridge systems are divided into two different categories—maritime operation systems (after ship bridge) (Figure 1. Maritime operation systems) and maritime navigation systems (Figure 2. Maritime navigation systems). The efficiency and effectiveness of maritime operations hugely influence mariner safety issues [1]. An increasing number of accident reports [2] identify dangerous system design characteristics and interactions among various embedded maritime operation systems as the main reasons for maritime accidents [3]. The usefulness of maritime operation systems is strongly linked to loss of life, significant property damage, or negative effects on the environment [4].

Maritime operation systems are complex systems. Individual operators cannot accomplish maritime operations alone, and cooperation among multiple operators and subsystems are required. The associated operation environments involve greater levels of complexity than a regular office. Maritime work is more or less similar to a society, and it involves not only technical work, but also social communication from task to task [5]. Operators, systems, operational behavior and social communication build a sociomaterial system in maritime operations. IT designers consider such complex sociomaterial systems as an infrastructural setup [6] in which economy, technology, and system stakeholders are involved [7]. However, according to Pomeroy and Jones [8], maritime systems are a combination of human operators, technical elements, and physical equipment. In addition, they point out the necessity of considering sociomaterial systems in the broadest sense when dealing with marine safety. The study presented in this article is limited to usability issues within the scope of maritime operations, and we consider an abridged complex sociomaterial system comprising human beings (operators on the ship bridge and deck operators who work on maritime operation tasks) and maritime operation systems (Figure 3. Ship bridge).


Figure 1(left). Maritime operation systems (Copyright: Kongsberg maritime, Norway); Figure 2. Maritime navigation systems (Copyright: Ocean Industry Concept Lab & Maritime Human Factors Lab, Norway).

Traditionally, usability is not concerned with safety, but with understanding interaction mechanisms and using this understanding to improve design [9]. Usability refers to efficiency, effectiveness, and satisfaction [10], and it is widely used to evaluate web pages[11], mobile information systems [12], and general physical ergonomic issues [13]. Maritime operation systems involve many interrelationships among of multiple subsystems for different maritime operation tasks and challenging work environments. Consequently, they are much more complex than other sociomaterial systems. For example, dynamic positioning systems, drilling systems, alarm systems for operations, and dragging oil and deep water systems are integrated for maritime operations. Operators face many displays/subsystems (Figure 1. Maritime operation systems), and they need to communicate with other operators within and outside the ship bridge (Figure 3. ship bridge). Such distributed collaboration and complex sociomaterial interaction poses interesting questions for researchers—whether current usability methods still make sense for increasingly complex sociomaterial maritime operations, and what are strengths and weakness of current usability methods?

The usability of complex sociomaterial systems is rarely studied in the context of maritime operations. There are a few studies pertaining to usability in the maritime domain, for example, a study on fishing vessels [14] and one on maritime navigation [5]. Nevertheless, these studies focus on physical equipment as opposed to systems. In addition, consideration of usability issues in the design of most large and complex systems in the maritime domain is largely absent, for example, Henique et al. [15] largely neglected cooperative IT work in their research.

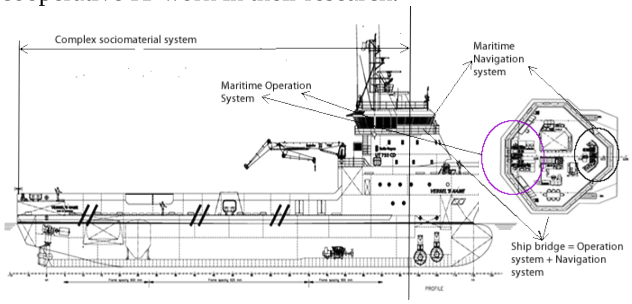


Figure 3. Ship bridge: Complex, sociomaterial maritime system (Copyright: unpublished resource, Maritime Human Factors Lab, no copyright restrictions)

This situation is understandable because we assume the current usability approach to be developed when the personal computer came into use, and it is applied to the Internet and various consumer electronics. Usually, there is no need for user cooperation when interacting with these products, and safety is not the first quality objective. However, maritime operation systems comprise many integrated subsystems with highly complex interrelationships, and, therefore, current usability methods are difficult to employ in such scenarios. However, maritime operation systems share strong similarities with other research domains such as aviation systems, maritime navigation, fishing systems, maritime rescue and coordination centers, and nuclear power plants [16]. Hence, it would be fruitful to borrow knowledge from other research domains to understand usability in highly cooperative work and complex systems.

Hornbæk [17] suggests that usability measures should be formulated on the macro, as well as the micro levels. In this manner, researchers can capture a global usability picture of complex systems. Thus, in the following text, we draw on the presentations described above to examine current research on usability in different disciplines. This consideration is necessarily both practical and theoretical because we aim at examining how current usability methods can be used practically at a micro level (individual and group usability testing) and the extent of theoretical understanding of usability for maritime operation systems at the macro level (a global usability understanding of entire maritime operation systems). While, it is still difficult for researchers to understand the relationships between humans and complex, sociomaterial maritime systems and their combinations from the two levels. Maritime operation system as an entire system for cooperatively operating by operators, it is important to obtain a picture of usability relations between each sub maritime task, which executed by each individual operator. Hence, interaction mechanism is a clear choice that can offer an opportunity to understand such relations, contrasts, problems, and opportunities in the maritime IT domain, e.g., the relations of several usability problems in one maritime task.

Section II we present the method we use in this state-of-art. Section III presents some cases within and outside the maritime domain. We apply current usability methods to our case of deep-water anchor handling in section IV. In section V we conclude that usability in maritime domain should be considered as interaction in complex environment.

II. METHOD

Getting a clearer idea of how to apply current usability methods to maritime operation systems is indispensable. Additionally, deliberating usability of entire maritime operation systems in a theoretical way is urgent because it would help ascertain unnecessary usability methods before the empirical studies.

Hence, we apply four research criteria to the research domain of usability evaluation in complex sociomaterial systems. Usability, ship bridge systems, maritime system, and sociomaterial system are the search keywords. We have searched databases such as ACM and IEEE Xplore Digital Library, Journal of Usability Studies, Marine Science and Technology Journal, Springer link, and Computer–Human Interaction Journal and Proceedings. Our goal is to map out the current usability methods used in maritime operation systems. Thus, by understanding usability methods from both other research fields and the maritime domain, we seek to find a way of promoting usability methods in the maritime industry. The paper includes a review of usability evaluation from the earliest to the most recent research on complex sociomaterial systems in control rooms, including aviation system, maritime navigation systems, fishing systems, maritime alarm systems, maritime rescue and coordination centers, and nuclear power plants. All work in these domains is highly cooperative among the operators, and several complex support systems are involved. In addition, in all these domains, the potential threats to human life and the environment in cases of abnormal system behavior are great.

III. CASES WITHIN AND OUTSIDE OF MARITIME DOMAIN

We chose the following cases because they cover most usability methods in terms of micro usability analysis. Moreover, the newest usability method—systems usability—attempts to understand systems at the macro level.

As complex sociomaterial systems, aviation systems attract considerable attention in this field. Mahemoff et al.

[19] evaluated an aviation system and described a pattern-based usability approach that was adapted from Mahemoff and Johnston's research [19]. They proposed that usability patterns should be robust; task–efficient; and effective in terms of user–computer communication, comprehensibility, and flexibility. This study concluded that heuristic evaluation, 'think aloud', and cognitive walkthrough are appropriate methods for a complex sociomaterial system and applied these methods to design an example alarm control system for flights. On similar lines, studies have been conducted for usability evaluation of systems such as the UK's air defense control [20], industrial process control [21][22], healthcare [23], aviation and space [24], transportation [25], and nuclear power plants [26].

In the maritime domain, usability is not the first priority of research. Most studies in this domain focus on human-centered design, with limited or no evaluation of system segments. In terms of usability in ship bridge systems, the most notable suggestions have been to design a useful ship bridge, navigation system, maritime mobile application, and so on. For example, researchers have used eye-tracking data to analyze usability issues on bridge systems [27]. Lützhöft et al. [28] conducted a series of navigation systems studies based on observation, interviews, questionnaires, and video recordings. They proposed that in a navigation system, large, shared interactive work surfaces could ensure good support for cooperative work planning and execution. In a parallel study, they present an application based on this type of a shared system to explore the potential of tabletops for maritime navigation.

In a study of maritime rescue and coordination centers, Mills [1] claimed that approximately 300 distress calls from both text entries and voice systems are sent in error every year; that is, the vessels sending these calls are not distressed but they did send emergency messages or distress alerts (e.g., 'Mayday'). Through heuristic evaluation and think aloud analysis, Mills found that the operators did not understand marine operation systems and made wrong system operation decisions. The most obvious fault was false alarms in the ships' systems [1]. In a follow-up study, Mills [14] discussed the usability problems of acoustic fishing aids on small fishing vessels, with a focus on data interpretation and comprehension. Through sequential heuristic analysis, the study pointed out that many operation errors occur because of the poor usability of interfaces. Also, the study found that errors occur because operators do not correctly understand the presented information [14]. Wilkinson [29] stated that improving the usability of user interfaces could help operators understand presented information and convert it into a correct decision or control action within a maritime setting. However, both Mills and Wilkinson did not elaborate on methods of improving the usability of these operation systems.

To understand the interaction mechanisms, human activity, and how users live with technology, Savioja et al. [30] conducted a study of nuclear power plants. They developed a method called 'systems usability'. This method builds on the activity theory [31]. Through this approach, researchers can understand and analyze different levels of operations and actions of individual users [30]. In practice, this method uses a predefined task. The evaluators observe the completion of this task to find certain measures such as errors and completion time. In addition, situation awareness is used to evaluate the user's performance. In a follow-up study, these researchers explained that they used the activity theory in complex systems evaluations for the following reasons [31]:

> …. Activity is understood as historically and culturally developed. Hence, the central aim in the analysis is to find out the current state of affairs but also their historical roots and possible trends from which development is proceeding. The approach suits well the needs of control room evaluation in a state in which hybrid technologies have been implemented and more profound modernizations are under design. In order to understand whether the development of tools is proceeding in a good direction, the wider historical context tools must be understood. (p. 259)

Two units of activity [32] are used for analyzing the work in nuclear power plants —object-oriented and mediated. By means of object-oriented activities, the systems-usability method analyses work execution sequences, way of acting, and experience in action. By means of mediated activities, this approach analyses the relationship between a subject and an object when mediated by tools (e.g. in the NPP studies, the user interfaces are the tools). The author's logic behind this approach lies in the elaboration of on mediation by distinguishing between two different functions of tools in an activity—instrumental and psychological. Another type of activity used is communicative [31]. Through these three functions (instrumental, psychological, and communicative) and object-oriented activities, the systems usability approach tries to cover a system's overall meaningful role in an activity, such as the manner in which humans conduct themselves in human–technology interactions and the global, society-defined purposes and objectives of a user's different task levels.

IV. APPLYING CURRENT USABILITY METHODS TO MARITIME OPERATION SYSTEMS

To examine how current usability methods could help researchers in the maritime domain, we apply these methods from abovementioned various research domains to our example of deep-water anchor handling operations (DWAHO). In DWAHO, two groups of operators on two vessels operate two maritime operation systems during anchor-handling tasks. An additional anchor-handling vessel (AHV)—the secondary AHV—is used to relieve some of the chain weight held by the main AHV (Orange unit, Figure 4. Deep-water anchor handling operations). In this operation, two systems (on two vessels) perform one shared task (positioning an oil platform, shown in orange in Figure 4). The main AHV does more than simply holding the chain. Before holding the chain, the main AHV must follow several procedures, including, among others, drawing anchoring arrangements, offshore installation draught during anchor handling, and measuring water depth [33].

The operators of the main AHV perform different roles during operation. For example, two or more operators conduct the different procedures. All operations use different subsystems simultaneously (Figure 4. Deep-water anchor handling operations), such as dynamic positioning systems, drilling systems, alarm systems for operations, and dragging oil and deep-water systems.

A. *Micro level of usability evaluation*

In an AHV, the interactions are not static. The combination of operators, operation systems, and ship deck operators change from task to task. Barrett [34] proposed that boundary relations such as boundary cooperation and strain, too, change from task to task. To test the usability of AHV systems, for each maritime task, we should consider a group of operators rather than an individual operator. Otherwise, the test would be too narrow and limited to determine the global usability of entire maritime operation systems.

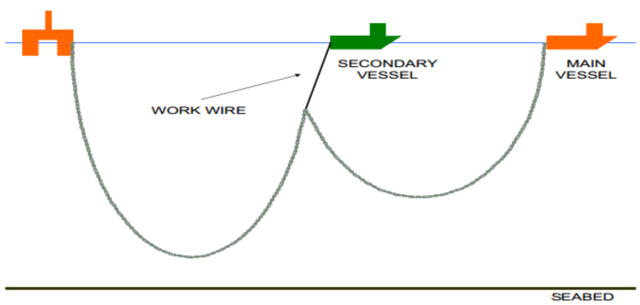


Figure 4. Main AHV is assisted by secondary AHV during DWAHO[33]

However, in aviation system studies, Mahemoff et al [19] did not discuss the mechanism of interaction or the evidence gathered for the method that can best contribute to knowledge about general issues of collaborative work and the potential role of technology and users. Similarly, Lützhöft et al [28] discussed design for marine collaborative settings in their maritime cooperative application study but with minimal stress on efficiency, effectiveness, and accuracy. Even though Lützhöft et al [28] focused on the interaction mechanisms, the usability approach was more or less neglected. The NPP-based studies [30][31] do not clarify the definition of a boundary when an entire environment allows for dynamic changes in groups, systems, subsystems, and combinations of these.

Anchor handling operation systems are special because they come from shipyards, thus requiring manufacturers to deliver assembled solutions, and these solutions are usually needed at different levels of integration, ranging from the physical proximity of equipment to full-scale data-level integration [35]. Furthermore, current system products are a an alternative manifestation [36] of component software elements. An evaluation of a different context of technology-in-use [37] in the maritime domain requires an approach different from merely replicating methods meant for other research domains, where most systems are not developed by decomposition.

B. *Current usability methods in system fragments*

Current usability methods mostly are conducted at the individual level. Nevertheless, there exist differences at micro levels of usability evaluation between other complex sociomaterial systems and anchor handling operation systems in terms of their natural work environment context. Anchor handling operation environments are usually extremely unstable. Wind and stormy waves invariably affect usability experts evaluating maritime operation systems, which increases the difficulty involved in evaluating these systems. Therefore, usability experts are required to account for natural environmental factors that may affect their evaluation outcomes. In addition, the data collection methods used in another domain's study cannot be duplicated directly because of the following reasons:

- The observation of maritime operations is different from the observation of other research domains. Most anchor handling operators may have undertaken only several intensive training courses before they move onboard for hands-on learning [38]. Generally, experienced operators assist the new operators in performing their duties. In contrast, operators from other research domains are trained for several months before they operate systems independently [39]. Thus, observation processes vary, and errors are easier to spot in maritime operations especially when operators need training/supervision for completing a task. However, as outside observers, usability experts should play a role as participant observers to develop a descriptive understanding of the way of life of the study group [40]. Whether the interface, computer screen, procedures, and analogue indicators are simply very usable or seemingly smooth in operation should be judged not by the standards of usability experts but by the participants.
- Most other studies are conducted in a simulator environment, and their usability evaluation is based on a new simulator in addition to the data obtained by methods of questionnaire and interview. In this context, it may be difficult for experienced operators to express realistic problems. However, the interfaces of anchor handling operation systems are computer-supported tools used in a real environment. Thus, it could be easier for an experienced operator to provide detailed descriptions of a real working place, as well as conversations between researchers and their participants [41]. Therefore, from the productivity viewpoint, it would be better if researchers developed an evaluation method for maritime IT designers, who decide on future uses for maritime new products. Again, environments of other research domains are steady compared with a ship bridge. Normally, maritime operation systems are operated in unstable working conditions at sea. In addition, bridge operators have to communicate with other operators on deck or other vessels and oil platforms from time to time. Therefore, a greater number of recorders and usability evaluators would be required compared with those needed in the systems-usability method.

- Eye tracking can help researchers identify obvious usability problems, but eye-tracking data can only tell the interests of users as opposed to explaining why users pay attention to some information and ignore other items. Usability experts cannot go deeper into the user's mind. It is difficult for these experts to capture personal meaning [31], so it would be difficult for researchers to evaluate user perceptions.
- In addition, think aloud, cognitive walkthrough, and heuristic evaluations are difficult to conduct in a situation in which the dynamics of the process largely determine the pace of the situation [31][16]. Nonetheless, knowledge and expertise are embodied inherently in complex sociotechnical systems [6]. Think aloud has proved useful for identifying usability issue–related collaborative problems in mobile collaborative systems within the maritime domain [42]. We believe it depends on the manner in which researchers conduct interviews and ask questions [41] in their conversations with maritime operators.

C. *Macro level of usability evaluation*

From the interaction viewpoint, in nuclear power plants studies, the understanding of 'object of activity' continues to be focused on individual work. This can be understood as a micro level of usability evaluation [30][31]. In the action–objective relationship, the focus is on the interactions between individual users and computer systems. The macro level of usability is evaluated using instrumental, psychological, and communicative analysis methods to investigate the manner in which a user interacts with user interfaces and environments in groupware. These ideas are explained using the term 'communicative' [31].

However, it is impossible to understand usability over an entire environment because a few aspects emerge in unfolding activities that cannot be understood in the absolute sense but only as 'relative activities' [43]. Evaluation of technologies-in-use cannot be carried out in a black box. Maritime operation is usually not only a collection of each individual operator's work and its reflections on group work but also cooperative work among individual operators through interactions. For example, multiple operators manipulate different systems on the ship bridge in anchor-handling operations. Moreover, these operators communicate with different operators on the ship deck during the same or different tasks at the same or different time. Different operators within and outside the ship bridge can be seen as different groups. Furthermore, maritime operation reflects different relations across each boundary between groups, including systems, operators, and social contexts evolving within these operational processes.

Hence, when considering systems usability in the maritime domain, we overlook group interactions and cross relations among different working groups. Schultze and Orlikowski [44] studied the performance perspective in globally distributed, immersive work and argued that in a set of activities and interactions engaged in by various actors, the actors' relationships and mobilisations produce certain effects, such as actions and interactions that are no longer independent but deeply connected and grounded. We concur with this understanding and want to apply it to the maritime domain. We cannot see each maritime operation task as already defined and fixed because there are some invisible operators present in each maritime operation, and cooperation across operator groups is dynamic. For example, in an anchor-handling operation, deck operators assist an operator on the ship bridge; however, in another task, these deck operators may play another role in different operational processes with another operator on the ship bridge. We should treat the group interactions on the ship bridge in terms of the performativity [18] of the systems engaged in the operators' practices. The performativity of maritime operation systems is sociomaterial [18][45], as established by different maritime tasks in which systems are designed and engaged in social practice of operators.

As sociomaterial practices, maritime operation on the ship bridge is a significant source of information about why, when, where, and how maritime operators interact with the anchor handling system and communicate with which operator on the ship deck. This information is much complex than an experimental result that normally takes place in a laboratory. Researchers have no capability to grab the relationships between systems, operators and their combinations. Moreover, only a piece of system can be studied in a laboratory. There is a huge gap between reality and virtual environment. Keeping these questions in mind, operators within and outside ship bridges cooperating in each specific maritime operation task could help understand those invisible activities. The usability measurement idea then moves and reconfigures the dynamic and transformed relationships among interaction issues within and across groups. Such an idea also extends and intensifies usability measurements within maritime operations as an interaction, which would be very helpful for designers, who could then clarify dynamic boundaries for groups, systems, and the associated social context.

D. *From usability problems to interaction in ecosystems of maritime operations*

After applying usability methods to a maritime case, we realize that current usability methods are insufficient for evaluating maritime operation systems. We assert that the term of usability makes less sense in maritime operation systems. The reason is that usability focuses more on individuals. Logic relations of each individual usability problem in a complex environment of maritime operations are overlooked. The relations between individual operators, system segments, and their combined interactions could not be fully understood in maritime operation systems. Hence, usability methods have little power to illustrate such relationship. In this case, researchers inadvertently lose focus on the entire maritime operating systems but pay more concentration on individuals, systems segments without deeply touching the interaction between operators and systems in maritime operations. On this occasion, usability problems should be considered as interaction issues in an ecosystem [43] whereby operators live in their working contexts to reveal interaction issues. In this manner, a usability issue is not standalone but instead integrated into an interaction. From an interaction viewpoint, we can gain a complete picture of the maritime operation system.

V. CONCLUDING REMARKS

Although we mapped the drawbacks of the abovementioned methods in evaluating complex systems, it does not mean that current usability evaluations are not useful; in fact, they may provide some useful data for understanding interaction in empirical investigation, such as observation and interview. To avoid the limitations of individual methods, Neale et al [46] recommend that a multi-method evaluation approach, including observation among other methods, be used in an environment in which multiple users use systems in a work setting. Ship bridge systems are typical complex sociomaterial systems, so it is necessary to develop new usability measures for this special context-in-use. The measures should not be restricted only to the micro level, covering only usability for individuals or group operators for one specific task, as is the case with most studies thus far. Instead, they should determine how specific technologies-in-use enact different sociomaterial environments. This is important in every maritime operation to ensure holistic evaluation of maritime operation systems.

Based on the understanding of dynamic relationships in and across group interactions, in the future, what we call a "usability issues" could be a construed as an 'interaction' through the identification of boundary relations among different working groups and working environments. This should be treated as fundamental 'interaction network [47]' from the maritime operation systems viewpoint, where human beings and their working environments are investigated holistically. To achieve these assumptions, empirical studies on ship bridges that further investigate the relationship between interaction issues are needed. Hence, we will collect research data aboard in the future and to investigate the interactions as our second step of this project.

To conclude, we suggest the following tips to help practitioners who plan to undertake interaction mechanism for maritime operation systems:

- When studying complex systems, concluding usability results for an entire system should not only test segments of the system. Mapping out usability relations of systems' segments is important. Hence, the focus should be shifted to interaction in ecosystems.
- Understanding maritime work of interaction in a complex system is important and useful for outlining interaction relationships in a big system picture. The work in maritime operation systems may involve multiple tasks and operators. The relationship of operators, tasks, different operator groups, and their combination should be understood as a whole.
- Humans and technology should be considered together. The relationship between human operators and technology is strongly connected. Understanding the ecology of this relationship will render a better understanding of interaction relations in interactions of series work as well as a complete system picture. Although we use sociomaterial practice [18] to interpret the relationships between humans and technology, other theories also can serve such analysis, e.g., actor-network theory (ANT) [48]. For example, ANT may better explain the network of operations in complex systems.

The future work of this project is to collect data on ship bridge rather than perform experiments in a laboratory. Our purpose is through interpreting the interaction in maritime operation systems to measure the current design of maritime systems in industries. In turn, we aim to use data analysis results to develop a way for the future evaluation of designing maritime systems.

ACKNOWLEDGEMENT

We would like to thank all members of the DESIGN group at the University of Oslo and Maritime Human Factors Lab at Ålesund for their guidance and support. Our thanks also go to the 37th Information Systems Research Seminar in Scandinavia for suggestions and constructive comments. Last, we would like to thank all anonymous reviewers for invaluable comments.

REFERENCES

[1] S. Mills, "To Live or Drown: When Information Systems become Critical," *Comput. J.*, vol. 38, no. 6, Aug. 1995, pp. 413–417.

[2] National Imagery and Mapping Agency, "Chapter 28 Maritime Safety Systems," United States Department of Defense, Technical, 2014.

[3] N. A. Stanton, P. Salmon, D. Jenkins, and G. Walker, *Human Factors in the Design and Evaluation of Central Control Room Operations*. CRC Press, 2009.

[4] J. C. Knight, "Safety critical systems: challenges and directions," in *Software Engineering, 2002. ICSE 2002. Proceedings of the 24rd International Conference on*, 2002, pp. 547–550.

[5] M. Lützhöft, "The technology is great when it works: Maritime technology and human integration on the ship's bridge," University of Linköping, 2004, Linköping, Sweden.

[6] R. Kling and R. Lamb, "IT and Organizational Change in Digital Economies: A Socio-Technical Approach," *Computers and Society*, Sep, 1999, pp. 17–25.

[7] E. Mumford, "A Socio-Technical Approach to Systems Design," *Requir. Eng.*, vol. 5, no. 2, Sep. 2000, pp. 125–133.

[8] R. Pomeroy and B. Jones, "Managing the human element in modern ship design and operation," in *Human factors in ship design and operation*, October 2012, London.

[9] J. M. Carrol, "Human-computer interaction: Psychology as a Science of Design," *Annu. Rev. Psychol.*, vol. 48, Feb. 1997, pp. 61–83.

[10] ISO 9241 -11, "Ergonomic requirements for office work with visual display terminals (VDTs) - part 11: guidance on usability." ISO, 1998.

[11] K. Pernice and J. Nielsen, "Usability Guidelines for Accessible Web Design."

[12] R. Budiu, “Usability Testing for Mobile Is Easy.” http://www.nngroup.com/articles/mobile-usability-testing/ access date: 06.01.2015
[13] K. Grøndal, “Introduction to: Usability, Ergonomics, Human Factors, MMI...,” 2009.
[14] S. Mills, “Usability problems of acoustical fishing displays,” *Displays*, vol. 16, no. 3, 1995, pp. 115–121,
[15] E. Henique, S. Lindegaard, and B. Hunt, “Design of Large and Complex Display Systems,” in *Handbook of Control Room Design and Ergonomics*, 0 vols., CRC Press, 2008, pp. 83–138.
[16] Y. Pan, S. Komandur, S. Finken, and H. P. Hildre, “An analytical review of usability evaluation for ship bridge multi-screen complex systems,” in *Human factors in ship design and operation*, London, 2014, pp. 109-114.
[17] K. Hornbæk, “Current practice in measuring usability: Challenges to usability studies and research,” *Int. J. Hum.-Comput. Stud.*, vol. 64, no. 2, Feb. 2006, pp. 79–102,
[18] W. J. Orlikowski, “Sociomaterial Practices: Exploring Technology at Work,” *Organ. Stud.*, vol. 28, no. 9, 2009, pp. 1435–1448.
[19] M. J. Mahemoff and L. J. Johnston, “Principles for a usability-oriented pattern language,” in *Computer Human Interaction Conference, 1998. Proceedings. 1998 Australasian*, 1998, pp. 132–139.
[20] E. Hey, “Case Study: Conducting large-scale multi-user user tests on the United Kingdom Air Defence Command and Control system,” *J. Usability Stud.*, vol. 1, no. 3, 2006, pp. 121–135.
[21] F. Nachreiner, P. Nickel, and I. Meyer, “Human factors in process control systems: The design of human–machine interfaces,” *Saf. Des.*, vol. 44, no. 1, Jan. 2006, pp. 5–26.
[22] P. Nickel and F. Nachreiner, “Evaluation of presentation of information for process control operations,” *Cogn Technol Work*, vol. 10, no. 1, 2007, pp. 23–30.
[23] Y. Dahl, O. A. Alsos, and D. Svanæs, “Fidelity Considerations for Simulation-Based Usability Assessments of Mobile ICT for Hospitals,” *Int. J. Hum.-Comput. Interact.*, vol. 26, no. 5, Apr. 2010, pp. 445–476.
[24] E. Huang, E. Mynatt, and J. Trimble, “Displays in the Wild: Understanding the Dynamics and Evolution of a Display Ecology,” in *Pervasive Computing*, vol. 3968, K. Fishkin, B. Schiele, P. Nixon, and A. Quigley, Eds. Springer Berlin Heidelberg, 2006, pp. 321–336.
[25] H. Karvonen, I. Aaltonen, M. Wahlström, L. Salo, P. Savioja, and L. Norros, “Hidden roles of the train driver: A challenge for metro automation,” *Cogn. Ergon. Situated Hum.-Autom. Collab.*, vol. 23, no. 4, Jul. 2011, pp. 289–298.
[26] L. Norros and M. Nuutinen, “Performance-based usability evaluation of a safety information and alarm system,” *Int. J. Hum.-Comput. Stud.*, vol. 63, no. 3, Sep. 2005, pp. 328–361.
[27] D. Papachristos, P. Koutsabasis, and N. Nikitakos, “Usability Evaluation at the Ship’s Bridge: A Multi-Method Approach,” in *4th International Symposium on Ship Operations, Management and Economics*, 2012, Athen, Greece.
[28] M. Lützhöft, M. Lundh, and T. Porathe, “On board ship management overview system - an information sharing system on board,” *Int. J. Mar. Des.*, vol. 155 - part C, 2013.
[29] G. R. Wilkinson, “III.—Ergonomics in Ship Design,” *J. Navig.*, vol. 27, no. 04, 1974, pp. 471–478.
[30] P. Savioja and L. Norros, “Systems Usability — Promoting Core-Task Oriented Work Practices,” in *Maturing Usability*, E. C. Law, E. Hvannberg, and G. Cockton, Eds. Springer London, 2008, pp. 123–143.
[31] P. Savioja and L. Norros, “Systems usability framework for evaluating tools in safety–critical work,” *Cogn. Technol. Work*, vol. 15, no. 3, Aug. 2013, pp. 255–275.
[32] A. Rizzo and M. Palmonari, “Context and Consciousness: Activity Theory and Human Computer Interaction,” *User Model. User-Adapt. Interact.*, vol. 8, no. 1–2, pp. 153–157.
[33] L. Wennersberg, “Modeling and Simulation of Anchor Handling Vessels,” Master thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2009.
[34] M. Barrett, *Boundary relations : technological objects and the restructuring of workplace boundaries*. 2007.
[35] M. Lützhöft and M. Lundh, “Maritime Application of Control Systems,” in *Handbook of Control Room Design and Ergonomics: A Perspective for the Future, Second Edition*, CRC Press, 2008.
[36] R. E. Grinter, “Recomposition: putting it all back together again,” in *Proceedings of the 1998 ACM conference on Computer supported cooperative work*, Seattle, Washington, USA, 1998, pp. 393–402.
[37] P. Thomas and R. D. Macredie, “Introduction to the new usability,” *ACM Trans Comput-Hum Interact*, vol. 9, no. 2, 2002, pp. 69–73..
[38] Maritime Technology and Operation AS, “General steps of operational compentence.” Høgskolen i Ålesund.
[39] United States Nuclear Regulatory Commission, “Part 52—Licenses, Certifications, and Approvals for Nuclear Power Plants.” NRC library.
[40] J. Blomberg, M. Burrell, and G. Guest, “An ethnographic approach to design,” in *The human-computer interaction handbook*, J. A. Jacko and A. Sears, Eds. L. Erlbaum Associates Inc., 2003, pp. 964–986.
[41] R. Madden, *Being ethnographic - A guide to the theory and practice of ethnography*. Sage Croydon, 2010.
[42] J. Kjeldskov, et al, “Evaluating the usability of a mobile guide: the influence of location, participants and resources,” *Behav. Inf. Technol.*, vol. 24, 2005, pp. 51–65(15).
[43] V. Kaptelinin and L. J. Bannon, “Interaction Design Beyond the Product: Creating Technology-Enhanced Activity Spaces,” *Human–Computer Interact.*, vol. 27, no. 3, Dec. 2011, pp. 277–309.
[44] U. Schultze and W. J. Orlikowski, “Research Commentary---Virtual Worlds: A Performative Perspective on Globally Distributed, Immersive Work,” *Info Sys Res.*, vol. 21, no. 4, 2010, pp. 810–821.
[45] W. J. Orlikowski, “The sociomateriality of organisational life: considering technology in management research,” *Camb. J. Econ.*, vol. 34, no. 1, 2010, pp. 125–141.
[46] D. C. Neale, J. M. Carroll, and M. B. Rosson, “Evaluating computer-supported cooperative work: models and frameworks,” in *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, Chicago, Illinois, USA, 2004, pp. 112–121.
[47] E. T. Meyer, “Socio-Technical Interaction Networks: A discussion of the strengths, weaknesses and futhre of Kling’s STIN model,” *IFIP Int. Fed. Inf. Process.*, vol. 223, no. Social informatics: An information Society for All?, pp. 37–48.
[48] B. Latour, “On actor-network theory. A few clarifications plus more than a few complications,” *Soziale. Welt*, vol. 47, 1996, pp. 369–381.

When Simple Technologies Makes Life Difficult

Pursuing experienced simplicity in welfare technology for elderly

Suhas Govind Joshi
Department of Informatics
Faculty of Mathematics and Natural Sciences, University of Oslo
Oslo, Norway
joshi@ifi.uio.no

***Abstract*— In this paper, we use the case of elderly living with welfare technology to demonstrate how technology intended to be simple often defeats its own end. We discuss why simplicity requires attention and consideration not only to the context-detached design but also to the design in use as applying analytic and imagined simplicity does not guarantee experienced simplicity. We provide examples and evaluation results to help argue for our perspective on simplicity and present five implications for design pursuing simplicity.**

***Keywords* — simplicity; elderly; welfare technology.**

I. INTRODUCTION

With the large-scale rollout of welfare technology many elderly find themselves living a life surrounded by technology. One of the technological devices found in the apartment of an 84 year old lady residing in a local care home in Oslo is an automated light sensor in her living room. Because of the small size of her apartment she sleeps with the door open, and when she turns in bed at night the sensor in the living room registers her movement and the light is activated throughout the apartment. Her solution to this was to cover the sensor with tinfoil (as illustrated in Figure 1).

Figure 1. Covering a sensor with tinfoil (Photo: S. Finken [1])

This observation exemplified how simple technology may end up making life difficult, and served as a trigger for us to explore the matter of simplicity. This paper investigates difficulties with making technology simple for others, in this particular case making welfare technology simple for the elderly. The discussion is grounded in data gathered with three different evaluation methods spanning over six months involving 45 participants, including 30 elderly with an average age of 86 years.

The paper is structured as follows. In Section 2, we give an analysis of simplicity in the literature, as well as our perspective on the matter. In Sections 3 and 4, we outline the research context and research methods of our study before presenting the results in Section 5. We end the paper with a discussion in Section 6 on why simplicity is challenging through five implications for design pursuing simplicity.

II. SIMPLICITY

Simplicity in its most elementary definition describes something with an uncomplicated quality or condition. Researchers have applied the concept of simplicity to various research studies within various disciplines of computer science. Over time, this vague definition of simplicity has made it applicable to different areas of computer science, and in several disciplines the term has evolved into an established term with a more refined and tailored use mainly applicable to that specific discipline or context. As a philosophical principle, simplicity can be differentiated into ontological, following the principle of parsimony, and syntactical (structural) simplicity, perceived as elegance [2]. Hence, the theoretical perspective of the researchers in the debate of philosophy of science can heavily influence how they perceive and apply such a term. Lee et al. [3] describes simplicity within the area of Human-Computer Interaction (HCI) as "not only simple page layout but also interface organization, functionality, structure, and workflow and framework". Following this definition, simplicity in HCI encompasses various elements and researchers tend to find their own perspectives and definitions to simplicity. One of the most cited authors on simplicity, Maeda [4], defines his ten laws of simplicity (reduce, organize, time, learn, differences, context, emotion, trust, failure and the one). On the other hand, Colborne [5] concentrates on only four strategies (remove, organize, hide and displace) in his discussion on simplifying devices and experiences. Simplicity has also been analyzed through the notion of minimalism by Obendorf [6] who defines four types of minimalism (functional, structural, compositional and architectural) and utilize this perspective on minimalism to discuss simplicity in HCI. However, as Picking et al. [7] points out, design principles are in general often formulated as brief guidelines that aim to cover wide areas of application and apply to multiple domains simultaneously; it

 ISBN: 978-1-61208-382-7

is difficult to use these guidelines consistently as they rarely specify which specific design choices to make. Since laws, strategies and principles for simplicity can serve as everything from minor inspirations to governing factors, Obendorf [6] have called for more differentiated and concretized definitions of how simplicity is understood, and exactly how it influences the design outcome.

Several researchers have pointed out the importance of simplicity as a design principle in systems designed for the older population [8]-[10], however prior studies [2] suggest that perceived simplicity is context-dependent and relies heavily on the users' previous exposure. Our understanding of simplicity is anchored in two main elements, namely *mastery* and *context*. Both of these elements revolve around the users' experience and perception of the system in use rather than the isolated and context-detached design itself; simplicity is a characteristic of a system that manifests itself once the intended users take use of the system in its appropriate context. When using simplicity as a design guideline, one should always envision the act of simplification resulting in positive effects on the mastery of the user in the desired context. Blindly following simplicity as a design principle, e.g., reducing or hiding elements because general rulebook on simplicity says so, ignores the true intention behind the design choice, namely disentangling the perceived complexity. However, analyzing the simplicity laws and principles of Maeda, Colborne and Obendorf one quickly register that these laws mainly consider simplicity as context-independent. All of Colborne's four principles encourage modification to the design detached from the eventual context. Similarly, Obendorf relies on minimalism which itself does not automatically ensure systems free of complexity; it only encourages basic design with deliberate lack of decoration without discussing the perceived simplicity. From Maeda's ten laws we can extract five laws considering the relational use of the system rather than the system itself, namely time, learn, context, emotions and trust. Only these laws reflect how we understand simplicity, i.e., rather than being a term of size, quantity or volume, it should first and foremost reflect the contextual experience. Thus, simplicity in a system is not something one adds to the design; it is something achieved once mastery is uncomplicated in its appropriate context.

Our view on simplicity aligns with the research of Eytam & Tractinsky [11] suggesting that the ability to design own complexities can be a desire among users. They define this contrast between advocated guidelines for simplicity and the observed behavior as the paradox of simplicity, and argue that simplicity is not defined in objective guidelines but rather be understood through how the users perceive simplicity. The explicit focus on the users' side of the interaction in HCI influences how we discuss the concept of simplicity how it is a matter of more than just reducing complexity; simplification is an intricate and dynamic design principle embracing factors such as mastery and context of use as examples of decisive factors of simplicity. This is also in line with [2] who suggest that simplicity as a design principle should be a complex and flexible design paradigm rather than a simple dichotomous variable, incorporating elements such as user interface design, as well as contextual factors (for example integration to other IS). Keay-Bright & Howarth [12] focus on designing intuitive interfaces and describe simplicity not as a compromise in richness or diversity of human experience, but rather a minimal interface that empowers the users to design their own complexities that ensures mastery.

III. RESEARCH CONTEXT

A. *Empirical context*

This study is part of a larger long-term research project focusing on newly acquired welfare technology in local care homes in Oslo Municipality. The particular local care home involved in this study consists of 91 individual apartments for the elderly (with an average age of 84 years) organized with common reception, cantina and recreation room. There is no medical services provided, and those in need organize their own arrangements with the district home care services, however the elderly have access to basic services such as hairdressing, foot therapist, gym and cinema. The goal of the local care home is to be a smart house, for example actively utilizing technology in order to prolong the time elderly can remain independent in their own homes before being admitted to a nursing home. Each individual apartment comes pre-installed with a set of new technologies, including automated lighting, heating and ventilation control, stove guard, electrical sockets with timers, motion sensors in all rooms, video calling, door locks with radio-frequency identification (RFID), and a customized tablet. Since the building opened in 2012, our research group has been present at this facility, and this local care home is an excellent arena to study existing technology. It also serves a venue where we experiment with new and alternative welfare technology.

B. *Technology under evaluation*

In this study, we included the tablet and some of the room control devices in the local care homes. The main objective was initially to concentrate solely on the tablet, however we feared that only studying this touch-based device would restrict the discussion of simplicity to an analysis of touch-screen interfaces rather than being an open discussion on how the user experience simplicity in the welfare technological devices that surround them. As a result, we included a set of devices in the room, i.e., light, temperature and ventilation systems, as well as the RFID door locking system.

1) *Tablet*

The tablet illustrated in Figure 2 comes pre-installed in all apartments and introduces a new way of arranging, planning and keeping an overview of everyday activities, as well as allowing residents to order meals from the downstairs cafeteria straight from the device. The tablet also provides basic opportunities for communication, namely telephoning

 ISBN: 978-1-61208-382-7

and text messaging, as well as entertainment services, e.g., radio and an Internet browser. However, the tablet only comes with one mode and offers few options for customization, hence flexibility and robustness is of great importance as it needs to support the daily activities of all residents and employees.

Figure 2. The tablet

2) Room controls devices

Some of the pre-installed technologies and devices in each apartment is lightning, heating and ventilation control in every room of the apartment. This includes automated motion-activated light sensors, automated thermostat and automated adjustment of ventilation. The three photos in Figure 3 depicts a close-up of the heating interface as well as the RFID door locking system used to access each apartment. The door locks automatically, but opens with a RFID-card, and represents an interface few had experienced before. Since all these devices come pre-installed there is no option for the residents to utilize other interfaces or interaction methods, e.g., traditional door locks with keys or two-button light switches, and these can all be seen as a part of the "welfare package" in each apartment. As a result, they were tested together during the evaluations, and we will refer to these devices as "room control devices" in this paper.

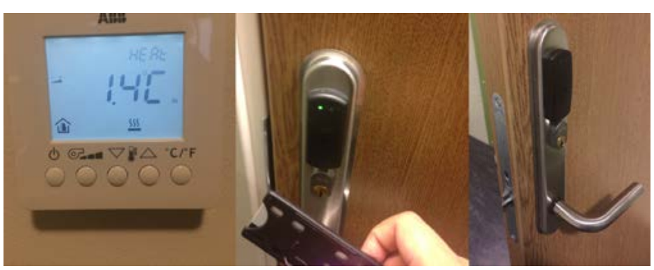


Figure 3. Heating control (left) and RFID door (right)

IV. RESEARCH METHOD

The data for this study was gathered over a six months divided into two phases. We were motivated by prior experiences with elderly and welfare technology [13][14] where findings suggest that giving enough time helps avoiding or eliminating bias. Three different methods of evaluation (Table I) were used during these two phases, and Figure 4 illustrates the outline of the research phases. We applied different methods of evaluation partly motivated by methodical triangulation, although the main reason was giving the participant more than just one opportunity to express their perspectives on simplicity. *The task-based group evaluation* allowed the participants to freely address simplicity issues during task walkthrough independent of schemas, heuristics or guidelines. Through the *simplicity evaluation* participants had a chance to evaluate the simplicity by grading pre-selected factors of simplicity, and during the *usability assessment* we did not ask them, but rather observed and measured them in order to discuss simplicity through their performance. The first phase included a task-based group evaluation, a simplicity evaluation and a usability assessment. The initial plan was to conduct these three activities during the first phase and then follow up with an equivalent usability assessment after six months with the same participants and the same usability criteria. However, due to the feedback and results discovered during the second usability assessment, we chose to repeat the simplicity evaluation as well.

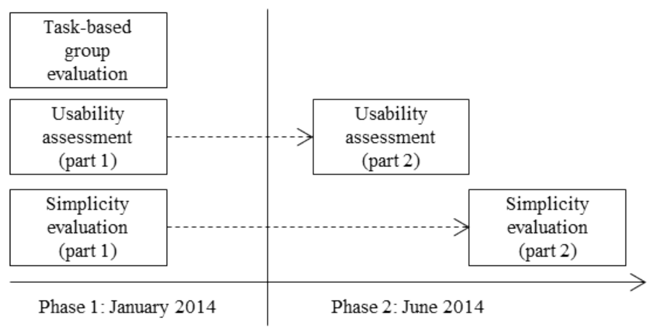


Figure 4. Outline of the research phases

TABLE I. OVERVIEW OF METHODS

#	Method	Participants Phase 1	Participants Phase 2	Participants Phase 1+2
A	Task-based group evaluation	21	-	21
B	Usability assessment	11	11	22
C	Simplicity evaluation	12	12	24

A. Task-based group evaluation

The task-based group evaluation was a part of a broad study where altogether 21 participants were engaged, namely 11 elderly, 7 employees and 4 experts. This dataset include several factors out of which some are not relevant for this study, although this evaluation has previously contributed to another study [13]. Nevertheless, the evaluation included a total of 6 sessions, 3 sessions with groups of elderly, 2 sessions with groups of employees, and 1 session with a group of HCI-experts. The sessions were structured as group walkthroughs of pre-selected representative tasks where the participants were asked to grade the severity of identified issues and then engage in a plenary discussion. Examples of representative tasks were ordering a meal and signing up for activities on the tablet and controlling lighting and ventilation in the room. During this session all participants labeled issues with predefined categories. The data included in this study are those issues labeled by the participants as “simplicity” issues. All participants were free to individually define what issues they considered to be simplicity issues.

B. *Usability assessment*

The usability assessment involved 11 participants; altogether, 7 elderly and 4 experts participated. The participants were given a set of 10 representative tasks to perform while completion time and error rates were measured and the sessions photographed. The tasks are listed in Table II. The tasks were distributed evenly between the tablet and the room control devices. Errors were counted and also divided into *deliberate errors* and *accidental errors*; the former represents errors where the user performed an action intentionally although performed the wrong action, while the latter represents unintentional actions. An example of a deliberate error is intentionally pressing the channel button on the television remote control when you want to adjust the volume because you in your best judgment consider the channel button to be the correct action for the desired outcome (i.e., adjust the volume), and you intentionally press that button. On the other hand, if you want to change the channel and while reaching for the correct button you unintentionally bump into the power button instead, then it is a case of an accidental error.

TABLE II. OVERVIEW OF PERFORMED TASKS

Task #	Task description
Task 1	Locking and unlocking the RFID door
Task 2	Playing a game on the tablet
Task 3	Browsing on the tablet
Task 4	Sending and receiving text messages on the tablet
Task 5	Listening to radio on the tablet
Task 6	Ordering food from the cafeteria on the tablet
Task 7	Activating room control devices with movement
Task 8	Setting and adjusting the ventilation
Task 9	Turning on and off wall and ceiling lighting
Task 10	Adjusting the heating level

These evaluations were carried out in the homes of 5 of the 7 participants, while 2 participants preferred to have the test conducted in an adjacent meeting room along with the experts. The usability assessment was repeated during the second phase in order to study changes in behavior, performance and satisfaction after six months. The conditions and environmental factors were similar between the two assessments with the exception that 1 additional elderly participant chose to not have the test in her apartment.

C. *Simplicity evaluation*

The goal of the simplicity evaluation was to provide the elderly with an opportunity to evaluate the simplicity without being restricted to certain tasks (as in method A) or tied to their performance (as in method B). Hence, the participants were asked only to grade the simplicity of the tablet and the room control systems. Each participant were given an individual oral and written explanation of each factor and was then asked to grade the simplicity factor from 1-5. The evaluation comprised 7 factors redefined from the 5 laws of Maeda coinciding with our perspective of simplicity, namely the symbiotic relationship between mastery and context. The 7 elements were *intuitivity*, *organization*, *memorability*, *error rate*, *time*, *learnability* and *trust*. Intiutivity reflects the perceived easiness when first approaching the system in the given context, while learnability and memorability describes the system’s ability to foster mastery and maintain it over time. With organization we did not look at organization of the interface, e.g., icon clutter, but studied how the system fitted within its context. We also included time, i.e., their experience on their own performance and error rate, i.e., how many errors they encountered, in order to study their own perspective on mastery.

D. *Participants*

The three methods involved 45 participants altogether and the participants divided into four user groups described in Table III. The elderly ($n = 30$) participated in all methods during both phases, while the usability experts ($n = 8$) participated during both phases of the simplicity evaluation and the usability assessment. Finally, the employees only ($n = 7$) participated in the task-based group evaluation. The elderly were recruited among the residents at the local care home and their age ranged from 79-94 ($\mu = 86$). Upon moving into this local care home, all elderly were cognitively cleared by medical experts, i.e., possessing at least an acceptable level of cognitive and reasoning abilities. They struggled with various medical conditions, e.g., reduced motor abilities or reduced vision, and they represented a broad range of social difficulties.

TABLE III. OVERVIEW OF PARTICIPANTS

User group	User role	Use frequency	Expertise	Participated in method #	N
The elderly	End-users	Every day	(none)	A, B, C	30
Daytime employees	End-users and trainers	Every day	Health and domain	A	4
Shift work employees	End-users and trainers	Once a week	Limited domain-expertise	A	3
Usability experts	None	One-time only	HCI and usability	A, B	8

V. RESULTS

A. *Task-based group evaluation*

Out of a total of 39 identified issues, 17 were considered simplicity issues by at least one of the user groups. Each group that had identified the issue was then asked to grade the severity of the issue as *minor (M)*, *serious (S)* or *critical (C)*. All identified issues are listed in Table IV. The *aggregated degree of seriousness* reflects the final level of seriousness assigned to the issue based on the grading of the groups. If there were disagreements between only two groups, the most serious grading took precedence; otherwise the number of occurrences decided this aggregated degree of seriousness. Out of these 17 identified issues 5 were labeled as critical issues, 7 were categorized as serious issues, and 5 were considered minor issues. The group of elderly reported a total of 14 issues, out of which 36 % were graded as minor. The similar percentage was lower for the two other groups, respectively 25 % for the employees and 27 % for the experts. Since both the employees and experts reported fewer issues overall that the other two groups, this implies that the employees and experts regarded identified issues as more severe that the elderly, with a percentage of 75 % (employees) and 73 % (experts) graded as either serious or critical against only 64 % for the elderly.

We also wanted to study the balance of simplicity, i.e., identify the level of simplicity where the system was neither too simple nor too complex. As a result, we also asked the participants to differentiate between issues they considered a result of the vendor making the interface or interaction *too simple*, i.e., a matter of oversimplification, and issues they considered *too complex* and wished were further simplified. 13 issues were considered a result of oversimplification and participants expressed usability issues due to interface, language, symbols etc., being too simple for their liking. 4 of the 5 critical and 6 of the 7 serious issues were labeled oversimplified. It should be noted that similar to the aggregated degree of seriousness, the expressed simplification desire is the aggregated evaluation of the group(s) who brought forward the issues, however all groups answered unanimously for all issues. As a result, their individual answers are not presented as with the degree of seriousness where we encountered variations between groups.

Most of the issues had a clear consensus on the grade of severity. Only those 3 cases where two groups addressed an issue and simultaneously gave it different grades did we encounter any disagreements. Rather than considering the grade of one group as more important than other, we chose instead to always use the highest grade. This was considered an acceptable solution by the participants; for example, the elderly labeled the highest number of issues as minor issue, but for 3 of the 5 issues that the elderly labeled as minor issues (#1, #10, #29) the aggregated grading was upgraded to serious since either the employees or the experts regarded the issue as serious. For the two remaining issues one was only reported by the elderly (#11) and one group disagreed with the elderly on the severity grade of the last issue (#28). Additionally, only in 3 cases were the issue only addressed by one group (out of which two were minor issues), and the overall consistency of the grading of the issues was therefore considered to be good.

B. *Usability assessment*

The usability assessment included 10 tasks (Table II) tested by 7 elderly and 4 experts in each of the two phases, and Figure 5-7 presents the completion time and error rate for each of the tasks in both phases. The completion time listed for each task is the average time spent by all 11 participants to complete the task, while the error rate is the average error rate for deliberate and accidental errors.

On average, the experts performed their tasks during the first phase within half the time of the elderly ($\mu_{experts} = 173.11$ against $\mu_{elderly} = 330.57$), and did so with half as

TABLE IV. IDENTIFIED SIMPLICITY ISSUES

Issue #	**Issue description**	**Aggregated degree of seriousness**	**Group 1 Elderly**	**Group 2 Employees**	**Group 3 Experts**	**Imbalance issue**
1	The device screen always stays on (even in standby mode)	**S**	M	S	S	Too simple
5	The phone icon color is misleading	**S**	S	M	S	Too complex
7	There is no indicator of remaining battery	**C**	C	C	C	Too simple
8	There is no indication of the device being charged or already fully charged	**S**	-	S	-	Too simple
10	The system signals two new messages when just one message arrive	**S**	M	S	-	Too simple
11	The system uses separate indicators to indicate the same message	**M**	M	-	-	Too complex
15	There is one phone number for texting (12-digit) and another for calling	**C**	C	S	C	Too complex
20	The default values in text boxes are misleading and unpractical	**S**	S	C	S	Too simple
21	It is impossible to grad the on-screen keyboard in certain views	**C**	S	-	C	Too simple
24	The language is inconsistent	**S**	S	S	-	Too simple
25	It is too easy to delete everything	**M**	-	M	M	Too simple
28	The events in the calendar are not chronologically ordered	**M**	M	S	M	Too complex
29	The duration of phone calls is missing	**S**	M	-	S	Too simple
34	There is no comment feature on activities and events	**M**	-	-	M	Too simple
35	The language is confusing	**M**	S	M	-	Too simple
36	The icons are confusing	**C**	S	-	C	Too simple
38	The notifications are misleading	**C**	C	S	-	Too simple

many deliberate ($\mu_{experts}$ = 1.82 against $\mu_{elderly}$ = 3.90) and accidental ($\mu_{experts}$ = 1.18 against $\mu_{elderly}$ = 3.18) errors. Their standard deviation also confirms a more consistent performance throughout the 10 tasks both time wise ($\sigma_{experts}$ = 11.90 against $\sigma_{elderly}$ = 36.66) and error wise ($\sigma_{experts}$ = 0.52 against $\sigma_{elderly}$ = 1.06 and $\sigma_{experts}$ = 0.32 against $\sigma_{elderly}$ = 0.59). The average completion time for all 10 tasks increased slightly between the first and second phase ($\Delta\mu$ = 8.29, $\Delta\sigma$ = 7.36) for the elderly. There is no clear consistency in how the user performs on average in each task. The completion time of 4 tasks went down with an average of 9.46 seconds, while the completion time of the remaining 6 tasks went up with an average of 15.98 seconds. The deliberate error rate dropped for 6 tasks ($\Delta\mu$ = 0.36) and increased for the other 4 tasks ($\Delta\mu$ = 0.46), and the accidental error rate increased for 4 tasks ($\Delta\mu$ = 0.29), dropped for 4 tasks ($\Delta\mu$ = 0.32) and remained unchanged for the remaining 2 tasks. However, there is no correlation between which tasks that went up in deliberate or accidental error rate. Only for one of the tasks (#4) did the sum of deliberate and accidental errors decrease when the completion time decreased. For the other 3 tasks, where the completion time dropped (#1, #2 and #10), one increased the sum of errors by 0.14 (#1) while the two other had no change in error rate even though the completion time decreased.

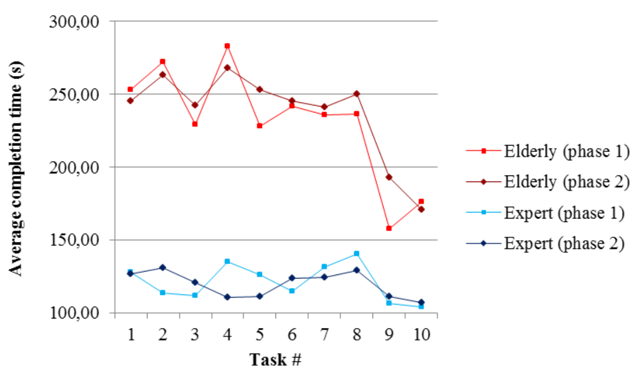


Figure 5. Overview of average completion time (s)

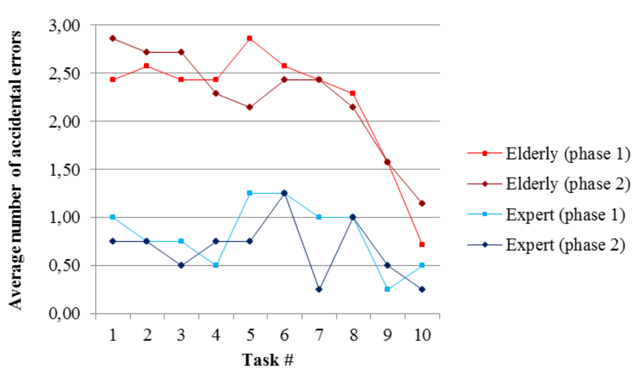


Figure 6. Overview of average number of accidental errors

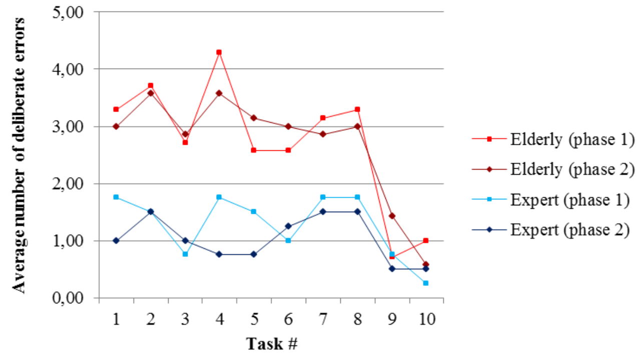


Figure 7. Overview of average number of deliberate errors

We registered that these two last tasks had the lowest completion time in both cases for all participants, as well as the lowest error rate (both deliberate and accidental) for both groups. A similar performance pattern was also registered among the experts, a group with less performance fluctuation than the elderly, and these were the two tasks with highest mean deviation in both phases for both groups. These two tasks were also the only tasks where the group of elderly matched the performance of the experts. The average difference in completion time between elderly and experts in phase 1 was 110.23 seconds (σ = 30.9) and 117.55 seconds (σ = 26) in phase 2, while the difference for task #9 and #10 were only 66.43 in phase 1 and 67.88 in phase 2. Similarly, the difference in deliberate error rate had an average of 1.46 (σ = 0.69) in phase 1 and 1.68 (σ = 0.73) in phase 2, while the difference for task #9 and #10 were only 0.48 in phase 1 and 0.41 in phase 2; the accidental error rate had an average difference of 1.4 (σ = 0.45) for phase 1 and 1.57 (σ = 0.48), compared to 1.2 difference in phase 1 and 0.55 in phase 2 for task #9 and task #10. Consequently, this anomaly is not a result of learning effect but rather a sign of tasks that were significantly easier than the rest.

C. *Simplicity evaluation*

Figures 8 and 9 present the results from both phases of the simplicity evaluation. During the first phase, there were clear differences in opinion between the participants. While the average score of the 12 participants ended up on the upper half of the scale, the deviation within the data was large (μ = 3.4 and σ = 0.79), and participant #10 gave 4.4 out of 5 on average for the 7 factors of simplicity, whereas participant #11 only gave 1.7 out of 5. The average score given to each of the 7 factors were much more evenly distributed with only half the deviation (σ = 0.4) despite some of the factors having a much higher internal deviation (e.g., memorability with μ = 3.0 and σ = 1.0). The second phase yielded results very similar to the first phase. There were few changes in how the users perceived and rated the 7 factors with the highest factor difference between the two phases being as low as 0.3 (intuitivity and trust), while the rest averaged at 0.15. However, almost all participants have

 ISBN: 978-1-61208-382-7

changed their perception of simplicity since the first phase. Participant #10 and #12 both end up with an average score 0.1 below their previous average, and for some participants, e.g., participant #6 with a 0.9 difference, the change in opinion is much more evident. 5 of the participants end up giving a higher average score during the second phase ($\Delta\mu = 0.53$), while the remaining 7 reduce their average score ($\Delta\mu = 0.37$). Hence, even though the number of participants increasing their score between the two evaluations is lower than those reducing it, the difference in their average score brings the total average up ($\Delta\mu = 0.1$). While the overall perception of simplicity does not necessarily change much, the reduced deviation between participants carefully suggest that their opinions have harmonized during the six months between the two evaluations ($\sigma_{phase2} = 0.51$ against $\sigma_{phase1} = 0.79$).

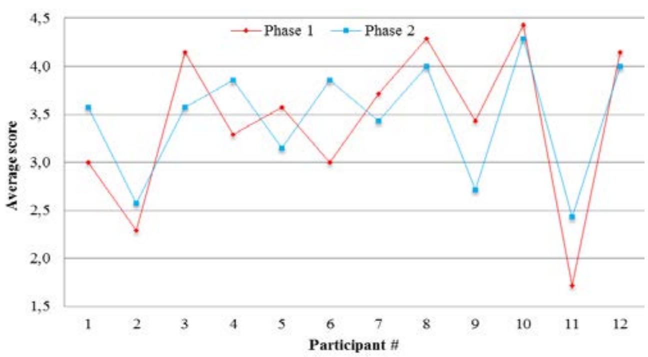


Figure 8. Average score given by each participant

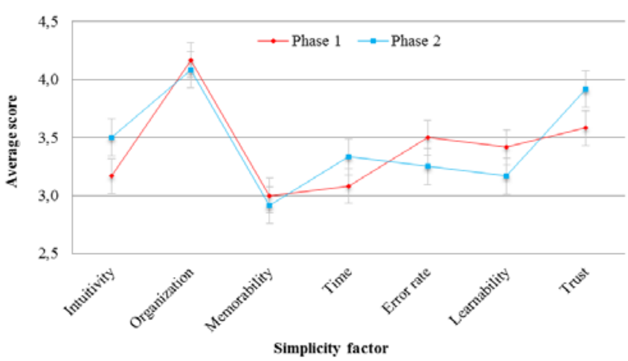


Figure 9. Average score given for each simplicity factor

VI. DISCUSSION

A. *Ensuring familiarity and transferability*

Mastery requires understanding and learning. It also relies heavily on the users' previous exposure, and design following simplicity should evoke a connection to prior experiences. Thus, the elderly rely heavily on transferring prior skills and knowledge in order to adapt a level of understanding and learning that nurtures mastery. One of the key challenges with both systems evaluated in our study was the lack of consistent metaphors. Several elderly with prior experience with devices similar to those used in our evaluation were unable to utilize prior knowledge due to metaphors not being consistent; simplicity also encompasses other design principles, e.g., consistency and affordance. Actions, icons, symbols and other metaphors should mediate experiences rather than direct [11]. And the diverse backgrounds of the elderly made us very aware of the difficulty of reducing complex information into simplified metaphors where everyone understands both the metaphors and the symbolic meaning or feeling they encompass. This challenge has been addressed by previous studies [15] who relied on a simplified design to trigger a nostalgic effect in order to help familiarizing metaphors.

In our studies, several elderly struggled with the tablet responding to their actions with unexpected outcomes. One example included elderly trying to use prior knowledge like familiarized gestures on the tablet, e.g., pinching and dragging to zoom or sliding actions to scroll, when visiting websites during task #3 (Table II). The system being of a different operating system than what they had previously used responded differently than expected; the slider scrolled the website in the opposite direction and the pinch and drag gesture were not recognized by the system at all. Another prominent example was the RFID doors automatically locking if they were closed, i.e., a contrast from the traditional method of locking doors, by turning a key. The doors were heavy and closed automatically, and once closed they would also lock automatically like a spring lock, only without any sound or click. It was especially confusing during the first evaluation as the elderly still had not memorized that the redundant key hole affording use of traditional keys (Figure 3) served no purpose, and repeatedly expected the door to be locked manually with a key after closing the door, when instead the door would automatically close and lock behind them. In fact, the accidental error rate for the task involving the doors (task #1 in Table II), was one of the tasks with highest combined average error rate was one out of only four cases where the deliberate error rate increased between the first and second phase. This was a matter of confusion and reported as one of the main issues responsible for the degree of learnability dropping between the two simplicity evaluations (see Figure 9). A third example included problems during text messaging (task #4 in Table II). When asked to send and receive text messages, several old and familiar metaphors were suddenly replaced by new unfamiliar metaphors where the elderly struggled with applying old knowledge to the new system. For example, the phone number was not their usual phone number, nor did it resemble a traditional phone number (issue #15), and the icons used to symbolize contacts and messages were not recognized (issue #36). The task of text messaging yielded the highest number of deliberate errors during both evaluations, and this was clearly a result of their attempt to perform actions associated with prior experience or applying old metaphors to the new

ISBN: 978-1-61208-382-7

system that were no longer compatible or purposeful. Through these three examples we discovered that the most confusing and frustrating situations arose when the elderly performed an action where the outcome was unclear or unfamiliar. Familiarity and transferability became strong indicators of the ability to master new systems; when actions became disconnected from their meaning, the purposefulness in the actions disappeared and mastery suddenly became a challenge.

B. Maintaining purposeful actions

In order to further discuss purposeful actions we gave the participants six months to familiarize themselves with the systems before asking them to evaluate the simplicity a second time. 3 participants (#1, #3, #6 in Figure 8) reported a higher average score during the second simplicity evaluation, suggesting a more positive attitude towards the 7 elements of simplicity we evaluated. As a result, we investigated whether this was a result of increased learning and understanding, or just a matter of increased use frequency. When discussing the mastery of the system, we need to distinguish between increased ease due to more frequent use and increased ease due to actions, metaphors and language suddenly making more sense. It was unanimously agreed upon that the participants reported a higher score as a result of increased frequency rather than actions, metaphors and language making sense. Confusing metaphors were still confusing and during the six months participants had learned certain use patterns by heart. To them, adopting strategies to avoid problems uncomplicated and improved the efficiency once memorized. However, it was evident that time did not contribute to increased understanding of metaphors, but rather resulted in incorporated strategies and workarounds. Confusing actions, metaphors and language remained confusing even after six months of use, also for those reporting a higher average score, and the increased perception bloomed out of the development of personal strategies for memorizing or working around troublesome tasks. This is an important finding as patience is often considered a virtue when elderly adapt to new technology, including in our own previous work [13][14]. In this study however, we observed that actions, metaphors and language confusing the ended up remaining confusing after six months as well; providing more time might heal all wounds, but it does not guarantee disentanglement of perplexities and disorientations.

Another argument for ensuring purposeful actions is to maintain good mapping. Natural mapping is understood as designing the interface in such a manner that the user can readily determine the relationship between the action and the outcome into the world [16]; i.e., a design where the user is able to associate cause with effect, thereby understanding expected output for provided input. As an example, the autonomy and intangibility of the automated light sensor evaluated during the usability assessment (task #9 in Table II) imposed several challenges to mapping. The physical zone in the room where movements were recognized was not clear, and there were no indications in the interface towards the intensity of the light or the duration of the light. One participant claimed that the best mapping for her was a traditional light switch where up meant on and down meant off in the middle in the room where the left switch controlled the lamp to the left and the right switch controlled the lamp to the right. Similarly, replacing traditional door keys with RFID cards to unlock doors had similar effects on the natural mapping; the users were unable to properly answer how long the door remained unlocked once the RFID card was scanned or determine the minimum required distance between the RFID card and the scanner on the door.

C. Adapting to evolving perceptions of simplicity

Trier & Richter [2] argues that the application of simplicity as a design guideline requires flexibility. Between the two phases we observed two participants undergo changes in their overall health level. There were significant differences in their cognitive and reasoning abilities. For example, one of these participants could no longer explain the numbers on the display used to adjust heating levels (Figure 3). She had a custom color marker that indicate up and down for temperature as the up- and down-facing arrows no longer served as metaphors for increasing and decreasing the room temperature. While the arrows and display offered sufficient explanation during the first evaluation, she could no longer explain the details of the system during the second evaluation, e.g., the meaning of "1.4^{o}C" on the display (as illustrated in Figure 3). Instead, she found that blue and red colors helped her remembering that if she pressed those buttons long enough it would eventually get colder or warmer. This exemplifies how typical aging symptoms, e.g., reduced cognitive capacities, clearly influenced both their performance and their assessment of simplicity. Related work [8] discuss how only paying attention to physical and perceptual characteristics of elderly end up struggling with coping with the cognitive behavioral characteristics and traits of becoming elderly. Consequently, we consider achieving simplicity among elderly especially difficult as the elderly undergo rapid cognitive, physical and social changes in their lives that alter their attitude and opportunities towards technology. As metaphors lose their abilities to aide us with understanding and interacting with the system, our perception of the simplicity of the system deteriorate over time. Simplicity is not a constant factor that remains the same throughout of life, but rather one of the dynamic and flexible factors that evolves along as we evolve; acquiring new knowledge, entering new contexts and adapting new technologies contribute to reshaping our view on simplicity and what we perceive as simple. Similarly, s changes in our lives can contribute to complicating systems we once considered simple; it often becomes a matter not only of preference, but also a matter of limited opportunities. Over a period of six months the perspectives of all the elderly participants changed in both the simplicity evaluation and the usability assessment. A design offering simplicity should therefore adapt according to the changing behavior and abilities of the elderly.

Cooper et al. [17] also discusses the phenomenon where visual simplicity leads to cognitive complexity due to an unbalanced reduction. Several participants struggled with adapting to new technology due to cognitive load and preferred to rely on old knowledge and metaphors instead; they preferred familiar technologies, even those comparatively inefficient and impractical, because they could rely on habits. Examples of such desires included installing old landline telephones rather than telephoning from the tablet even though the latter was free, and using old televisions with large physical buttons instead of new flat screen television even though it involved getting out of the couch every time to change channel. A frequent counter-argument is that this behavior is a result of their attitude towards technology in general rather than a matter of cognitive overload, however their attitude during the rest of discussions clearly suggested that they were positive towards technology but struggled with adapting to certain aspects of the system, in this particular case it was the misleading colors (#5), the two separate phone number (#15) and the confusing language (#35) that caused the perceived complexity (Table IV). If those aspects of the systems are metaphors intended to bridge the gap between the system and prior experiences, achieving mastery can become difficult, sometimes also impossible. As a result, we argue that design striving for simplicity should be open to seemingly inefficient and impractical features if they evoke positive stimuli for the users, e.g., allowing them to take advantage of old habits rather than adapting new ones.

D. *Avoiding forcing ways of reasoning*

By oversimplifying technology, we limit the users' freedom and make decisions on their behalf by forcing them into predefined patterns of behavior that do not necessarily comply with their needs. The participants in our study disliked the predefined settings and missed working with a system that could adapt or be customized to fit their cognitive and bodily capabilities. Similar to studies of Eytam & Tractinsky [11], several participants desired the ability to design their own complexities. Our principal example was the tablet which did not offer any customization options or the option to install custom application with services that the system did not currently offer. Once one participant discovered a way to override the system and install own application, in this case a video chat application, several others asked for instruction on how to do so as well. This case exemplified how the intention of simplifying the system by removing seemingly undesired features became a restriction of the users' desires. By directing, limiting or forcing decisions on the elderly, the outcome might end up being stigmatizing rather than inspiring [18]. For the elderly who feel they are losing control and influence over their own life, this stigma through oversimplification may further assume a role as a reinforcing factor counteracting dignity and integrity by depriving them of their opportunity and right to autonomy [13]. This may again influence the ability to learn how to operate such systems as more general suggestions on simplicity in learning advocates the use of environments where users feel good and able. From their own results, Keay-Bright & Howarth [12] conclude that environmental factors that stimulate and encourage without prejudgment is a vital requirement for learning. Besides decelerating or even preventing the process of mastering, inhibiting learning has also proven to result in negative experiences for the elderly. The feeling of helplessness that comes with aging makes the elderly more aware of their own dependability, and previous findings from our studies showed several participants felt deprived of their independence due to oversimplified and restrictive systems limited their opportunity to function at their best level [13].

E. *Balancing the simplicity*

The phenomenon of systems involving simplification measurements that end up having the opposite effect is often referred to as fake simplicity. Colborne [5] describes fake simplicity as the idea unable to ever meet its initial promise, instead just making everything unnecessarily complex and less effective. One example was the microwave of one of the participants that instead of using time or watt as input, used pictures of a pizza slice and a cup of tea to signal the duration and strength. Another example mentioned by a participant was his washing machine with only predefined programs where neither duration nor temperature was specified. Oversimplification can prevent mastery by concealing important components of the interaction thereby preventing the user from learning the relationship between action and effect. It also demonstrates how mastery requires balance. On one hand, the system needs to foster mastery through a design that is perceived as free of complexities; on the other hand, the system should encourage mastery by challenging and exciting the user and simultaneously avoiding oversimplified and condescending interfaces. Finding this balance where users are both presented with challenging tasks and at the same time provided with enough help to solve them helps us preventing that the system tips over in either direction.

During the task-based group evaluation, the participants were asked to identify simplicity issues as either too simple or too complex systems. As a result, they were asked to clarify whether it was a case of lack of simplicity or abundance of simplicity, i.e., a complex issue that could benefit from simplification or an issue that was simplified to such an extent that it had become oversimplified and demeaning. Surprisingly, 13 out of 17 issues were classified by the participants as matters of oversimplification, i.e., that the simplification of the interface or interaction resulted in either poor usability or led undesired user experiences. The most important finding from these results was that simplicity is not a principle where “one size fits all”. One argument presented by an elderly lady for not liking the phone function of the tablet was that with tablets and mobile phones, the action of answering a call required an additional step. With a traditional land line phone, picking up the phone initiated the call, while on newer device she would first have to press an answer button and then pick up the phone, thereby complicating it for her by introducing additional step. Secondly, the internal disagreement between the groups further suggests that the elderly might have a different outlook on simplicity relatively compared to the two other groups, thereby demonstrating a variation not only between individuals but also between groups of individuals. What remains a matter of simplicity for the elderly seems to deviate from what the employees and experts consider simplicity issues further suggesting that simplicity in use is different from analytic simplicity or imagined simplicity. Achieving simplicity without simultaneously weakening the functionality is one of the great struggles of designers, and it is vital to find this point of intersection where constructive simplification suddenly begins to defeat its own end. Simplicity is not only a matter of aesthetics; it is also a matter of balanced functionality.

VII. CONCLUSION

This study was motivated by the old lady covering up her automated light with tinfoil because the intended simplicity ended up complicating her life. In this paper, we have demonstrated additional examples of simple technology aggravating the lives of elderly, thereby illustrating how we believe simplicity in context-detached design to be different from experienced simplicity; analytic and imagined simplicity does not ensure simplicity in use. We argue that simplicity is anchored in *mastery* and *context* and that simplicity should (1) build on familiarity and the ability to utilize old knowledge to help mastering the system; (2) ensure purposeful actions where the user can understand and learn to master the system; (3) adapt along with the evolving contextual factors; (4) avoid limiting the users to predefined patterns of behavior and allow them to use and master the system as they find appropriate; and (5) find the balance where the design is simple enough to be understood and learned, yet challenging enough to allow users to progress towards mastery. Only by doing so, we can achieve mastery in the intended context of use, which is what we believe simplicity to be.

REFERENCES

[1] S. Finken and C. Mörtberg, “Performing Elderliness – Intra-actions with Digital Domestic Care Technologies”, in ICT and Society, K. Kimppa, et al., Editors. 2014, Springer Berlin Heidelberg. p. 307-319.

[2] M. Trier and A. Richter, " I Can Simply…"-Theorizing Simplicity As A Design Principle And Usage Factor, in ECIS 2013. 2013. p. 2-3.

[3] J. Lee, D. Lee, J. Moon, and M.-C. Park, “Factors affecting the perceived usability of the mobile web portal services: comparing simplicity with consistency”. Information Technology and Management, 2013. 14(1): p. 43-57.

[4] J. Maeda, “The laws of simplicity”. 2006: MIT press.

[5] G. Colborne, “Simple and usable web, mobile, and interaction design”. 2010: New Riders.

[6] H. Obendorf, “Minimalism: Designing Simplicity”. 2009: Springer.

[7] R. Picking, V. Grout, J. Mcginn, J. Crisp, and H. Grout, “Simplicity, consistency, universality, flexibility and familiarity: the SCUFF principles for developing user interfaces for ambient computer systems”. International Journal of Ambient Computing and Intelligence (IJACI), 2010. 2(3): p. 40-49.

[8] H. Akatsu, H. Miki, and N. Hosono, “Design principles based on cognitive aging”, in Human-Computer Interaction. Interaction Design and Usability. 2007, Springer. p. 3-10.

[9] F. H. A. Razak, N. A. Razak, W. A. Wan Adnan, and N. A. Ahmad. “How simple is simple: our experience with older adult users”. in Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction. 2013. ACM. p. 379-387.

[10] S. Sulaiman and I. S. Sohaimi. “An investigation to obtain a simple mobile phone interface for older adults”. in Intelligent and Advanced Systems (ICIAS), 2010 International Conference On. 2010. IEEE. p. 1-4.

[11] E. Eytam and N. Tractinsky. “The Paradox Of Simplicity: Effects Of User Interface Design On Perceptions And Prefence Of Interactive Systems”. in MCIS 2010. 2010. p. 30.

[12] W. Keay-Bright and I. Howarth, “Is simplicity the key to engagement for children on the autism spectrum?”. Personal and Ubiquitous Computing, 2012. 16(2): p. 129-141.

[13] S. G. Joshi, Emerging ethical considerations from the perspectives of the elderly, in Ninth International Conference on Cultural Attitudes in computer-Human Interactions, M. Strano, H. Hrachovec, and S. Fragos, Editors. 2014. p. 1-15.

[14] S. G. Joshi and A. Woll, “A Collaborative Change Experiment: Telecare as a Means for Delivery of Home Care Services”, in Design, User Experience, and Usability. User Experience Design for Everyday Life Applications and Services. 2014, Springer. p. 141-151.

[15] M. Nilsson, S. Johansson, and M. Håkansson. “Nostalgia: an evocative tangible interface for elderly users”. in CHI'03 Extended Abstracts on Human Factors in Computing Systems. 2003. ACM. p. 964-965.

[16] D. A. Norman, “The design of everyday things”. 2002: Basic books.

[17] A. Cooper, R. Reimann, and D. Cronin, “About face 3: the essentials of interaction design”. 2007: John Wiley & Sons.

[18] L. Rosenberg, A. Kottorp, and L. Nygård, “Readiness for Technology Use With People With Dementia The Perspectives of Significant Others”. Journal of Applied Gerontology, 2012. 31(4): p. 510-530.

Identifying User Experience Elements for People with Disabilities

Mingyu Lee, Sung H. Han, Hyun K. Kim, Hanul Bang
Department of Industrial Management & Engineering
POSTECH
Pohang, Republic of Korea
{mk3215s2, shan, emokubi, hanul89}@postech.ac.kr

***Abstract*—This study aims to identify the elements of the user experience (UX) of mobile products and services for people with disabilities. Although many researchers have emphasized UX in designing new products and services, common understanding of UX for those with disabilities is absent. This study identified UX elements for people with disabilities by analyzing previous studies. A total of 45 articles from the literature were analyzed, and as a result, UX elements for people with disabilities were identified. The results can be used as criteria for developing new products/services or evaluating existing products/services.**

Keywords—User experience (UX); UX elements; Disabled people; Usability; Affect; User value.

I. INTRODUCTION

User experience (UX) refers to all of experiences resulting from interactions that a user has with a product or service [1][2]. With growing interest in UX in the field of Human-Computer Interaction (HCI), many studies have been conducted in both industry and academia [3].

However, existing studies of UX have rarely considered people with disabilities. People with disabilities have different abilities to sense external information, so their experience with products or services can be different from that of non-disabled people. Understanding the UX of people with disabilities should be a great help to the development of products and services.

This study aimed to identify the UX elements of people with disabilities. We defined UX of people with disabilities as an experience that consists of aspects of interaction between the disabled and products/services which are influenced by assistive technologies. We investigated the elements based on the definition. This study analyzed relevant literatures on usability for people with disabilities and UX of general users, and identified the elements of UX for those with disabilities. During surveying the literatures, the specific types of disabilities were not considered to collect diverse literatures. Mobile devices have been regarded as important products for people with disabilities, so this study focused on the UX of mobile devices/services.

Section 2 presents the existing studies on people with disabilities and their limitations. Section 3 explains how to derive the UX elements for people with disabilities in the study, and presents the results of UX elements. In Section 4, the study analyzes the differences between the UX elements for people with disabilities and that for general users. We present the characteristics of UX elements for people with disabilities.

II. EXISTING STUDIES ON USABILITY FOR PEOPLE WITH DISABILITIES

Many studies related to the design of products and services for people with disabilities have been conducted and suggested a variety of design concepts related to usability or accessibility. The concept of barrier-free design emerged in the 1950s. It aimed to remove obstacles in houses or buildings for those with physical disabilities [4]. Accessible design focuses on standardizing designs for people with physical limitations to maximize the number of potential users [5]. Universal design is the most recent concept and has as its goal designing products or environments usable by all people without any need for adaptation or specialized design [6]. Inclusive design is a similar concept to accessible design, and design for all is another term used to refer to universal design.

Many studies suggested the concept related to people with disabilities, but they just focus on improving the usability of products and services for people with disabilities. Similar to non-disabled people, however, those with disabilities also experience user affect [7] and user value [8]. It is necessary to consider the comprehensive UX of those with disabilities, not only usability. This study attempted to identify the UX elements for people with disabilities through a literature review.

III. IDENTIFICATION OF UX ELEMENTS FOR PEOPLE WITH DISABILITIES

First, we surveyed the literature using keywords related to UX for people with disabilities, such as assistive technology, accessible design, universal design, and universal usability. Various types of literature, including journals, conference proceedings, magazines, reports, and books, were considered (Table 1), and 34 articles were collected in the literature survey from Google Scholar and SCOPUS databases. UX elements for people with disabilities have not been identified before, so evaluation criteria of products/services for people with disabilities or design considerations for them were considered as the candidates of UX elements in the study. As a result, a total of 49 candidates for the UX elements were identified, such as accessibility, perceivability, coequalness, and independence.

TABLE I. REPRESENTATIVE SOURCES OF THE SURVEY

Journals
- Disability & Rehabilitation - Technology and Disability - American journal of physical medicine & rehabilitation - Journal of Visual Impairment & Blindness - International Journal of Design - Design studies - International Journal of Industrial Ergonomics - International Journal of Human-Computer Interaction
Conferences
- Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI Conference) - International Conference on Universal Access in Human-Computer Interaction - Proceedings on the conference on Universal Usability - The ERCIM Workshop on User Interfaces for all
Books/Reports
- Design and Use of Assistive Technology - Plenum Series in Rehabilitation and Health - The universal access handbook - Universal Access in Human Computer Interaction - Design Council Report

Second, we collected the UX elements for general users to get the insights for UX elements for people with disabilities. 11 journal papers and reports which state the UX elements for general users comprehensively were collected and 92 candidates for the UX elements were identified.

Then, we integrated all of the candidates and removed candidates not appropriate to people with disabilities and mobile products or services based on three criteria; importance of elements, relevance to disabilities and mobile products/services, and relevance to the purpose of using a product/service. For example, translucency and durability were excluded. Those were important aspects of mobile devices to general users, but color and shape, which were essential characteristics of product to recognize a product, were much important than translucency to people with disabilities. Durability usually was not the purpose of using a mobile product or a service.

Lastly, we grouped the elements with similar characteristics or meanings. Based on the existing UX studies, the UX elements were classified as usability, affect, and user value [9]. The UX of people with disabilities would also follow the same classification. Finally, the UX elements of people with disabilities and their definitions were identified as presented in Table 2.

TABLE II. THE UX ELEMENTS OF PEOPLE WITH DISABILITIES AND THEIR DEFINITIONS

Elements		Definitions	Examples or similar concepts
Usability	Accessibility	Degree to which a product or a service is enable the user to approach or operate	Accessible size, Input assistance, Visibility, Audibility
	Effortlessness	Ability of a product or a service to require no effort of the user to use it	Efficiency, Effectiveness
	Flexibility	Degree to which a product or a service can accommodate to changes in tasks or environments	Adaptability, Interoperability
	Informativeness	Degree to which the product is informational and giving all the necessary information to the user in a proper manner	Comprehensiveness, Explicitness
	Learnability	Degree to which a product or a service is enable the user to learn how to use it	Memorability, Predictability, Consistency, Familiarity, Intuitiveness
	Simplicity	Degree to which the way that a product or a service works looks simple and uncomplicated	Modelessness
	User support	Ability of a product or a service for the user to use it easily	Helpfulness, Error prevention, Recovery, Feedback, Easy to installation
User Value	Attachment	Ability for the user to have subjective value on a product or a service by giving special meanings to it	Preciousness, Affection, Reminiscence
	Customer need	Degree to which functions of a product or a service satisfy the user's functional needs	Comfort, Convenience, Usefulness/Utility, Intelligence, Security, Trust
	Identity	Ability for the user to percept the distinct personality of an individual by using a product or a service	Self-esteem, Self-respect, Self-satisfaction
	Independence	Ability for the user to have confidence to achieve something without any aid of somebody	Self-determination, Autonomy
	Relaxation	Feeling of being relaxed or pleased by interacting with a product or a service	Pleasure, Fun, Enjoyment, Taking a rest
	Sociability	Degree to which a product/service satisfies the users' desire that they want to interact with society as a member	Social emotion, Social value, Social belonging, Relationship, Friendship
Affect	Sensory affect	Primitive and direct images by interacting with a product or a service	Shape, Color, Brightness, Sense of grip, Texture, Heaviness
	Descriptive affect	Impressions of a product or a service that the users would describe based on their experience	Delicacy, Simplicity, Rapidity, Rigidity
	Evaluative affect	Attitudinal or judgmental images about a product or a service	Attractiveness, Reliability, Comfort, Convenience

IV. Discussion

There are some differences between UX elements for general users and people with disabilities. With regard to usability, accessibility should be more emphasized for people with disabilities than for general users. People with disabilities tend to have an insufficient sensory or physical ability to use mobile phones or services, thus it is necessary to make mobile phones or services easy to use for them. In other words, accessibility is the prerequisite of usability. Accessibility has also been emphasized in studies on design, such as accessible, universal, and barrier-free design [10][11].

More diverse elements of user value were derived than those for general users. Products or services with assistive technology enable people with disabilities to do daily activities, such as reading a book, taking a walk, and making a phone call to someone, which they cannot do in their daily life, and this fact led diversity of user value elements. Specifically, independence and identity are unique and major elements of user value, which are also main elements of quality of life [12][13]. There are few studies on the affect of people with disabilities while usability was widely considered for designing products or services. However, affect should not be overlooked in the design process in the sense that people with disabilities use more diverse sense organs than general people while getting information.

This study composed the UX elements with consideration for various types of disability. However, the elements or importance of each element can differ according to types and severity of disability. For instance, a visually impaired person mostly feels affect from a sound, while a hearing impaired person mainly from a vision. Therefore, different UX elements according to types of disability should be considered before designing products or service with consideration of types of disability.

The study conducted the literature survey systematically to define and identify UX elements for people with disabilities. To verify the UX elements that identified in the study, experiments or interviews with people with disabilities should be conducted in the future study.

V. Conclusion

This study analyzed the literature on disability and UX, and identified the elements of the UX of people with disabilities. Previous studies on designing products and services focused mainly on product and service usability, but this study suggested that user value and affect should be considered in the design stage. The UX elements for those with disabilities can be used as criteria for developing new products/services or evaluating existing products/services.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2013R1A1A2013231)

References

[1] P. M. Desmet and P. Hekkert, “Framework of product experience,” International Journal of Design, vol. 1, no. 1, pp. 57-66, 2007.

[2] ISO CD 9241-210 Ergonomics of human-system interaction. Part 210: Human-centred design processfor interactive systems, ISO, 2008.

[3] J. Park and S. H. Han, “Defining and Analyzing User Value of Smartphones”, Proceedings of Spring Conference of the Ergonomics Society of Korea 2012, 2012, pp. 312-315.

[4] W. F. Preiser and K. H. Smith, Universal design handbook 2nd ed., McGraw Hill Professional, USA, 2010.

[5] ISO/IEC Guide 71 Guidelines for standards developers to address the needs of older persons and persons with disabilities, CEN/CENELEC Guide 6, ISO, 2002.

[6] The Center for Universal Design. *Environments and products for all people.* [Online]. Abilable from: http://www.ncsu.edu/ncsu/design/cud/about_ud/about_ud.htm 2014.12.22

[7] J. H. Park and H. Y. Ryoo, “Emotion of People with Visual Disability for Enhancing Web Accessibility,” Journal of Korean Emotion and Sensibility, Vol. 11, No. 4, pp. 589-598, 2008.

[8] L. Demers, R. D. Wessels, and R. Weiss-Lambrou, B. Ska, L. P. De Witte, “An international content validation of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST),” Occupational Therapy International, Vol. 6, No. 3, pp. 159-175, 1999.

[9] J. Park, S. H. Han, H. K. Kim, Y. Cho, and W. Park, “Developing elements of user experience for mobile phones and services: survey, interview, and observation approaches,” Human Factors and Ergonomics in Manufacturing & Service Industries, Vol. 23, No. 4, pp. 279–293, 2013.

[10] G. Dewsbury and M. Edge, “Designing the Home to Meet the Needs of Tomorrow,” Open House International, Vol. 26, No. 2, pp. 33-42, 2001.

[11] H. Petrie and N. Bevan, The evaluation of accessibility, usability and user experience, The universal access handbook, pp. 10-20, 2009.

[12] M. Wehmeyer and M. Schwartz, “The relationship between self-determination and quality of life for adults with mental retardation,” Education and training in mental retardation and developmental disabilities, Vol. 33, pp. 3-12, 1998.

[13] M. J. Scherer and L. A. Cushman, “Measuring subjective quality of life following spinal cord injury: a validation study of the assistive technology device predisposition assessment,” Disability & Rehabilitation, Vol. 23, No. 9, pp. 387-393, 2001.

Adaptive Content Presentation Extension for Open edX

Enhancing MOOCs Accessibility for Users with Disabilities

Sandra Sanchez-Gordon
Department of Informatics and Computer Sciences
National Polytechnic School
Quito, Ecuador
sandra.sanchez@epn.edu.ec

Sergio Luján-Mora
Department of Software and Computing Systems
University of Alicante
Alicante, Spain
sergio.lujan@ua.es

***Abstract*—In this paper, we propose a three-layer architecture to extend the Massive Open Online Courses (MOOCs) platform Open edX to enhance course content accessibility for users with disabilities. Because of their open nature and global scope, MOOCs are a great opportunity for people with disabilities that might not be able to engage in learning otherwise. The goal of the proposed extension is to enhance MOOCs' accessibility by adapting course content to student needs, preferences, skills and situations. In this approach, the user does not need to know what adaptations should be applied to the MOOC to make it more accessible for them. The user only needs to keep updated their accessibility preferences in their user profile. The extension automatically applies all the necessary adaptations as commanded by the adaptive engine and provides the presentation layer with the content best suited for the user.**

Keywords-massive open online course; MOOC; accessibility; adaptive content presentation; Open edX.

I. INTRODUCTION

With a brief six-year history, Massive Open Online Courses (MOOCs) constitute a relatively new model of e-learning. From the educational perspective, a MOOC is an online course with no entry requirements, no participation limits, and free of charge. Due to their open nature and global scope, MOOC courses can potentially benefit hundreds of thousands of diverse users. The first MOOC to get really massive was "Introduction to Artificial Intelligence", offered in 2011 by Sebastian Thrun of Stanford University and Peter Norving of Google, with an enrollment of 160,000 participants [1]. The success of this course promoted the development of more courses and the emergence of MOOC platforms to host new courses. Currently, the main MOOC platforms are Coursera (10.5 million students), edX (3 million students), Udacity (1.5 million students), MiriadaX (1 million students) and FutureLearn (800.000 students) [2].

In 2014, Bohnsack and Puhl [3] conducted a study that determined that none of the current MOOC platforms is fully accessible: most lack of correct HTML syntax (e.g., language definition, heading structure, and labels in input fields) and accessible design (e.g., clean interface, keyboard navigation, links to skip to main content). This study concluded that accessibility was not in focus when these platforms were built, thus excluding people with disabilities and not fulfilling the goal of MOOCs being open to everyone.

In 2013, edX opened sourced its platform so developers worldwide can build enhancements [4]. At the time of writing this paper, several organizations worldwide have adopted Open edX to launch their own MOOC initiatives [5]. According to Shah, Open edX is becoming the "de facto platform for organizations and groups who are looking to host their own MOOCs" [3].

Open edX provides an opportunity for developing an extension to enhance accessibility applying adaptive user interface techniques. In the literature review, regarding adaptive user interfaces for users with special needs we highlight a classical work from Stephanidis et al. [6], a follow-up work by the same author about Universal Access [7], and the works of Liu et al. [8] and Sloan et al. [9]. Regarding the developing of extensions for Open edX, we found only a proposal to develop a learning analytics extension [10]. So far, we have not been able to find research on applying adaptive user interface techniques to content delivered to MOOC users according to their particular accessibility needs. With an adequate architecture, future MOOC platforms will be able to overcome content accessibility barriers for the benefit of learners, both able and disabled.

The rest of this paper is organized as follows: Section II describes MOOC users accessibility needs. Section III explains the proposed three-layer architecture to develop an adaptive content presentation extension for Open edX. Section IV presents conclusions and future work.

II. MOOC USERS ACCESIBILITY NEEDS

Not only users with permanent disabilities (e.g., blindness, low vision, deafness, hard of hearing, motor and cognitive issues) can benefit from accessible MOOCs. All of us could potentially experiment a temporary or environmental disability at some point in our lives. For example, difficulties distinguishing colors due to lighting conditions (low vision) and impediments to hear due to lack of headphones in noisy environments (low hearing) or must-be-quiet places such as hospitals or libraries (deafness). Moreover, most of us will develop combined disabilities as we age naturally. Also, in the context of a MOOC, people learning in a language different from their own might face difficulties due to their level of proficiency in the course language (e.g., non-native speakers read at slower speed, which leads to information overload and cognitive issues)

[11]. This is significant, since 2014 statistics from the MOOC aggregator Class Central indicates that 80% of MOOCs are offered in English and in distant second place is Spanish with 8.5% [3]. Finally, the typical age range of MOOC students is 16 to 88 years old, with a growing tendency of elderly users engaging in lifelong learning for intellectual stimulation and social engagement. Elderly students face several accessibility barriers to access course content due to diminishing capacities such as vision decline, hearing loss, decremented motor skills and cognition issues [12].

To promote accessibility, the World Wide Web Consortium (W3C) created the Web Accessibility Initiative (WAI) to develop guidelines for web content (WCAG), authoring tools (ATAG), and browsers and other user agents (UAAG) [13]. These guidelines are a good starting point to understand users' accessibility needs.

Although currently there are accessibility options in most operating systems, special-purpose applications and assistive technologies for several disabilities, most of them require the user to explicitly invoke them. Also, the potential negative psychological effects caused by the introduction of an assistive technology that change how a user interacts with a computer may lead to the user rejection of that assistive technology or computer use altogether [7][9]. In this work, we propose meeting the MOOC users accessibility needs with an approach based in user profiling and the use of questionnaires that combines explicit user-invoked adaptations with automatic adaptations.

III. Adaptive Content Presentation Arquitecture

Adaptive content presentation involves personalizing the contents delivered to the user to enhance their accessibility and usability [6]. To successfully achieve this, it is necessary an accurate detection of the user accessibility needs through user profiling and a mechanism that allows transparent selection and presentation of the appropriate adaptations according to the registered needs [9].

The proposed extension will provide functionality for both MOOC authors and users. On one hand, the solution will allow course authors using the extended platform to configure parameters and define features so the course can adapt to diverse potential learners. On the other hand, the extended platform will allow users of MOOCs hosted in it to manage their accessibility user profile by selecting a combination of accessibility issues that best suit their current life situation and optionally taking quick questionnaires to define specific accessibility preferences (e.g., text size, color contrast, line spacing). The use of an accessibility user profile represents an improvement compared with current approaches used in websites and web applications, where the user must manually select specific technical adaptations. Nevertheless, the Accessibility Preferences user interface has an "Advanced Options" feature that provides more savvy users with freedom to select specific adaptations if desired. Figure 1 shows a user interface prototype to select accessibility preferences.

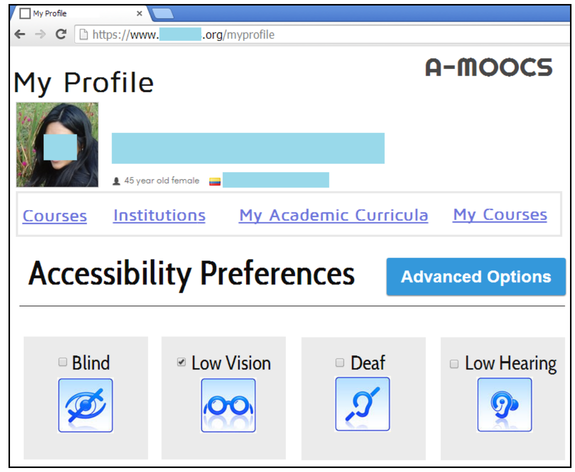


Figure 1. User interface prototype to select accessibility preferences

In this approach, the user does not need to know what adaptations should be applied to the MOOC to make it more accessible for them. User needs to know only their reality and keep it updated in their profile. From that, the proposed adaptive content presentation extension automatically applies all the necessary adaptations.

The architecture of the adaptive content presentation extension is composed of three layers, as illustrated in Figure 2.

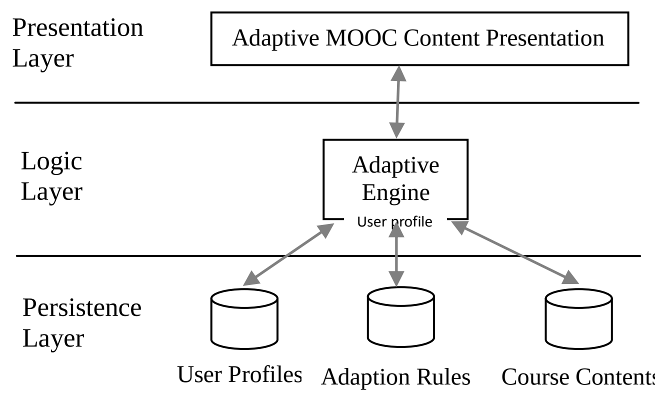


Figure 2. Three-Layer Architecture for the Extension

The Presentation Layer receives the course content in the appropriate format from the Logic Layer, where the adaptive engine resides. To select the appropriate content format, the adaptive engine scans the user profile and applies the necessary adaption rules, as explained by Stephanidis et al. [6]. Figure 3 shows an extract of the adaption rule sequence to be executed if the user accessibility preferences profile indicates dyslexia, based on the guideline developed by De Santana et al. [14]

```
FOR UserPreference[i]
{IF UserPreference[i] EQUALS dyslexia THEN LineSpacing=1.5
AND TextFont=Serif.Arial AND TextSize=12 AND
TextJustification=Unjustified AND …} NEXT i
```

Figure 3. Extract of adaption rule sequence for dyslexia

The Persistence Layer contains three databases for storing user profiles, adaptation rules and course contents. The course contents' database must contain several alternative formats for the same content, as illustrated in Figure 4. The user must be able to access any of the available alternative formats for any course content if desired.

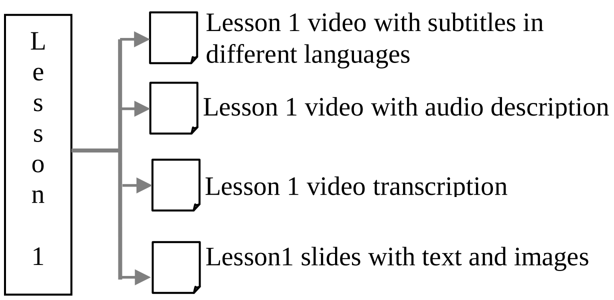


Figure 4. Example of alternative formats for same course content

Also, the extended platform must support different use scenarios, such as a student with combined disabilities (e.g., elderly student) or several students with different disabilities accessing the MOOC together. Figure 5 shows a lesson adapted to users with dyslexia, low vision and blindness.

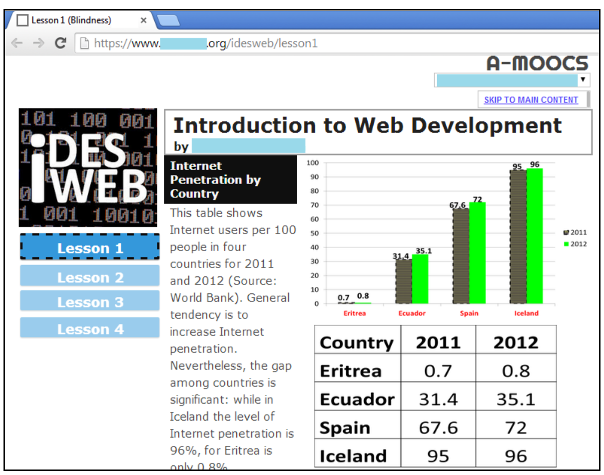


Figure 5. Lesson interface adapted for dyslexia, low vision and blindness

In this case, three adaption rules sequences have been applied: the dyslexia rule sequence explained earlier; the low vision rule sequence (i.e. text with contrast ratio of 4.5:1, border pattern to indicate current focus, data on top of graph bars); and the blind rule sequence (i.e. link to skip to main content, graph with alternative text, table readable by a screen reader containing the graph data).

IV. Conclusion and Future Work

Software solutions must adapt to users, not the other way around. This is the best cost-effective way to design a solution, especially in scenarios where large numbers of diverse users are expected, as is the case in MOOCs. That is why, in this paper, we proposed an adaptive content presentation extension for Open edX that will allow MOOCs hosted in this extended platform to adapt to the specific needs, preferences, skills, and situations of learners, both able and disabled. More important, with this extended platform the particular situation of a person with disabilities may go unnoticed for both instructors and peer students, so the person with disabilities can be treated truly equally, hence assuring real inclusiveness.

Continuing research is essential for increasing accessibility of MOOCs. As future work, we intend to develop a detailed design of the proposed architecture, create a prototype of the extension for Open edX, and perform both user-based and expert-based evaluations.

References

[1] S. Sanchez-Gordon and S. Luján-Mora, "MOOCs gone Wild", in Proceedings of the 8th International Technology, Education and Development Conference, March 2014, pp. 1449-1458.

[2] W. Shah, MOOCs in 2014: Breaking Down the Numbers. edSurge. [Retrieved: December 2014] https://www.edsurge.com/n/2014-12-26-moocs-in-2014-breaking-down-the-numbers

[3] M. Bohnsack and S. Puhl, "Accessibility of MOOCs", In Computers Helping People with Special Needs, Springer International Publishing, 2014, pp.141-144.

[4] edX, The Open edX, 2014. [Retrieved: December 2014] http://code.edx.org/

[5] gitHub, Sites powered by Open edX, 2014. [Retrieved: December 2014] https://github.com/edx/edx-platform/wiki/Sites-powered-by-Open-edX

[6] C. Stephanidis et al., "Adaptable and adaptive user interfaces for disabled users in the AVANTI project", in Intelligence in Services and Networks: Technology for Ubiquitous Telecom Services, Springer, 1998, pp. 153-166.

[7] C. Stephanidis, "Adaptive techniques for universal access", in User Modeling and User-Adapted Interaction, 11 (1-2), 2001, pp. 159-179.

[8] J. Liu, C. K. Wong, and K. Hui, "An adaptive user interface based on personalized learning", in IEEE Intelligent Systems, 18 (2), 2003, pp. 52-57.

[9] D. Sloan, M. Atkinson, C. Machin, and K. Li, "The potential of adaptive interfaces as an accessibility aid for older web users", in Proceedings of the International Cross Disciplinary Conference on Web Accessibility, ACM, April 2010, doi: 10.1145/1805986.1806033

[10] J. Ruiz, H. Pijeira Díaz, J. Ruipérez-Valiente, P. Muñoz-Merino, and C. Delgado Kloos, "Towards the development of a learning analytics extension in open edX", in Proceedings of the 2nd. International Conference on Technological Ecosystems for Enhancing Multiculturality, ACM, October 2014, pp. 299-306, doi: 10.1145/2669711.2669914

[11] S. Sanchez-Gordon and S. Luján-Mora, "Web Accessibility Requirements for Massive Open Online Courses", in Proceedings of 5th International Conference on Quality and Accessibility of Virtual Learning, May 2014, pp.530-535.

[12] S. Sanchez-Gordon and S. Luján-Mora, "Web accessibility of MOOCs for elderly students", in Proceedings of the 12th International Conference on Information Technology Based Higher Education and Training, October 2013, pp. 1-6, doi: 10.1109/ITHET.2013.6671024

[13] W3C. "WAI Guidelines and Techniques", 2011. [Retrieved: January 2015] http://www.w3.org/WAI/guid-tech.html

[14] V.F. de Santana, R. de Oliveira, L.D.A. Almeida, and M.C.C. Baranauskas, "Web accessibility and people with dyslexia: a survey on techniques and guidelines", in Proceedings of the International Cross-Disciplinary Conference on Web Accessibility, ACM, April 2012, doi:10.1145/2207016.220704

Expressive Humanoid Face: a Preliminary Validation Study

Nicole Lazzeri, Daniele Mazzei, Alberto Greco,
Antonio Lanatà and Danilo De Rossi

Research Center "E. Piaggio"
University of Pisa
Pisa, Italy
Email: daniele.mazzei@centropiaggio.unipi.it

Annalisa Rotesi

Faculty of Psychology
University of Florence
Florence, Italy

***Abstract*—Non-verbal signals expressed through body language play a crucial role in our daily communications. Facial expressions, in particular, are the most universal signs to express innate emotional cues. Human faces convey important information in social interactions, which help us to better understand our interlocutor. Nowadays, humanoids and social robots are becoming increasingly similar to humans both aesthetically and expressively. However, their visual expressiveness is a crucial issue in making these robots more realistic and intuitively perceived as human-like. This paper presents a preliminary study aimed at evaluating the capability of a humanoid to perform facial expressions in terms of recognition rate and response time in comparison with humans' ability. Results showed that the recognition rate of human and robot expressions did not reveal differences while the physical robot can convey expressions better than its 2D photos and its 3D models. Moreover, the results showed that both human and robot positive expressions were better recognized than the negative ones.**

Keywords–Facial expressions; emotion perception; humanoid robot; expression recognition; social robots.

I. Introduction

Human beings communicate in a rich and sophisticated way through different channels, i.e., sound, vision, touch and smell. In particular, in human social relationships visual information plays a crucial role. Human faces convey important information both from static features such as identity, age and gender, and from dynamic changes such as expressions, eye blinking and muscular micro-movements. The ability to recognize and understand facial expressions of an interlocutor allows us to establish and manage empathic links, which drive our social relationships.

Charles Darwin was the first to observe that basic expressions such as anger, disgust, contempt, fear, surprise, sadness, happiness, are universal and innate [1]. Indeed, human beings are able to recognize faces and read facial expressions almost unconsciously and with little or no effort [2].

In the last years due to rapid advances in robotics and computer graphics, more and more interactive robots and agents have become common in our daily lives. The rapid growth of robotics has made possible the development of a new class of emphatic machines known as social robots. These innovative agents are used in various fields ranging from entertainment to human assistance and health care [3]. Hashimoto and his collegues [4] proposed an android robot called SAYA as a teacher in elementary and university classes. SAYA was remotely controlled by an opertor makeing the robot able to perform facial expressions, head and eye movements, and utterances. The experimental studies showed that SAYA was more accepted by elementary school children than by university students enhancing their interests for the class.

The ability to express emotions is clearly becoming fundamental for a social robot's believability [5] driving the research on the design of user-friendly social robots able to reproduce human-like facial expressions [6]. Costa and her team [7] presented a perceptual study with ZECA, a robotic child able to perform facial expressions and gestures. ZECA was used in human-robot interaction studies with children with autism demonstrating their capability to recognize the robot expressions.

Humanoids and social robots are usually high-cost products typically used in academia and research fields only. On the other hand virtual avatars are widely used as social characters for games, storytelling and tutoring [8]. However, humanoids and avatars differ in a fundamental aspect: the *embodiment*.

This work is based on the hypothesis that highly anthropomorphic robots with physical embodiment are able to convey expressions and socially interact easier and more intuitively than avatars and 3D models [9][10]. Indeed, the embodiment could help robots to express their emotions by means of a physical and real aspect, which is absent in a screen.

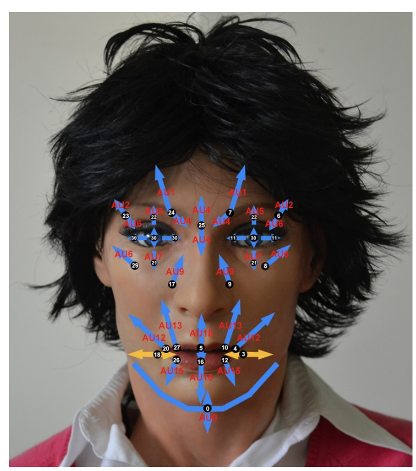


Figure 1. The FACE robot with references of the servo motor positions and of the corresponding FACS AUs.

Current research literature aims at evaluating facial expressions on robots and virtual avatars with different approaches

 ISBN: 978-1-61208-382-7

[11][12][13]. Our work aimed at studying the capability of a realistic humanoid robot to show facial expressions in comparison with 2D pictures and 3D models of itself and of a female human. Our robot was built with a female appearance in order to take advantage of the higher expressivity of female expressions as demonstrated by Adolph et. al. [14]. In our experiments, participants were asked to evaluate three set of facial expressions through questionnaires, as previously done by Becker-Asano and Ishiguro [15] and in other similar works [16][17] and their answers were evaluated in terms of recognition rate and response time.

This paper is structured as follows: Section II describes the material used to create the stimuli for the experiment; Section III presents the method and the protocol of the experiment and its setup; Section IV explains the statistical analysis and the related results about the facial expression recognition; in the end, Section V summarizes the results of the experiment drawing a general conclusion.

II. MATERIALS

The material used for the experiment included various stimuli: the FACE (Facial Automaton for Conveying Emotions) robot; the FACE robot avatar and a set of 2D and 3D pictures of a women performing facial expressions.

A. The robot FACE

FACE is an android female face used to study human-robot interactions with a focus on non-verbal communication, developed in collaboration with Hanson Robotics [18][19][20]. FACE consists of a passive body with a realistic facial system made of an artificial skull covered by a porous elastomer called Frubber™. FACE is animated by 32 servo motors positioned inside the skull and in the upper torso (Figure 1).

In this study the attention was focused on recognizing the six basic emotions considered as 'universally accepted' by Paul Ekman [21], i.e., happiness, sadness, anger, fear, disgust and surprise. The FACE's original facial expressions were manually created using the Hybrid Engine for Facial Expression Synthesis module (HEFES) [19] following anatomical facial expressions guidelines (Artanatomia) [22].

To standardize the methodology for creating the FACE's facial expressions, we adopted the Facial Action Coding System (FACS) developed by Ekman and Friesen [23]. Using FACS, a facial expression can be decomposed into Action Units (AUs), which are defined as observable independent facial movements. The FACE's servo motors are positioned similarly to the major facial muscles therefore it is possible to find a correspondence between them and the AUs (Figure 1).

B. Synthetic 2D and 3D stimuli

The stimuli chosen for the experiment were 2D images and 3D models of the robot FACE and of a female human.

An image of each of the 6 basic emotions plus the neutral face was used to create the set of 7 2D photos for FACE. The set of 7 3D models of FACE was created using the Autodesk 123D catch® program, which generates 3D models taking as inputs one hundred photos acquired moving around the robot from the left to the right side covering about 180°.

The set of human 2D photos and 3D models was taken by selecting a female subject (item bs103) from the Bosphorus Database [24], a 2D/3D collection of FACS-based facial expressions acquired using a structured-light 3D scanner [25]. Figure 2 shows 2D photos and 3D models of FACE and the human subject used in this experiment.

Figure 2. The stimuli used in the experiment: 2D photos (first row) and 3D models (second row) of (a) FACE expressions and (b) human expressions

Due to technical problems of a servo motor corresponding to the buccinator muscle (motor n. 4), FACE was partially enable to raise the left part of the smile obtaining an ambiguous happiness expression. Motor n. 4 is used only in the happyness expression, consequently we excluded the data relative to the happyness from the analysis.

III. METHOD

A. Experimental setup

Participants were seated comfortably at a desk about $0.5m$ far either from a TV screen (Size: 32 inch, Frame rate: 100Hz, Resolution: 1920 x 1080) or the robot. The experiment setup included one laptop for controlling the robot FACE and one laptop for controlling the animation on the TV screen.

10 participants (7 males, 3 females) aged $19 - 31$ years (mean age 24.1 ± 3.4) were recruited for the experiment. All participants attended scientific disciplines at University of Pisa (IT), were native Italian speakers and had either normal or corrected to normal vision. All participants gave written informed consent for the experiment.

B. Experimental Protocol

The protocol of the experiment was organized in 3 phases:

- **First phase:** each participant had to recognize 14 2D photos of facial expressions: 7 photos of humans from the Bosphorus database and 7 photos of FACE, in random order (different for each participant);
- **Second phase:** each participant had to recognize 14 3D models of facial expressions: 7 3D models of humans from the Bosphorus database and 7 3D models of FACE, in random order (different for each participant);
- **Third phase:** each participant had to recognize 6 basic expressions performed by the robot FACE in random order (different for each participant).

TABLE I. CONFUSION MATRIX OF THE RECOGNITION RATES (IN PERCENTAGE) OF SEVEN (FOR HUMANS) AND SIX (FOR THE ROBOT) FACIAL EXPRESSIONS WITH PRESENTED MODELS (COLUMNS) AGAINST SELECTED LABELS (ROWS). THE HIGHEST VALUES ARE SET IN BOLD. THE COLUMN LABELS ARE A=ANGER, D=DISGUST, F=FEAR, N=NEUTRAL, SA=SADNESS AND SU=SURPRISE.

Confusion matrix (N=10)																	
	Human 2D photos						**Human 3D models**						**Physical robot**				
	A	**D**	**F**	**N**	**Sa**	**Su**	**A**	**D**	**F**	**N**	**Sa**	**Su**	**A**	**D**	**F**	**Sa**	**Su**
Anger	20	10	0	0	0	0	**40**	0	0	0	0	0	**50**	20	0	0	0
Disgust	0	**40**	0	0	10	0	0	**30**	0	0	0	0	10	**50**	10	0	10
Fear	0	0	30	0	0	0	10	**30**	30	0	0	0	10	0	**60**	0	10
Neutral	0	0	0	**80**	0	0	0	0	0	**90**	0	0	/	/	/	/	/
Sadness	0	20	0	10	**20**	0	0	0	0	10	**40**	0	0	0	0	**50**	0
Surprise	0	0	**60**	0	0	**100**	0	0	**60**	0	0	**90**	0	0	20	0	**80**
Pride	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Embarrassment	0	10	0	0	0	0	0	20	10	0	20	0	0	0	0	0	0
Pain	10	20	0	0	**20**	0	0	10	0	0	30	0	10	0	0	10	0
Pity	0	0	0	0	**20**	0	10	0	0	0	0	0	0	10	0	40	0
Contempt	**30**	0	0	10	10	0	10	0	0	0	0	0	20	20	0	0	0
Interest	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Shame	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
Excitement	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
I do not know	30	0	10	0	**20**	0	30	10	0	0	0	0	0	0	10	0	0
No answer	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

The Unity 3D® software was used as front end animation tool to show both 2D photos and 3D models. In each phase, the participant had at most 30 seconds to recognize the expression. A C# program was developed to orchestrate the expression times as follow: at most 10 seconds was given to observe the expression followed by at most 20 seconds to answer. In the first and second phase, after the first 10 seconds or whether the participant pressed “Enter” on the keyboard (before the end of the first 10 seconds), i.e., the participant was ready to give an answer, a black screen appeared on the TV. After pressing “Enter” the participant had to choose one of the possible answers of the questionnaire (listed in Table I). The response time was also recorded on the “Enter” key pressing.

In the third phase, after the first 10 seconds or whether the participant selected the answer on the screen (before the end of the first 10 seconds), i.e., the participant was ready to answer, the robot performed the neutral expression. In this case, the participant evaluated 6 instead of 7 different facial expressions since the neutral expression was used as “black screen”. To answer, the participant had to select an option directly on the screen through a software tool running on a laptop. The response time was recorded on the mouse clicking.

IV. Data analysis

The set of expressions considered in the analysis included anger, disgust, fear, sadness and surprise. As mentioned in Sec. II-B, the happiness expression of the robot was ambiguous due to technical problems with a servo motor therefore it was excluded from all datasets in the data analysis.

The facial expression recognition rates were analyzed using the Cohen’s kappa [26], a statistical measure of inter-rater reliability used to examine the agreement between observers on the assignment of categories of a categorical variable. The Cohen’s kappa ranges from -1.0 to 1.0, where large numbers mean better reliability, values near zero suggest that agreement is attributable to chance, and values less than zero signify that agreement is even less than that which could be attributed to chance. According to Landis and Koch [27], with a significance level of 0.05, kappa can be classified according to the following: $k <= 0.00$ less than chance agreement, $0.01 < k < 0.20$ slight agreement, $0.21 < k < 0.40$ fair agreement, $0.41 < k < 0.60$ moderate agreement, $0.61 < k < 0.80$ substantial agreement, and $0.81 < k <= 1$ almost perfect agreement.

The Kolmogorov-Smirnov test [28] and the analysis of variance were applied to the datasets of the response times. The Anova-1way parametric test with a post-hoc Bonferroni test [29] was used to examine the category differences. The statistical inference was carried out using the OriginLab software [30].

A. *Are facial expressions of the robot perceived as well as the expressions of humans (expressed as 2D photos or 3D models)? Yes*

Table I shows the confusion matrix of the participants’ answers for 2D human photos, 3D human models and robot FACE’s expressions. The Cohen’s kappa of the three categories showed a homogeneous expression evaluation with the best level of agreement for the expressions performed by the physical robot: $K_{Hum2D} = 0.570$ ($p < 0.001$) 95% CI (0.350, 0.789), $K_{Hum3D} = 0.606$ ($p < 0.001$) 95% CI (0.401, 0.809) and $K_{Robot} = 0.701$ ($p < 0.001$) 95% CI (0.530, 0.871). For all three categories, the best recognition rate was achieved for the surprise expression. Human anger, fear and sadness were not so well understood when shown both as 2D photos and 3D models.

Figure 3 shows a trend of increasing recognition rate for stimuli that gradually become more realistic, i.e., from human 2D photos to human 3D models up to the physical robot. This supports our hypothesis about the importance of the embodiment in conveying expressions.

All participants were instructed to choose a label for each expression as soon as they recognized it and their response time was recorded. Table II shows the means and the standard deviations of the response time for each expression in the three categories: human 2D photos and 3D models and FACE robot.

The Anova-1way parametric test did not find significant differences between the three categories (F(2,68) = 0.55309,

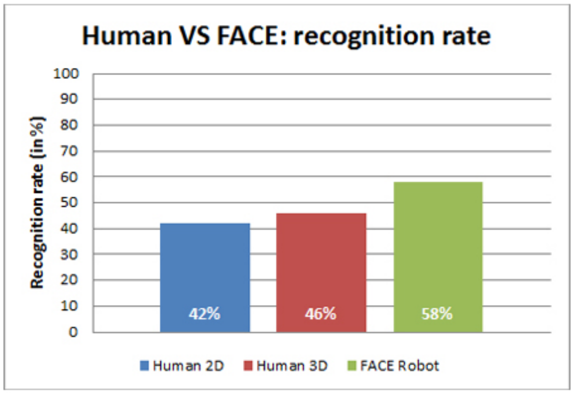


Figure 3. Recognition rate (%) of the human 2D photos, human 3D models and robot FACE.

TABLE II. MEANS AND STANDARD DEVIATIONS OF THE PARTICIPANTS' RESPONSE TIME (S) IN RECOGNIZING FACIAL EXPRESSIONS OF THE HUMAN AND THE ROBOT FACE.

	Response time in seconds (N = 10)					
	Human 2D		**Human 3D**		**Robot**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Anger	4.09	0.60	9.53	5.94	8.53	4.43
Disgust	8.52	4.04	10.79	3.30	12.32	6.78
Fear	9.71	6.90	10.25	1.34	9.02	3.66
Sadness	19.31	7.84	10.90	6.53	16.02	0.96
Surprise	8.13	5.99	9.13	3.39	9.65	4.93

p = 0.57774, $\alpha = 0.05$) (Figure 4). This data confirm that the facial expressions performed by FACE were perceived similar to human expressions.

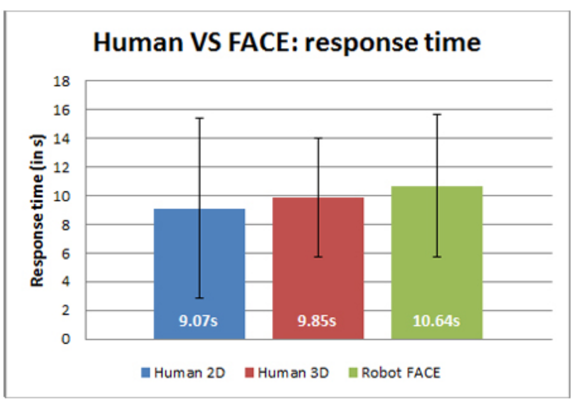


Figure 4. Response time (s) of the human 2D photos, human 3D models and robot FACE expression recognition.

B. *Is there a valid and useful reason to create and develop a realistic humanoid robot instead of using its 2D photos or its 3D models? Yes*

Previous results demonstrated that FACE can convey expressions that are recognized with a similar rate of human 2D photos and 3D models. Moreover, our study tried to investigate if the embodiment of the FACE robot is an added value, which could help people to better understand and interpret the expressed emotional status.

Table III shows the confusion matrix of the participants' answers for the robot expressions shown as 2D photos, 3D models and physical robot. As in the previous case, for all three categories, the best recognition rate was achieved for the surprise expression. The expressions performed by the robot were less confused than those shown in 2D photos or 3D models. The level of agreement between the participants was comparable for the three categories of the facial expressions performed by the FACE robot with the best level of agreement for stimuli performed by the physical robot as in the previous case: $K_{FACE2D} = 0.519$ ($p < 0.001$) 95% CI (0.284, 0.752), $K_{FACE3D} = 0.604$ ($p < 0.001$) 95% CI (0.375, 0.832) and $K_{Robot} = 0.701$ ($p < 0.001$) 95% CI (0.530, 0.871).

A comparison between the robot stimuli that gradually become more realistic, i.e., from 2D photos to 3D models up to the physical robot, shows a trend of increasing recognition rate for stimuli performed by the physical robot (Figure 5). This suggests that the embodiment of the robot conveys the expressions better than 2D photos and 3D models.

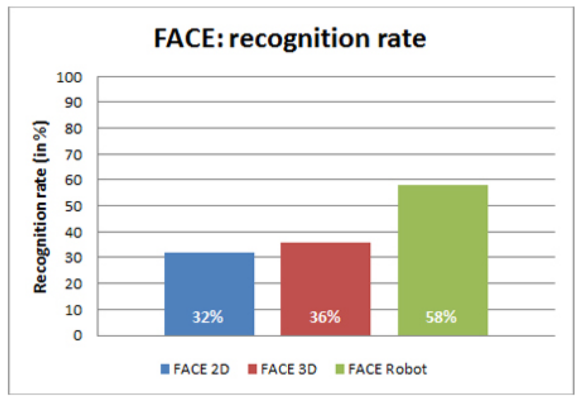


Figure 5. Recognition rate (%) of the robot 2D photos, robot 3D models and robot FACE expressions.

Table IV shows the response time means and standard deviations for each expression in the different categories of FACE: 2D photos, 3D models and the physical robot.

TABLE IV. MEANS AND STANDARD DEVIATIONS OF PARTICIPANTS' RESPONSE TIME (S) IN RECOGNIZING FACIAL EXPRESSIONS OF ROBOT 2D PHOTOS, ROBOT 3D MODELS AND THE PHYSICAL ROBOT. (* Only one response.)

	Response time in seconds (N = 10)					
	FACE 2D		**FACE 3D**		**Robot**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Anger	16.49	11.20	6.17	*	8.53	4.43
Disgust	12.31	13.89	7.15	3.35	10.55	5.78
Fear	9.64	7.27	13.36	9.11	9.68	3.87
Sadness	13.87	3.29	11.73	6.82	16.02	0.96
Surprise	13.63	3.17	9.93	4.08	9.31	4.20

The Anova-1way parametric test could not distinguish between the three distributions (F(2,57) = 1.66754, p = 0.19778, $\alpha = 0.05$). The time for recognizing an expression performed by the robot was comparable to the one required to recognize the same expression as 2D photo or 3D model (Figure 6).

Figure 7 shows that the surprise expression achieved the best recognition rate in comparison with each negative expression with a difference of at least of 25%.

TABLE III. CONFUSION MATRIX OF THE RECOGNITION RATES (IN PERCENTAGE) OF ROBOT FACIAL EXPRESSIONS WITH PRESENTED MODELS (COLUMNS) AGAINST SELECTED LABELS (ROWS). THE HIGHEST VALUES ARE SET IN BOLD. THE COLUMN LABELS ARE A=ANGER, D=DISGUST, F=FEAR, N=NEUTRAL, SA=SADNESS AND SU=SURPRISE.

Confusion matrix (N=10)																	
	FACE 2D photos						FACE 3D models						Physical robot				
	A	**D**	**F**	**N**	**Sa**	**Su**	**A**	**D**	**F**	**N**	**Sa**	**Su**	**A**	**D**	**F**	**Sa**	**Su**
Anger	20	10	10	0	0	0	10	20	0	0	0	0	**50**	20	0	0	0
Disgust	**30**	30	**20**	0	0	20	30	**40**	0	0	0	0	10	**50**	10	0	10
Fear	10	0	**20**	0	0	0	0	0	30	0	0	0	10	0	**60**	0	10
Neutral	0	0	0	**40**	0	0	0	0	0	**40**	0	0	/	/	/	/	/
Sadness	0	0	0	0	30	0	0	0	0	0	**30**	0	0	0	0	**50**	0
Surprise	0	0	0	10	0	**60**	10	0	20	0	0	**70**	0	0	20	0	**80**
Pride	0	0	0	30	0	0	0	0	0	20	0	0	0	0	0	0	0
Embarrassment	0	0	10	0	10	0	0	0	10	0	0	0	0	0	0	0	0
Pain	0	0	0	0	0	0	0	0	0	0	10	0	10	0	0	10	0
Pity	0	0	10	0	**40**	10	0	0	**40**	0	10	0	0	10	0	40	0
Contempt	20	**50**	10	10	10	10	**40**	**40**	0	10	0	0	20	20	0	0	0
Interest	0	0	0	10	10	0	10	0	0	20	10	10	0	0	0	0	0
Shame	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Excitement	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
I don't know	10	10	10	0	0	0	0	0	0	10	**30**	20	0	0	10	0	0
No answer	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

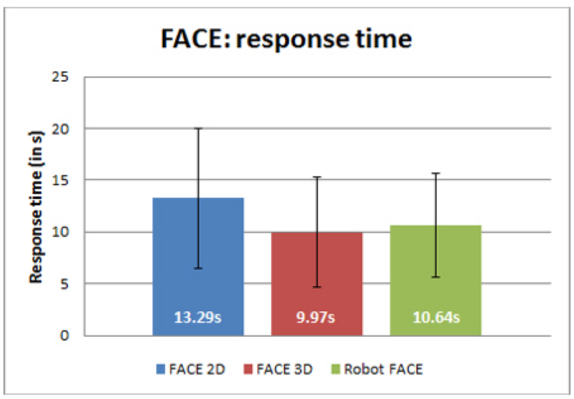


Figure 6. Response time (s) in recognizing the facial expressions of FACE 2D photos, 3D models and physical robot.

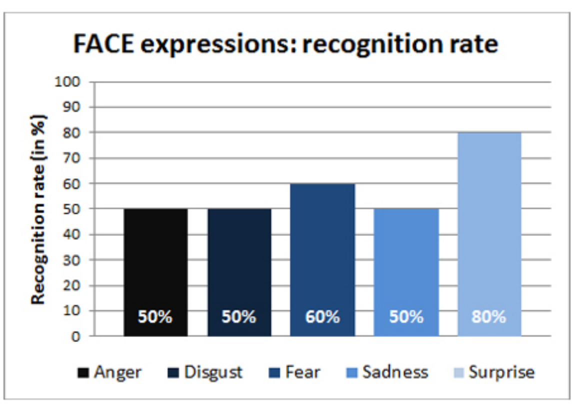


Figure 7. Recognition rate (%) of the FACE expressions.

V. CONCLUSION

Our study aimed at investigating (1) if the recognition rate of facial expressions performed by FACE were similar to the ones achieved with humans stimuli and (2) if there were differences in recognizing facial expressions performed by FACE using its 2D photos, 3D models or the robot itself.

The final dataset used in the analysis did not include the happiness expression because it was considered ambiguous due to an abnormal functioning of a servo motor.

In regard to the first question, our preliminary results demonstrate that the recognition rate of human expressions is similar to the one of the robot expressions. This supports our hypothesis that the robot is able to convey emotion through facial expressions as well as human 2D photos and 3D models.

Concerning the second question, we found that the physical robot can convey expressions better than its photos and its 3D models. We could hypothesize that this study support the direction of the contention that the dynamic and the embodiment of social humanoids improve the recognition and discrimination of emotions in comparison with 2D pictures and 3D displays [31][32][33].

Usually positive expressions may not require the analysis of the entire face to be recognized since they can be characterized by a single feature, such as a smiling mouth for the happiness [34]. This phenomenon makes the recognition of the expression simpler and then faster [35]. Our results confirmed this phenomenon. A comparison between the recognition rates of 2D and 3D human expressions showed that the surprise expression was generally recognized better than the negative expressions. Indeed, anger was often confused with contempt or not recognized at all, disgust was confused with fear or pain while fear with surprise and sadness with pity or pain. Even in the case of 2D and 3D robot expressions, the best recognition rate was achieved for surprise while anger and disgust were often confused with disgust or contempt, fear with disgust or pity and sadness with pity or not recognized at all.

In conclusion, we based our experiment on the hypothesis that the embodiment of highly anthropomorphic robots could help them to express their emotions by means of a physical aspect, which is absent in a virtual character on a screen. Our results found that there is a general tendency to better recognize expressions performed by the physical robot than the ones shown as 2D photos and 3D models. The embodiment of the robot and its dynamics could be an added value to help people to better understand and interpret the emotional status of a robot.

This work represents a preliminary study of the emotion conveying capability of our robot and its results are encouraging for future experiments. These results highlighted that generating facial expressions is a challenging task that requires high-fidelity reproduction therefore future developments will concern improving the performance of the robot in expressing emotions. In addition to the exclusion of the expression of happiness, two factors may have influenced the statistical analysis: the small size of the sample and the extended forced-choice paradigm. Thus, new effective experimental tests will be designed to be more effective. Moreover, this study give us the foundations for the setup of a therapeutic scenario in which the FACE robot will be used as emotional display in the autism treatment.

ACKNOWLEDGEMENTS

We would like to thank Andrea Guidi, and Niccolò Albertini, for thei help during the preparation of the experiments. This work was partially founded by the European Commission within the Project EASEL (Expressive Agents for Symbiotic Education and Learning) (FP7-ICT-611971).

REFERENCES

[1] C. Darwin, The Expression of the Emotions in Man and Animals, P. Ekman, Ed. London: Harper Collins, Feb. 1998, ISBN: 0-00-255866-1.

[2] J. E. LeDoux, The emotional brain. New York: Simon and Shuster, 1996.

[3] H. I. Christensen and et al., A Roadmap for U.S. Robotics From Internet to Robotics. Computing Community Consortium, Mar. 2013.

[4] T. Hashimoto, N. Kato, and H. Kobayashi, “Development of educational system with the android robot saya and evaluation,” International Journal of Advanced Robotic Systems. Special Issue Assistive Robotics, vol. 8, 2011, pp. 51–61.

[5] B. Siciliano and O. Khatib, Springer Handbook of Robotics, B. Siciliano and O. Khatib, Eds. Springer, 2008, ISBN: 978-3-540-23957-4.

[6] C. Pèlachaud, “Some considerations about embodied agents,” 2000.

[7] S. Costa, F. Soares, and C. Santos, “Facial expressions and gestures to convey emotions with a humanoid robot,” in Social Robotics, ser. Lecture Notes in Computer Science, G. Herrmann, M. Pearson, A. Lenz, P. Bremner, A. Spiers, and U. Leonards, Eds. Springer International Publishing, 2013, vol. 8239, pp. 542–551, ISBN: 978-3-319-02674-9.

[8] R. Gockley, J. Forlizzi, and R. Simmons, “Interactions with a moody robot,” in Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, ser. HRI ’06. New York, NY, USA: ACM, 2006, pp. 186–193, ISBN: 1-59593-294-1.

[9] D. Hanson, “Why we should build humanlike robots,” 2001.

[10] U. staff, “A ubiquity interview with david hanson,” Ubiquity, vol. 2006, no. May, 2006, p. 1, ISSN: 1530-2180.

[11] M. Moosaei and L. Riek, “Evaluating facial expression synthesis on robots,” in Proceedings of the HRI Workshop on Applications for Emotional Robots at the 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2013.

[12] M. Blow, K. Dautenhahn, C. L. Nehaniv, D. C. Lee, and A. Robotcub, “Perception of robot smiles and dimensions for human-robot interaction design,” in Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06). IEEE Press, 2006, pp. 469–474.

[13] J. Saldien, K. Goris, B. Vanderborght, J. Vanderfaeillie, and D. Lefeber, “Expressing Emotions with the Social Robot Probo,” International Journal of Social Robotics, Aug 2010, pp. 1–13, ISSN: 1875-4791.

[14] D. Adolph and W. A. GEORG, “Valence and arousal: a comparison of two sets of emotional facial expressions,” American Journal of Psychology, vol. 123, no. 2, 2010, pp. 209–219.

[15] C. Becker-Asano and H. Ishiguro, “Evaluating facial displays of emotion for the android robot geminoid f,” in 2011 IEEE Workshop on Affective Computational Intelligence (WACI), 2011, pp. 1–8.

[16] C. Bartneck, J. Reichenbach, and A. Van Breemen, “In your face, robot! the influence of a character’s embodiment on how users perceive its emotional expressions,” in Confrence on Design and Emotion, 2004.

[17] C. D. Kidd and C. Breazeal, “Effect of a robot on user perceptions,” in Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004, pp. 3559–3564.

[18] D. Hanson, “Exploring the aesthetic range for humanoid robots,” in Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science, (Vancouver, Canada, Jul. 2006, pp. 16–20.

[19] D. Mazzei, N. Lazzeri, D. Hanson, and D. De Rossi, “HEFES: a Hybrid Engine for Facial Expressions Synthesis to control human-like androids and avatars,” in Proceedings of 4th IEEE RAS EMBS International Conference on Biomedical Robotics and Biomechatronics, BIOROB, 2012, pp. 195–200, ISSN: 2155-1774.

[20] D. Mazzei, L. Billeci, A. Armato, N. Lazzeri, A. Cisternino, G. Pioggia, R. Igliozzi, F. Muratori, A. Ahluwalia, and D. De Rossi, “The FACE of autism,” in The 19th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2010). IEEE Computer Society publisher, 2010, pp. 791–796, ISSN: 1944-9445.

[21] P. Ekman, “Are there basic emotions?” Psychological Review, vol. 99, no. 3, Jul. 1992, pp. 550–553.

[22] “ARTNATOMY,” 2014, URL: http://www.artnatomia.net/uk/artnatomy2014.html [accessed: 2014-12-29].

[23] P. Ekman and W. V. Friesen, “Measuring facial movement,” Journal of Nonverbal Behavior, vol. 1, no. 1, Sep. 1976, pp. 56–75, ISSN: 0361-3496.

[24] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, and L. Akarun, in Biometrics and Identity Management, B. Schouten, N. C. Juul, A. Drygajlo, and M. Tistarelli, Eds. Berlin, Heidelberg: Springer-Verlag, 2008, ch. Bosphorus Database for 3D Face Analysis, pp. 47–56, ISBN: 978-3-540-89990-7.

[25] A. Savran, B. Sankur, and M. Taha Bilge, “Comparative evaluation of 3d vs. 2d modality for automatic detection of facial action units,” Pattern Recognition, vol. 45, no. 2, Feb. 2012, pp. 767–782, ISSN: 0031-3203.

[26] J. Cohen, “A Coefficient of Agreement for Nominal Scales,” Educational and Psychological Measurement, vol. 20, no. 1, 1960, p. 37.

[27] J. R. Landis and G. G. Koch, “The measurement of observer agreement for categorical data,” Biometrics, vol. 33, 1977, pp. 159–174.

[28] N. Smirnov, “Table for estimating the goodness of fit of empirical distributions,” The Annals of Mathematical Statistics, vol. 19, no. 2, 1948, pp. 279–281, ISSN: 00034851.

[29] H. Scheffé, The Analysis of Variance, ser. A Wiley publication in mathematical statistics. Wiley & Sons, 1999, ISBN: 9780471345053.

[30] O. Corporation, “Originlab,” 2012, URL: http://www.originlab.com/ [accessed: 2014-12-29].

[31] T. Wehrle, S. Kaiser, S. Schmidt, and K. R. Scherer, “Studying the dynamics of emotional expression using synthesized facial muscle movements,” Journal of Personality and Social Psychology, vol. 78, no. 1, 2000, pp. 105–119.

[32] Z. Ambadar, J. W. Schooler, and J. F. Cohn, “Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions,” Psychological Science, vol. 16, no. 5, 2005, pp. 403–410, ISSN: 09567976.

[33] E. Bould and N. Morris, “Role of motion signals in recognizing subtle facial expressions of emotion,” British Journal of Psychology, vol. 99, 2008, pp. 167–189.

[34] R. Adolphs, “Recognizing emotion from facial expressions: psychological and neurological mechanisms.” Behavioral and cognitive neuroscience reviews, vol. 1, no. 1, Mar. 2002, pp. 21–62.

[35] J. Leppänen and J. Hietanen, “Positive facial expressions are recognized faster than negative facial expressions, but why?” Psychological Research, vol. 69, 2004, pp. 22–29.

A User-Centered Approach for Social Recommendations

F. Colace, M. De Santo and L. Greco

Dipartimento di Ingegneria dell'Informazione,
Ingegneria Elettrica e Matematica Applicata
University of Salerno
Fisciano, Italy
Email: {fcolace,desanto,lgreco}@unisa.it

F. Amato, V. Moscato, F. Persia and A. Picariello

Dipartimento di Ingegneria Elettrica e
Tecnologie dell'Informazione
University of Naples
Naples, Italy
Email: {flora.amato,vmoscato,fabio.persia,picus}@unina.it

***Abstract*—Recommender Systems represent useful tools helping users to find "what they need" from a very large number of candidates and supporting people in making decisions in several contexts. In this paper, we propose a novel user-centered and social recommendation approach in which several aspects related to users, i.e., *preferences*, *opinions*, *behavior*, *feedbacks*, are considered and integrated together with *items' features* and *context information* within a general framework that can support different applications using proper customizations (e.g., recommendation of news, photos, movies, travels, etc.). Preliminary experiments on system accuracy show how our approach provides very promising and interesting results.**

***Keywords*–Recommender Systems; Sentiment Analysis; Context-Awareness.**

I. Introduction

In the era of *Big Data*, we are assisting to an explosive and amazing increase of digital information and, as a consequence, more and more huge data collections of different nature are widely available to a large population of users. In addition, the widespread diffusion of the most popular social networks, multimedia repositories, news archives, travel websites, e-commerce portals, and so on, has constrained users necessarily to deal with this ocean of information to find "what they need".

In the last decade, *Recommender Systems* have been introduced to facilitate the browsing of such collections leveraging several features to provide useful recommendations: user preferences and past behavior, preferences and past behavior of the user community, items' features and how they can match user preferences, user feedbacks, context information and how recommendations can change together with the context. To accomplish their goals, the last generation of recommender systems is usually composed by one or more of the following components [1]: (i) a *pre-filtering* module that selects for each user a set of objects that are good candidates to be recommended (such objects usually match user preferences and needs or satisfy context contraints); (ii) a *ranking* module that assigns to every user a rank related to each candidate object using recommendation techniques (that can exploit in several ways items' features and users' preferences, feedbacks and behaviors); (iii) *post-filtering* module that dynamically excludes, for each user, some items from the recommendations' list on the base of the collected feedbacks and context information.

In this paper, we propose a novel *user-centered* approach that provides *social* recommendations, capturing and exploiting several aspects related to users: *preferences* (usually coded in the shape of items' metadata), *opinions* (textual comments to which it is possible to associate a particular sentiment), *behavior* (in most cases, logs of past items' observations made by users), *feedbacks* (usually expressed in the form of ratings). All these features are considered and integrated together with *items' features* and *context information* within a general and unique recommendation framework that can support different applications using proper customizations (e.g., recommendation of news, photos, movies, travels, etc.), overcoming problems related to the availability and quality of user profiles and ratings.

As motivating example, we can consider a user who desires to have information about the coming soon movies. In this case, the list of suggested items should consider the following information: (i) user preferences in terms of movies' metadata (e.g., favorite genre, director, stars, etc.); (ii) item features (i.e., movies' metadata) and their similarity (e.g., a semantic relatedness based on a movie taxonomy); (iii) user behavior in terms of the sequence of items that in the past the community of users have observed and positively rated; (iv) user feedbacks in terms of the user community ratings; (v) user opinions in terms of the average sentiment that items have aroused on the user community; (vi) context information: in this case, in terms of item features that satisfy the search criteria (e.g., coming soon movies shown in theaters near the user) or that have a good similarity with respect to the item that the user has selected and is currently watching.

For instance, we can imagine a user who prefers the adventure and fantasy genres and who has among his/her favorite actors Ian McKellen and Hugh Jackman; the system can initially suggest as first items to watch the X-Men saga movies, together with other titles. After the pre-filtering stage, the candidate items matching user preferences are initially ranked on the base of the related social *popularity* (e.g., number of accesses through past users' paths, average rating and sentiment from users' reviews, etc.). Successively, if the user selects to read more information about one of X-Men movies and rates it positively, a list of filtered items is then provided with the most popular movies that have a certain similarity in terms of metadata with the X-Men movie.

 ISBN: 978-1-61208-382-7

Eventually, if the user chooses to limit the search to the coming soon movies (constraint on a metadata value) and selects his/her position as context information, all the best movies - in according to a social view - matching user preferences that are showing in the next days in theaters near the user will be finally proposed (e.g., “X-Men: Days of Future Past” and “Captain America: The Winter Soldier”).

The paper is organized as follows. Section 2 discusses the state of the art of other similar systems, while Section 3 describes the proposed strategy for recommendation. Section 4 illustrates a system customization in the domain of movie recommendation and reports some preliminary experimental results, and provides a comparison with other recommendation techniques. Finally, Section 5 gives some concluding remarks and discusses future work.

II. Related Works

Recommender Systems represent a meaningful response to the problem of information overload since the mid-1990s [2], when the early works on this topic have been proposed. The main aim is to predict user’s preferences and make meaningful suggestions about items that could be of interest [1].

In *content-based* approach, the system recommends an item to a user relying on the ratings made by the user himself for *similiar* items in the past. The similiarity between items is often computed by the use of information retrieval and filtering techniques [3]. However, a critical drawback of this approach is *overspecialization*, since the systems only recommend items similiar to those already rated by the user. In *collaborative filtering* [4], the recommendation is performed by filtering and evaluating items with respect to ratings from other users. Typically, users are asked to rate items and a similarity between their profiles is also computed to be used as a weight when making recommendations for highly rated items [5]. An important limitation of collaborative filtering systems is the *cold start problem*, that describes situations in which a recommender is unable to make meaningful recommendations due to an initial lack of ratings, thus degrading the filtering performance. Content-based filtering and collaborative filtering may be combined in the so called *hybrid* approach that helps to overcome limitations of each method [6]. A recommendation strategy eventually should be also able to provide users with the more relevant information depending on the *context* [7][8][9] (i.e., user preferences, user location, observed objects, weather and environmental conditions, etc.) as in *Context Aware Recommendation Systems*.

Performance of recommender systems is strictly related to the *availability* and *quality* of user profiles and ratings and an important improvement to overcome such problems lies in the possibility to embed *social* elements into a recommendation strategy [10]. In such a context, customer *opinion summarization* [11] and *sentiment analysis* [12] techniques represent effective augmentations to traditional recommendation strategies, for example by not recommending items that receive a lot of negative feedbacks [10]. Finally, a recent category of recommenders, named *Large Scale Recommender Systems* (LSRS) [13], calls for new capabilities of such applications to deal with very large amount of data with respect to scalability and efficiency issues. Distributed computing of recommendations and parallel *matrix factorization* are the most diffused approaches to cope with such a problem.

Our work exploits sentiment classification techniques based on the *Latent Dirichlet Allocation* (LDA) to refine and enrich a context-aware and hybrid recommendation strategy that some of the authors have proposed for recommendation in multimedia browsing systems [14][15][16]. We thus obtained a user-centered approach in which several aspect of a user (preferences, opinions, feedbacks, behaviors) are simultaneously considered together with item features and context information within a unique and general framework able to efficiently scale with the increase of data.

III. The Recommendation Strategy

The basic idea behind our proposal is that when a user is browsing a particular items’ collection, the recommender system: (i) determines a set of useful *candidate* items for the recommendation, on the base of user actual needs and preferences (*pre-filtering stage*); (ii) opportunely assigns to these items a rank, previously computed exploiting items’ intrinsic features and users’ past behaviors, and using as refinement, social information in the shape of users’ opinions and feedbacks (*ranking stage*); (iii) dynamically, when a user “selects” as interesting one or more of the candidate objects, determines the list of the most suitable items (*post-filtering stage*), also considering other context information expressed by users in the shape of constraints on items’ features.

A. Pre-filtering Stage using user preferences

In the *pre-filtering* stage, our aim is to select for a given user u_h a subset $O_h^c \subset O$ (O being the set of items) containing items that are good “candidates” to be recommended: such items usually have to match some (static) user preferences and (dynamic) actual needs. Each item subject to recommendation may be represented in different and heterogeneous feature spaces and the first step consists in clustering together “similar” items, where the similarity should consider all (or subsets of) the different spaces of features. To this purpose, we employ *high-order star-structured co-clustering* techniques - that some of the authors have adopted in a previous work [16] - to address the problem of heterogeneous data pre-filtering. The pre-filtering stage leverages the clustering results to select a set of items by using the user’s profile, which is modeled as sets of descriptors in the same spaces as the items’ descriptors.

B. Ranking Stage using user behavior and items similarity

We use a technique that some of the authors have proposed in previous works - combining low and high level features of items, past behavior of individual users and overall behavior of the whole user “community” [14][15] - to provide useful recommendations during the browsing of multimedia collections. Our basic idea is to assume that when an item o_i is chosen after an item o_j in the same user *browsing session* (and both the explored items have been positively rated or have captured attention of users for an adequate time), this event means that o_i “is voting” for o_j. Similarly, the fact that an item o_i is “very similar” in terms of some intrinsic features to o_j can also be interpreted as o_j “recommending” o_i (and viceversa). Thus, we are able to model a browsing system for the set of items O as a labeled graph (coding both browsing sessions of the different users and similarity between items’ pairs by means of a set of proper matrices), and to compute the *recommendation grade* $\rho(o_i)$ for an item $o_i \in O_h^c$ related to a given user u_h in a similar manner to *Google Page Rank* algortihm [15].

C. Refining items ranks using user sentiments and feedbacks

We used the sentiment extraction technique as an improvement of the approach presented by some of the authors in a previous work [17], where the LDA has been adopted for mining the sentiment inside documents. In our view, the knowledge within a set of documents can be represented in a compact fashion by the use of a complex structure - the *mixed Graph of Terms* (mGT) - that contains the most discriminative words and the probabilistic links between them. The mGT is built starting from a set of comments belonging to a well-defined knowledge domain and manually labeled according to the sentiment expressed within them. In this way, the mGT contains words (and their probabilistic relationships) which are representative of a certain sentiment for that knowledge domain.

For the ranking refinement, we introduce two probabilities P^+ and P^- which express the probability that a sentiment, extracted from the set of comments related to a given item, is "positive" or "negative" (the probabilities P^+ and P^- also take into account the overall rating and trustiness of users). Such probabilities are then combined with the overall rank of an item by a proper function that increases the recommendation grade value if the sentiment within item's comments is positive, in the opposite decreases it in the case of negative mood.

D. Post-Filtering Stage using context information

In this stage, we have introduced a *post-filtering* method for generating the final set of "real" candidates for recommendation. The set of candidates includes the items that have been accessed by at least one user within k steps from a selected object o_j and the items that are most similar to o_j according to the results of a *Nearest Neighbor Query* ($NNQ(o_j, O^c_h)$) functionality.

The ranked list of recommendations is then generated by ranking the candidates set for each object o_j selected as interesting by user u_h. Finally, for each user, all the items that do not respect possible context constraints are removed from the final list. In our model, *context constraints* are expressed in terms of assigned values to the elements of particular sub-classes of features that the recommended items have to satisfy.

IV. Preliminary Experimental Results

A. Using the system for recommending movies

We have opportunely customized our system in order to provide recommendation services for users that are interested in coming soon movies. The design choices are briefly reported in the following: (i) we consider as data source the *IMDB* web site, collecting about 10,000 items; (ii) as items' metadata, we consider for each movie information related on *title*, *genre*, *stars*, *description*, *year*, *director* and *list of theaters* (characterized by name and location) in which they are coming; (iii) for each movie, available users' preferences, comments and feedbacks have been captured, also exploiting correlated public information from Social Networks (i.e., Facebook); (iv) users' behaviors have been reconstructed considering the available log with time-stamped information of users that have positively rated or watching for a certain time some items in the same browsing session.

B. Accuracy Computation

We decided to perform for the movie recommendation problem an evaluation based on *accuracy* metrics [18]. We used the dataset provided by [19], which makes available data collected by the *MovieLens* [20] recommender system. Through its website, MovieLens collects the preferences expressed by a community of registered users on a huge set of movie titles. The adopted dataset contains (i) explicit ratings about 1682 movies made by 943 users, (ii) demographic information about users (age, gender, occupation, zip code), and (iii) a brief description of the movies (title, release year, genres). We then determined user preferences by considering the known preferences of similar users (e.g., same age, occupation, etc.) from social networks and extended items' features by considering IMDB metadata, as well as, we exploited IMDB users' comments.

The experiments have been conducted on a collection of about 1,000 movies, rated by a subset of 100 users: each of them had rated at least 150 movies and at most 300, assigning to each movie a score between 1 (*"Awful"*) and 5 (*"Must see"*). Additionally, using the *timestamp* information, we were able to reconstruct usage patterns for each user and consequently the browsing matrices.

We used the *Mean Absolute Error* (*MAE*) and the *Root Mean Square Error* (*RMSE*) as metrics in our experiments. In our case, *MAE* and *RMSE* are defined as: $MAE = \frac{1}{N}\sum_{u,i,j}|r^j_{ui} - \hat{r}^j_{ui}|$ and $RMSE = \sqrt{\frac{1}{N}\sum_{u,i,j}(r^j_{ui} - \hat{r}^j_{ui})^2}$; where r^j_{ui} is the actual rating that the user u has given to item i for the item j, $\hat{r}^j_{ui}$ is the system predicted rating, and N is the total number of test ratings.

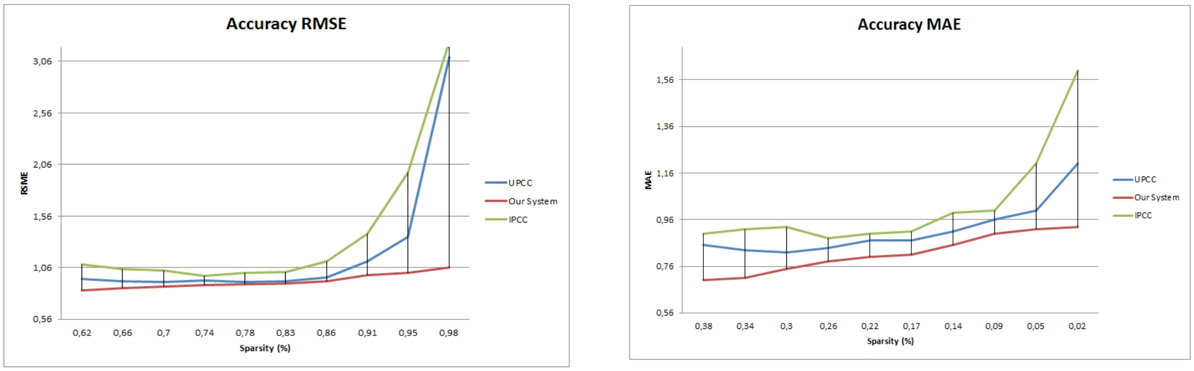
We compared the achieved accuracy of the predictions computed by our recommender system with the *UPCC* and *IPCC* [5] approaches (which reliable implementation can be obtained leveraging machine learning libraries provided by the *Apache Mahout* framework).

Fig. 1 shows the trend of RMSE and MAE for our system as well as for the UPCC and IPCC algorithms, as the sparsity of the rating matrix increases. Our approach outperforms UPPC and IPPC ones for each value of items' sparsity - and especially for higher values - showing how social information can improve recommendations. This is also due to the use of the items' *similarity matrix*, which provides useful information to the algorithm, in order to compute meaningful predictions even if a user's browsing session data is not available. Thus, our approach does not suffer from the *cold start problem.*

C. Considerations on efficiency

In order to evaluate the efficiency of our recommender system, we have measured execution times w.r.t. the execution times of other state-of-the-art methods. As the recommendation grades computation can be performed in off-line manner and the related updates are correlated to the insertion of a new item (or an update of its features) or to a new user, the running time is essentially dependent on the size of candidate items' set obtained in the pre-filtering stage. In general, we observed that the average computation times for all methods are comparable and it takes at most few seconds to obtain useful recommendations also for large sets of candidates.

 ISBN: 978-1-61208-382-7



(a) RMSE comparison

(b) MAE comparison

Figure 1. Comparison in terms of MAE and RMSE between our approach and UPCC and IPCC respectively

For scalability issues, we decided to use a distributed approach using *Apache Hadoop* framework to compute the items' clusters in the pre-filtering stage.

V. Conclusions and Future Work

We described a user-centered and social recommendation approach in which several aspects related to users - i.e., *preferences*, *opinions*, *behavior*, *feedbacks* - are considered and integrated together with *items' features* and *context information* within a general framework. We focused on a particular case study and implemented a recommender system based on our innovative approach, which is able to help users to choose coming soon movies having IMDB as main data source. Then, we investigated the effectiveness of the proposed approach in the considered scenarios, through the evaluation of the *accuracy*. Summing up, future efforts will be devoted to (i) extending the experimental evaluation to larger datasets, also considering the *stability* of recommendations, (ii) applying our approach to other kinds of data from heterogeneous collections and comparing it with other recent approaches of the literature.

References

[1] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds., Recommender Systems Handbook. Springer, 2011.

[2] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "Grouplens: An open architecture for collaborative filtering of netnews." ACM Press, 1994, pp. 175–186.

[3] M. Pazzani and D. Billsus, "Content-Based Recommendation Systems," 2007, pp. 325–341. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-72079-9_10

[4] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," IEEE Transactions on Knowledge and Data Engineering, vol. 17, 2005, pp. 734–749.

[5] X. Su and T. Khoshgoftaar, "A survey of collaborative filtering techniques," Advances in Artificial Intelligence, vol. 2009, 2009.

[6] A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock, "Methods and metrics for cold-start recommendations," in Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ser. SIGIR '02. New York, NY, USA: ACM, 2002, pp. 253–260. [Online]. Available: http://doi.acm.org/10.1145/564376.564421

[7] P. Dourish, "What we talk about when we talk about context," Personal and ubiquitous computing, vol. 8, no. 1, 2004, pp. 19–30.

[8] K. Kabassi, "Personalisation systems for cultural tourism," in Multimedia services in intelligent environments. Springer, 2013, pp. 101–111.

[9] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, "Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering," in Proceedings of the fourth ACM conference on Recommender systems. ACM, 2010, pp. 79–86.

[10] X. Zhou, Y. Xu, Y. Li, A. Josang, and C. Cox, "The state-of-the-art in personalized recommender systems for social networking," Artif. Intell. Rev., vol. 37, no. 2, Feb. 2012, pp. 119–132. [Online]. Available: http://dx.doi.org/10.1007/s10462-011-9222-1

[11] L. Zhuang, F. Jing, and X.-Y. Zhu, "Movie review mining and summarization," in Proceedings of the 15th ACM International Conference on Information and Knowledge Management, ser. CIKM '06. New York, NY, USA: ACM, 2006, pp. 43–50. [Online]. Available: http://doi.acm.org/10.1145/1183614.1183625

[12] X. Ding, B. Liu, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in Proceedings of the 2008 International Conference on Web Search and Data Mining, ser. WSDM '08. New York, NY, USA: ACM, 2008, pp. 231–240. [Online]. Available: http://doi.acm.org/10.1145/1341531.1341561

[13] H.-F. Yu, C.-J. Hsieh, S. Si, and I. S. Dhillon, "Parallel matrix factorization for recommender systems," Knowledge and Information Systems, 2013, pp. 1–27.

[14] M. Albanese, A. d'Acierno, V. Moscato, F. Persia, and A. Picariello, "A multimedia semantic recommender system for cultural heritage applications," in Semantic Computing (ICSC), 2011 Fifth IEEE International Conference on. IEEE, 2011, pp. 403–410.

[15] M. Albanese, A. dAcierno, V. Moscato, F. Persia, and A. Picariello, "A multimedia recommender system," ACM Transactions on Internet Technology (TOIT), vol. 13, no. 1, 2013, p. 3.

[16] I. Bartolini, V. Moscato, R. G. Pensa, A. Penta, A. Picariello, C. Sansone, and M. L. Sapino, "Recommending multimedia visiting paths in cultural heritage applications," Multimedia Tools and Applications, 2014, pp. 1–30.

[17] F. Colace, M. De Santo, and L. Greco, "A probabilistic approach to tweets' sentiment classification," in Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on, 2013, pp. 37–42.

[18] G. Adomavicius and J. Zhang, "Stability of recommendation algorithms," ACM Transactions on Information Systems (TOIS), vol. 30, no. 4, 2012, p. 23.

[19] "http://www.grouplens.org."

[20] "http://movielens.umn.edu."

Scalable Projection-Type Three-Dimensional Display by Using Compensation of Geometric Disotrtion

Youngmin Kim, Sunghee Hong, Sangkyun Kim, Hyunmin Kang, Jisoo Hong, Sangwon Lee, and Hoonjong Kang
Realistic Media Platform Research Center
Korea Electronics Technology Institute
Seoul, Korea
e-mail: rainmaker@keti.re.kr

Abstract— **We proposed an image compensation method of geometric distortion in multi-projection-type three-dimensional display. Projected images from optical modules could be distorted by toed-in configuration of array; we analyzed this relationship by using homography matrix. To verify our method, we designed zigzag configuration of multi-projectors and applied our proposed method. Experimental results will be provided to verify the proposed method.**

Keywords-Three-dimensional display; Geometrical optics

I. INTRODUCTION

Projection-type display has been widely used in three-dimensional display field because it can be spatially multiplexed and is separated from the screen. The most representative projection-type three-dimensional display approach is Holografika, which adopted multiple projectors to provide adequate number of views to the viewers. Holografika used a specially arranged array of optical modules and asymmetrically diffusive screen, where the large angle is Vertical Field of View (FOV) and the horizontal diffusion angle is equal to the angle between the optically neighboring modules and it corresponds to the angular resolution [1]-[4]. However, the projected images from a specially arranged array of optical modules can be distorted as the number of optical modules increase because the optical modules are positioned as toed-in configuration. Each point of the asymmetrically diffusive screen transmits the distorted images into different directions, so three-dimensional image can be deteriorated [5][6].

In this paper, we propose a compensation method of geometrical distortion in projection-type three-dimensional display by using multi-aperture optics. In Section II, we will analyze geometrical distortion due to perspective effects. In Section III, the compensation method will be discussed and practical approach of the compensation method will be described. In Section IV, we will provide simulated results from the compensation method using homography matrix and experimental results.

II. SCALABLE DIRECTIONAL-VIEW DISPLAY BY USING MULTIPLE PROJECTORS

Three-dimensional display technology has been developing since the unprecedented success of the three-dimensional movie 'AVATAR'. Among the three-dimensional display technologies, holography is the only way to express three-dimensional whole information of objects because the wave-front of object can be reconstructed. However, until now, it is practically hard to develop holography because of massive three-dimensional information contents manipulation and absence of high definition display device. As an alternative, super multi-view display was expounded by enthusiastic researchers. The advantage of the super multi-view display can provide motion parallax as well as binocular disparity because the interval among the views is small enough to provide smooth motion parallax.

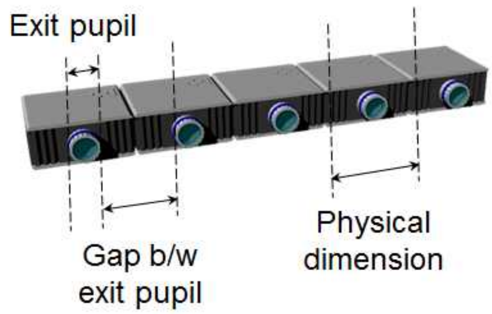


Figure 1. Relationship among the exit pupil, physical dimension, and gap between exit pupils in the array of conventional projection optical modules

Recently, we designed scalable directional-view display technique which is theoretically similar to super multi-view display by using compact and convenient subminiature projectors [5]. The reason why the super multi-view display is easy to be applied practically is effective size of physical projection optical modules. The exit pupil from commercially available projectors is too small compared with physical overall size of projector as shown in Figure 1. When an array of small exit pupil of projection lens is adopted in the super multi-view display, the rearranged three-dimensional rays from asymmetrically diffusive screen present an appearance of a black striped pattern like discontinuous image [6].

To solve this issue, it seems necessary to consider a use of multi-dimensional alignments of projection optics. In the previous paper, we designed zigzag configuration of projection optics by using disassembled commercial projectors for the three-dimensional display as shown in Figure 2. Since the horizontal interval of projection lens of the proposed method is same as the exit pupil of the projection lens, this configuration allows us to provide continuous directional-view images. Furthermore, to enlarge FOV of this configuration, a curved array of projection optics, so-called toed-in configuration, was installed in the proposed method.

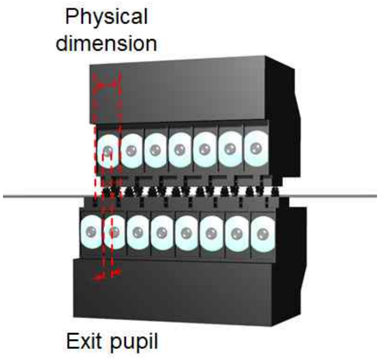


Figure 2. Schematic of proposed method by using an array of disassembled commercial projectors.

III. GEOMETRICAL DISTORTION IN TOED-IN CONFIGURATION

In the toed-in configuration of projection-type three-dimensional display, each projected image can be distorted due to their relative position from the screen. Therefore, it is important to compensate such distortions. In this section, the key issues regarding multiple projectors-asymmetrical diffusive screen calibration will be discussed. As shown in Figure 3, the fixed screen and multi-directional optical modules are adopted in a typical scheme of multiple projection-type three-dimensional display. We assumed that 1) the asymmetrically diffusive screen is flat although this method can be extended to non-planar screens, 2) the projection angle of the images emitted from each optical module is same, and 3) the projectors and the screen can be modeled by perspective transforms.

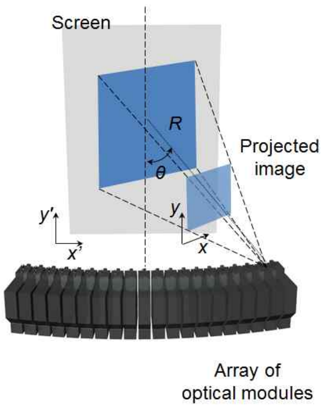


Figure 3. Schematic of the proposed method for the distortion analysis.

Consider a point (*x, y*) in the projector image plane. This point will be projected to unknown point (*x', y'*) on the screen. Primitive goal in this section is to find out a relationship between two corresponding points. We are able to exploit the fact that all of unknown point (*x', y'*) on the screen, and this can be established by a 3×3 homography matrix between the corresponding points of two different planes. Therefore, it can be expressed by a single projective transform [7],

$$H = \begin{bmatrix} R & 0 & 0 \\ 0 & R\cos\theta & 0 \\ \sin\theta & 0 & R\cos\theta \end{bmatrix} \tag{1}$$

with eight degrees of freedom; R denotes the distance between the center of screen and each projectors, θ indicates the angle between normal vector of the screen plane and projector image plane. The simple example (R = 500 mm, θ = 30º, and the pixel size of the screen plane is 0.2 mm) of above case can be shown in Figure 4.

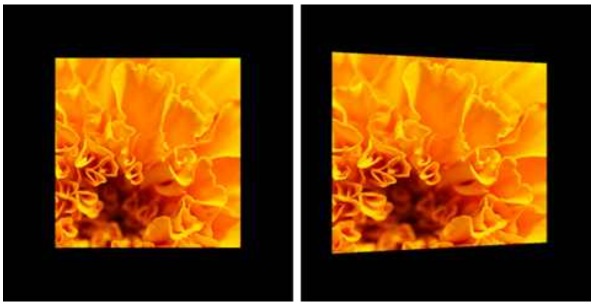


Figure 4. An original image of the projection optical module (left) and distorted image due to the toed-in geometry (right).

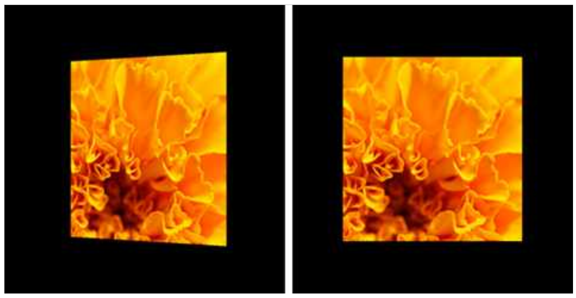


Figure 5. A distorted image from far right position (left) and compensated image by using homography matrix (right).

Since the homography matrix H represents the relationship of geometrical distortion between two coordinates (the projector image plane and the screen plane), the distortion can be compensated by multiplying coordination of original image by inverse matrix of homography. The relationship between pixels of original image and compensated image can be given by

$$(x_p, y_p) = (\frac{R\cos\theta}{R - x_0 \sin\theta} x_0, \frac{R}{R - x_0 \sin\theta} y_0) \quad (2)$$

where (x, y) and (x', y') denote the pixel position of the original image and compensated image, respectively. Using (2), the simulation was performed as shown in Figure 5. We can confirm that the distortion compensation results on the screen are same as original image when the distortion compensation image was projected on the screen.

IV. EXPERIMETNAL RESULTS AND DISCUSSION

To verify our method based on compensation of geometrically distorted images in multi-aperture three-dimensional display, we designed a curved array of multiple projection optical modules with 60º of field of view, which was called as toed-in configuration as well. As a projection optical module, commercially available pico-projectors with the resolution of 1280(H) × 720(V) (native resolution: 640 (H) × 480(V)) were used. This projector does not satisfy the requirement which was mentioned in Section 2, so we disassembled these projectors to provide scalable directional rays. The angular resolution of the toed-in configuration is 0.75º and the exit pupil of the projection optical module was 3.5 mm, as shown in Figure 6. Each projection optical module was connected with single-board computer, which was especially optimized in graphic performance and cost effectiveness.

Figure 6. Experimental setup for verifying our compensation method (up) and enlarged image of toed-in configuration of disassembled projection optical modules (down).

To experimentally verify that the proposed method can successfully compensate directional-view images, we coordinated the parameters of the cameras in the virtual space by using Unity 3D. The field of view of the virtual cameras was set to be same as that of projection optical modules. The captured directional-view images in the system were rearranged for pseudoscopic-orthoscopic conversion. We applied the homography matrix into the rearranged directional-view images for compensation as shown in Figure 7. However, the images still need compensation practically since all projection optical modules are not under the same conditions. So, we compensated this slight difference by using an image capturing device. We set the reference image from center of the optical module, and the projected images from other optical modules were captured. The captured images were compared with reference image from center optical module, and finally we could acquire compensated image in front of the asymmetrically diffusive screen. These findings, therefore, verify the compensation method of the proposed method. Based on the compensated results, we can apply this method for the toed-in configuration. Unlike Holografika method [1][2], this toed-in configuration provides improved FOV because the angle made by toed-in configuration could increase total FOV of the system. Therefore, we could establish distortion-free scalable three-dimensional display by using multi-aperture optics with enhanced FOV.

Figure 7. Experimental results before geometrical compensation (up) and after correction (down).

V. CONCLUSION

In this paper, we proposed compensation method of geometrical distortion in three-dimensional display by using multi-aperture optics. To verify our compensation method, we designed a toed-in configuration of projection optical modules and the contents was acquired by means of computer generated method. Captured directional-view images were compensated by using proposed homography matrix and a detailed compensation was performed by comparing with reference images from center of projection optical module. Further research directed toward real object contents acquisition will be required.

ACKNOWLEDGEMENT

This research was supported by the IT R&D program of MSIP/Giga Korea. [GK13D0200, Development of Super Multi-View (SMV) Display Providing Real-Time Interaction].

REFERENCES

[1] T. Balogh et al., “The holovisio system new opportunity offered by 3D displays”, In I. Horvth and Z. Rusk, editors, Proceedings of the TMCE 2008, pages 1-11. IEEE.

[2] T. Balogh, “Method and apparatus for displaying 3D images,” U.S. patent 6,999,071 (14 February 2006).

[3] Y. Kajiki, H. Yoshikawa, and T. Honda, “Hologram-like video images by 45-view stereoscopic display,” In Proceedings of SPIE **3012**, 1997, pages 154-166.

[4] T. Honda, D. Nagai, and M. Shimomatsu, “Development of 3-D display system by a fan-like array of projection optics,” In Proceedings of SPIE **4660**, 2002, pages 191-199.

[5] Y. Kim et al., “Scalable directional-view display and contents acquisition approach”, In Imaging and Applied Optics 2014, OSA Technical Digest (online), paper Jtu4A.5.

[6] Y. Kim, J. Hahn, and B. Lee, “Hologram-like display without crosstalk and discontinuity problem by using a holographic diffuser”, In Proceedings of International Conference on Fiber Optics and Photonics, 2008, pages 421-424.

[7] R. Sukthankar, R. G. Stockton, and M. D. Mullin, “Smarter presentations: Exploiting homography in camera-projector systems”, In Proceedings of International Conference on Computer Vis. 2001, pages 247-253.

Distributed Collaborative Construction in Mixed Reality

Christian Blank, Malte Eckhoff, Iwer Petersen, Raimund Wege and Birgit Wendholt
UAS Hamburg
Hamburg, Germany
Email: {first name}.{family name}@haw-hamburg.de

***Abstract*—Distributed collaboration, portable mobile applications, natural user interfaces and comprehensive systems have been identified as future research directions in recent reviews about mixed reality in construction. On the other hand, current research in the mixed reality field addresses movement and anthropometric realism as critical success factors for an immersive virtual environment. Advances in object tracking, online (human) 3D reconstruction and gestural interfaces accompanied by wearable mobile displays provide us with the technological base to contribute to the challenges in both areas. In this paper, we propose a comprehensive immersive environment for a distributed collaborative construction process in a mixed reality setup. Participants on remote sites, solely equipped with smart see-through glasses, are cooperating in the construction of a virtual 3D model combining real (tangibles) and virtual objects. We consider our solution to give most suitable support for a distributed collaborative construction task by increasing the immersion of the environment, i.e.: (1) creating the impression of real collaboration by mirroring the behavior of participants in a common virtual scene; (2) providing more natural interaction through freehand gestures; (3) increasing the physical experience of the user through wearable 3D displays and construction with tangibles.**

Keywords*–*Mixed Reality; Computer Supported Collaborative Work; Natural User Interaction.

I. Introduction

As has been shown by Rankohi et al. [1] and Chi et al. [2] mixed reality (MR) has been widely adopted in the construction field over the last decades. Hence, the same authors identify distributed collaboration, portable mobile applications and natural user interfaces (NUI) as future research topics. From the MR point of view, Dionisio et al. [3] consider the degree of realism in virtual environments as a relevant subject for future investigations. The closer gestural interfaces are to physical interactions - referred to as movement realism - and the higher the lifelikeness of virtual characters and humans - referred to as anthropometric realism - the higher the acceptance of the environment. Recent advances in 3D object tracking, 3D reconstruction, natural user interfaces and mobile MR devices, provide the means to bridge the gap between real and virtual collaboration tasks.

Given the latest technologies and research topics, we propose an environment for distributed collaborative construction where virtual and real processes converge because of (1) realistic images of participants in a common scene, (2) natural gestural interaction and (3) better physical experience through wearable displays and tangible interaction.

To be more specific, a scenario will illustrate the functionality of the environment. Participants on remote sites, solely equipped with optical see-through glasses, are cooperating in the construction of a virtual 3D model combining real (tangibles) and virtual objects. In the bottom cut-out of Figure 1 the virtual object is represented as transparent block (a partially completed marble track). The real object, a cube tagged with markers, can be attached to the virtual marble track and then gets replicated into a virtual piece of the track. Each participant will see the constructed model as overlay to the construction scene from an individual perspective. Manipulations of one participant will be transferred directly to all others. Since conflicting actions cannot be completely avoided in a distributed environment, we assume that all participants behave cooperatively. To provide best possible support for cooperation, each participant will see 3D reconstructed representatives of the others in the virtual scene, depicted as gray shaded characters in Figure 1. All manipulations on virtual objects and the replication of real into virtual objects are performed in a gestural manner. Equipped with mobile see-through glasses each participant will be able to inspect the virtual model from different perspectives by moving around in the scene.

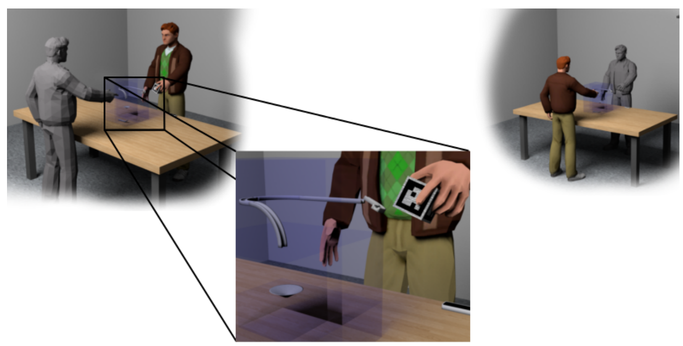


Figure 1. Overview of the proposed system. Two participants on two remote sites work collaboratively on a 3D model of a marble track.

In the remainder of this paper, section II briefly discusses the related work. Section III sketches the overall system design and introduces the system components. Section IV proposes a solution to ensure model consistency in a distributed construction process. Section V presents the components which are responsible for scene visualization. Section VI deals with the interaction techniques. Section VII sketches our solution for 3D online reconstruction of humans. Finally, section VIII will outline the major open work packages and future directions.

II. Related Work

A number of 3D MR construction environments have been developed in recent years. Salim [4] uses tangible building blocks and physical gestures to construct virtual urban landscapes on a 3D simulation table. Though already employing marker-less object tracking and object reconstruction, the solution is designed as a stationary, single user application. As opposed to [4], we employ marker-based object tracking with visible markers like in [5] or [6] for efficient identification and more reliable 3D pose detection. Though different proposals exist to estimate the object pose from non-coplanar feature

 ISBN: 978-1-61208-382-7

points like in [7] and [8], we decided to implement an algorithm using coplanar feature points as described in [9] allowing for object pose calculation with only one recognized marker.

MirageTable [10] is an environment to combine the virtual and real world in a consistent virtual scene. It supports the construction task as combination of real and virtual objects and collaboration of users on a common model supplying 3D representations of the participants in the scene. It also allows for physically-realistic freehand gestures to manipulate virtual objects. Moreover, marker-less object tracking, object reconstruction and replicating real objects into virtual ones are contained. Though very close to our proposal, they follow a stationary approach with a stereoscopic projection. Here movement, fast on-line reconstruction of humans and dynamically changing perspectives are not considered. Since we propose a mobile setup, where participants are allowed to move around in the scene, there is a need for complete 3D models of all remote participants. Due to constant changes in perspectives, the reconstruction task needs to be performed in near real-time. Though Tong et al. [11] have shown the feasibility of using multiple consumer-grade depth cameras for reconstruction, their approach is too slow to create real-time dynamic meshes. As opposed to Alexiadis et al. [12], whose real-time reconstructed 3D models contain a high number of vertices, which is unsuitable for later streaming, in our solution, the data volume of the reconstruction process can be adapted in an early stage. This allows us to balance performance with mesh quality. As far as gestural interaction is concerned our work joins physical and interpreted gestures to achieve consistent device free interaction. For physical gestures we adapt the work of Song et al. [13] and Hilliges et al. [14]. For interpreted gestures we extend the template based approach of Kristensson et al. [15] for 2D gestures to 3D spatial interaction.

MixFab [16] is a MR environment for gesture-based construction of 3D objects in a stationary setting with a see-through display. Real objects are scanned by means of a depth sensor and can be combined with virtual ones. Manipulations range from joining real and virtual objects to deforming virtual by means of real objects. Having focus on mixed construction manipulations and gestures, MixFab does not support collaborative tasks, nor does it provide a mobile solution.

Mockup Builder [17] is a semi-immersive environment for freehand construction of 3D virtual models on a stereoscopic multi-touch table. The focus is on appropriate, convenient hand-gestures and thus, an excellent foundation for further development of our gestural interface.

III. System Design

The collaborative distributed environment consists of a couple of client instances. A client instance defines the environment which will run at all remote sites. Each instance consists of multiple components, which are communicating via a network middleware. Figure 2 shows the component dependencies, their attached sensor devices and the overall data flow. The reconstruction component is responsible for online 3D reconstruction of participants, whose results, the user meshes, are distributed to all other remote client instances for displaying purposes. The components tangible tracking and gesture recognition support the input side of the user interface. With tangible tracking, real objects can be incorporated in the construction process. Both, the construction logic and the scene visualization component are continuously informed about the actual positions of the real objects in the scene. The construction logic uses the position information to decide whether a real object's position is suitable for joining with the virtual construction model. The scene visualization tracks the real object's position by means of a virtual replicate. The gesture recognition component identifies gestures and informs about physical interactions with virtual objects. The construction logic component maps gestures onto model actions like joining or separating objects with/of the model. Physical interactions with virtual objects are appropriately reflected in the construction scene and model. The construction logic component ensures model consistency with respect to domain constraints and among all concurrent manipulations of client instances. The scene visualization component creates a consistent view of all output related data yielded by other components and renders the display data with respect to the view-ports, which are reported by the mobile display devices.

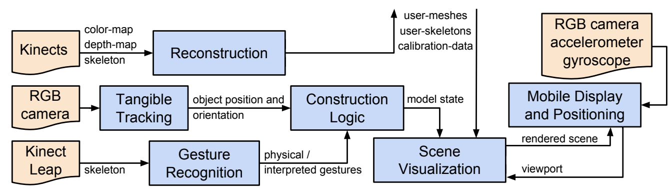


Figure 2. Client architecture

The backbone of component communication is the network middleware, which serves as an abstraction layer for the technical network. To achieve better scalability and extensibility, we have chosen a design of loosely coupled components and an event-driven and message-based communication style. To achieve location transparency, a service registry decouples network addresses from component services. Communication between client instances takes place in two ways: event-based, when actions of one client affect the underlying construction model and continuously, when user meshes are exchanged among instances.

IV. Construction Logic

The construction logic uses data from several components and controls the model construction based on domain-specific constraints. This module has to ensure model consistency and executes and resolves conflicting user actions in a concurrent distributed environment.

Concept: To ensure model consistency on a logical level, the very basic idea is to represent each entity as a building block with joints. Constraints are expressed in terms of joints, which specify criteria for valid connections. Only entities whose joints have matching criteria can connect to each other.

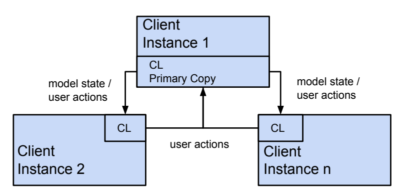


Figure 3. Distribution architecture. The construction logic (CL) component of each instance reports user actions to the primary copy, which synchronizes the model state and user actions among all client instances.

To enforce model consistency in a concurrent distributed environment in the first instance, user actions can only be

Copyright (c) IARIA, 2015. ISBN: 978-1-61208-382-7

performed on a primary copy (see Figure 3). Actions need to be executed in a strictly serialized and deterministic way to ensure consistent views for all users.

Current State: In an experimental setup, where the user can create a virtual marble track, the constraint-based construction approach has been validated successfully for different entity types and one unique constraint type. Results are shown in Figure 4 parts (b) and (c).

Future Work: The construction logic has to be further developed in several respects: (1) The logic itself has to be extended in order to cope with realistic domain models. (2) A solution for synchronizing replicated client-side model copies needs to be developed in order to enable consistent distributed construction.

V. Scene Visualization and Mobile Display and Positioning

Concept: The scene visualization component has to merge data from several components like (1) object positions from the tracking component, (2) model state from the construction logic and (3) user meshes from the reconstruction component. It will render the data into a consistent 3D scene and distribute the scene to all mobile displays of one client instance suitable for their individual view-ports.

Current State: Unity 3D is used as the engine for scene rendering and libGDX to display the rendered scene on Android-based mobile devices. All mobile devices permanently calculate their individual view-port and send it to the visualization component. The rotation matrix of a mobile module is calculated from an integrated accelerometer and gyroscope. A prototypical implementation for an Android tablet has been completed. The implementation for optical see-through glasses is currently under development. A more complicated task is to determine the head position. For now, a marker in the real scene represents the origin of the world space and gets tracked via a RGB camera of the mobile module.

Future Work: Using a marker to determine the head position, requires that users always look in the direction of the marker. A better solution will be global head-tracking for view-port-position calculation. This service will be integrated into the reconstruction component. Also, displaying user meshes in the scene is still an open task.

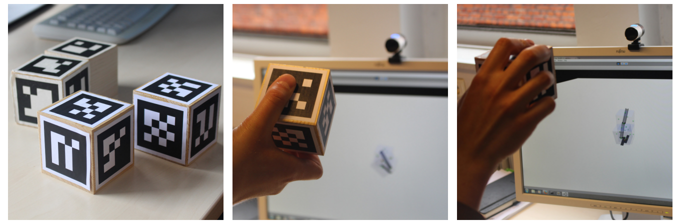
VI. Natural User Interface

Following Wachs et al. [18], who emphasize the role of natural user interfaces for intuitive and more natural interaction, we decided to exclusively use tangible and gestural interaction in our system. This section introduces the two subsystems responsible for tangible and gestural support.

A. Tangible Tracking

Unlike the common understanding of tangibles to serve as input controller we consider tangibles as real objects that become part of the construction model. Thus, tangible tracking for us means object tracking. For rapid prototyping purposes a marker-based solution has been implemented.

Current State: The marker-based approach uses cubes as representatives of domain entities, i.e., components of a marble track. Cubes carry unique, rotation-invariant markers on each side. A webcam is used as input device. Markers are continuously tracked in the image frames and marker positions are determined. Because of uniqueness and rotation invariance of the markers the basic orientation and the identity of objects can be determined. A coplanar POSIT algorithm [9] is used to estimate position and rotation of the marker relative to the camera. Finally, the actual position and rotation of the cube in the world system are calculated.



(a) Cubes represent marble track pieces. A cube carries unique markers, one marker for each side of a piece.

(b) Tracking: Moving cubes into the scene will create corresponding virtual marble track pieces. Virtual pieces follow the movement of the cubes.

(c) Constructing: Virtual pieces may join on valid connections which can be established through a gesture.

Figure 4. Constructing with tangibles.

The solution is capable of recognizing up to 5 unique objects in an area of 0.8 m in front of the camera with an update rate of 50 events per second. Occlusion is not handled.

B. Gesture Recognition

To support freehand 3D interaction with virtual objects a gesture recognition subsystem for interpreted as well as physical gestures is currently under development. Physical gestures are the virtual counterparts for interacting with objects in the real world, i.e., human movement has direct, realistic impact on virtual objects. Interpreted gestures are abstractions for movement patterns. These might be pointing gestures for menu item selection or object-related gestures like scaling virtual objects. The gesture recognition subsystem should provide a suitable abstraction layer to handle different kinds of input devices and multiple sensors. It should also be able to handle multiple users at the same time.

Concept: The gesture recognition subsystem consists of two major components: (1) Trame, a component for device abstraction in order to handle multiple sensor input and (2) the core recognition component for gesture detection, see Figure 5. Trame transforms sensor data into a common skeleton model. A controller of the core component dispatches the input to user related pipelines. These are processed in parallel, so that gestures of different users can be processed with interactive response times. Skeleton preprocessing will be used to smooth jitter, extract arm and hand positions, etc.

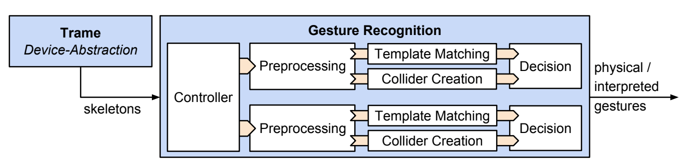


Figure 5. Overview of the gesture based interaction module with device abstraction and gesture recognition.

Template matching [15] enhanced with the observation of the third dimension and an extended set of input joints is responsible for detecting interpreted gestures. In parallel, a collider object, representing hand and arm movement, will be calculated in order to cope with physical gestures. In the decision step, interpreted gestures trigger corresponding events, which get distributed in the environment. In any case, a collider object will be provided for further processing.

(a) point cloud (b) wireframe (c) vertex colored triangle mesh

Figure 6. Reconstruction result using a single camera and organized fast mesh triangulation

Current State: Currently, Trame, the abstraction layer for sensor input, is implemented and supports Leap Motion and the Microsoft Kinect sensor.

Future Work: The implementation of the gesture recognition module for both kinds of gestures in a concurrent multi-user environment is one of the next goals to achieve. Afterwards usability studies need to be performed in order to verify the key assumption, that providing physical and interpreted gestures will give users an interface for interacting in a fast and natural way.

VII. User Reconstruction

When multiple users are operating in the MR environment, conflicting actions are imminent. We suppose that realistic virtual representatives of participants in a common scene will support cooperation since intentions of others might better be perceived. In technical terms, this means generating a closed textured polygon mesh for each user and visualizing it in the common 3D scene independently from the view-port.

Concept: As apposed to Alexiadis et al., who triangulate first and then combine multiple camera data - resulting in several mesh cleaning steps - this work proposes to first combine the data from multiple cameras and then to apply the mesh triangulation algorithm. Depending on the selected mesh triangulation algorithm some preparation steps may be necessary. The reconstruction pipeline therefore consists of a point processing step and a mesh triangulation step. The first step combines the data, and prepares the point cloud for the successive mesh triangulation. For example, a KD-tree of the point cloud is needed, when a moving-least-squares algorithm is used for triangulation. Multiple consumer-grade depth cameras are to be placed around a participant and calibrated to be able to transform the point cloud data into a common coordinate system.

Current State: Using a single depth camera a polygon mesh can be reconstructed at about 30 fps. The preparation step performs a background separation based on a thresholding approach only. The mesh is then triangulated with the organized fast mesh algorithm, which exploits point neighborhood relations known from the depth image. This naive meshing algorithm is not guaranteed to produce a closed mesh, as can be seen in Figure 6. While for a single camera this is a very efficient approach, it is not applicable to a point cloud assembled from data of several cameras.

Future Work: Future work includes the implementation of multi-camera management and a calibration method. Also needed is an evaluation of different processing pipelines for different meshing algorithms and related filtering steps in terms of speed and reconstruction quality. For example, moving-least-squares variants, greedy projection triangulation and marching cubes reconstruction are the next triangulation methods that will be evaluated. A solution for efficient mesh streaming is currently under investigation.

VIII. Conclusion and Future Work

In the preceding sections, we have outlined the architecture and components for a distributed collaborative construction environment. For each component, the current working state has been presented. Since we are at an early project stage, a number of open tasks have to be completed in the near future. In parallel, we are discussing realistic scenarios with manufacturing engineers and designers in order to verify our initial hypothesis, that virtual human representation and more natural interaction in conjunction with increased physical experience contribute to a better support for distributed construction. To end up with sound statements about our contribution to human computer interaction, we are planning a couple of comparative user studies with domain experts where we will investigate, whether (1) mixed construction has better acceptance and leads to better performance than pure virtual construction, (2) realistic 3D representations of participants better supports collaborative construction than employing self-animated avatars, (3) pure gestural and tangible interaction outranges more traditional interaction styles.

References

[1] S. Rankohi and L. Waugh, “Review and analysis of augmented reality literature for construction industry,” Visualization in Engineering, vol. 1, no. 1, 2013, p. 9.

[2] H.-L. Chi, S.-C. Kang, and X. Wang, “Research trends and opportunities of augmented reality applications in architecture, engineering, and construction,” Automation in Construction, vol. 33, no. 0, 2013, pp. 116 – 122, augmented Reality in Architecture, Engineering, and Construction.

[3] J. D. N. Dionisio, W. G. Burns, and R. Gilbert, “3d virtual worlds and the metaverse: Current status and future possibilities,” ACM Comput. Surv., vol. 45, no. 3, Jul. 2013, pp. 34:1–34:38.

[4] F. Salim, “Tangible 3d urban simulation table,” in Proceedings of the Symposium on Simulation for Architecture & Urban Design, ser. SimAUD ’14. San Diego, CA, USA: Society for Computer Simulation International, 2014, pp. 23:1–23:4.

[5] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, “Automatic generation and detection of highly reliable fiducial markers under occlusion,” Pattern Recognition, vol. 47, no. 6, 2014, pp. 2280–2292.

[6] M. Fiala, “Designing highly reliable fiducial markers,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 32, no. 7, 2010, pp. 1317–1324.

[7] L.-J. Qin and F. Zhu, “A new method for pose estimation from line correspondences,” Acta Automatica Sinica, vol. 34, no. 2, 2008, pp. 130–134.

[8] D. F. DeMenthon and L. S. Davis, “Model-based object pose in 25 lines of code,” International Journal of Computer Vision, vol. 15, 1995, pp. 123–141.

[9] D. Oberkampf, D. F. DeMenthon, and L. S. Davis, “Iterative pose estimation using coplanar feature points,” Computer Vision and Image Understanding, vol. 63, no. 3, 1996, pp. 495–511.

[10] H. Benko, R. Jota, and A. Wilson, “Miragetable: Freehand interaction on a projected augmented reality tabletop,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ser. CHI ’12. New York, NY, USA: ACM, 2012, pp. 199–208.

[11] J. Tong, J. Zhou, L. Liu, Z. Pan, and H. Yan, “Scanning 3d full human bodies using kinects,” IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 4, 2012, p. 643650.

[12] D. S. Alexiadis, D. Zarpalas, and P. Daras, “Real-time, full 3-d reconstruction of moving foreground objects from multiple consumer depth cameras,” IEEE Transactions on Multimedia, vol. 15, no. 2, 2013, pp. 339–358.

[13] P. Song, H. Yu, and S. Winkler, “Vision-based 3d finger interactions for mixed reality games with physics simulation,” in Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry. ACM, 2008, p. 7.

[14] O. Hilliges, D. Kim, S. Izadi, M. Weiss, and A. Wilson, “Holodesk: direct 3d interactions with a situated see-through display,” in Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems. ACM, 2012, pp. 2421–2430.

[15] P. O. Kristensson, T. Nicholson, and A. Quigley, “Continuous recognition of one-handed and two-handed gestures using 3d full-body motion tracking sensors,” in Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces, ser. IUI ’12. New York, NY, USA: ACM, 2012, pp. 89–92.

[16] C. Weichel, M. Lau, D. Kim, N. Villar, and H. W. Gellersen, “Mixfab: A mixed-reality environment for personal fabrication,” in Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems, ser. CHI ’14. New York, NY, USA: ACM, 2014, pp. 3855–3864.

[17] B. R. De AraúJo, G. Casiez, J. A. Jorge, and M. Hachet, “Special section on touching the 3rd dimension: Mockup builder: 3d modeling on and above the surface,” Comput. Graph., vol. 37, no. 3, May 2013, pp. 165–178.

[18] J. P. Wachs, M. Kölsch, H. Stern, and Y. Edan, “Vision-based hand-gesture applications,” Commun. ACM, vol. 54, no. 2, Feb. 2011, pp. 60–71. [Online]. Available: http://doi.acm.org/10.1145/1897816.1897838

Perceptional Approach to Design of Industrial Plant Monitoring Systems

Mehmet Göktürk
Department of Computer Engineering
Gebze Institute of Technology
gokturk@gyte.edu.tr

Mustafa Bakır, Burak Aydoğan, Mehmet Aydın
Process Improvement and Software Division
TÜPRAŞ İzmit Petroleum Refinery
{mustafa.bakir,burak.aydogan,mehmet.aydin@tupras.com.tr}

***Abstract*— In this paper, we have investigated the perceptional attitudes of a massive scale industrial plant staff towards process monitoring systems and focused on human factors that are influential in design of plant monitoring systems. The study is a part of a plant-wide monitoring system which is under development, aimed to help staff to monitor processes and plant performance in detail. The paper first gives focused introduction on large scale enterprise and plant monitoring and management information systems, then discusses human machine interaction relevance of these systems to the staff performance and perception. Examples from the literature and previous work are presented. Primary human factors in proactive monitoring and highly automated systems are briefly discussed. A design survey study follows the introduction and relevant literature sections. The survey is designed to collect perceptional status of the staff against such systems, including the perception of their current performance. Results are analyzed and discussed in order to enhance system design decisions in such plant monitoring systems. We believe that such perceptual studies, performed before attempting to implement large scale monitoring systems that are highly interactive to the existing staff, should be considered as essential part of the design process. The results of this study is being used as inputs in implementation of a recent petroleum plant monitoring system.**

Keywords-Plant Monitoring Systems; Management Information System; Perception, Human Performance.

I. INTRODUCTION

Large scale industrial plants comprise challenging environments for information technology specialists and software developers. Small scale plants do usually include few number of independent technical systems and simple information architecture. On the contrary, large scale plants are made up of various subsystems each probably built by different vendors at different times, using various technologies and diverse information architectures. Large scale plants are therefore significantly heterogeneous environments. Since assimilation of plant-wide information and trends completely by each user/staff is relatively a hard task in such heterogeneous environments, any lack of awareness about certain parameters could result in serious consequences and losses.

Modern process control systems are highly connected to plant-wide information systems and able to push significantly detailed data to upper information layers for control and monitoring activities. Such plants have highly automated processes with independent computerized process controllers that are responsible for running individual processes in optimum performance. High degree of automation and heterogeneous structure usually impairs visibility of parametric information about processes.

Often, information exchange between these automation and users who are responsible from running the plant are inadequate. This can eventually cause inefficiencies in plant performance and even result in life threatening catastrophes. The famous Three Mile Island nuclear accident (1979), is one example where poor user-system information exchange caused catastrophic plant failure [1][2].

Munro and Tilyard indicate the problem of user interaction in industrial environments as follows: “The industry’s strength has been in finding technical or hardware solutions while its weakness has been at the people end of the business in maximizing and consolidating the gains from the technologies” [3].

Monitoring a complex system is generally a hard task. A human operator needs to be aware of what is going on in a plant to a certain degree where his job is performed without any consequences or loss of production. However, required degree of awareness and level of detail that need to be provided remain unknown for most cases. The term “situational awareness” is used to describe this condition in literature. There have been numerous studies regarding situational awareness of the user in aviation and military areas due to their relative importance due to life critical nature of the operations [4][5]. As plant technologies advance, importance of plant monitoring also becoming significant both due to increased life critical nature and economical impacts in case of possible failures. When plant staff is unable to interact with and control correctly the required parameters, it is usually attributed to situational awareness problem. This is due to the fact that a normal behaving staff would think to act positively and in parallel to task descriptions in all situations.

Researcher David Hopkins summarizes major reasons of most situational awareness problems into four categories in terms of their consequences on user actions [6]:

 ISBN: 978-1-61208-382-7

1. The user has a full appreciation and understanding of the situation but an inability to take action. This is rare but may occur, such as, in extreme fatigue where the human can appreciate a situation but is too tired to do anything about it.

2. The user may have an adequate perception of all the relevant stimuli but a failure to appreciate their meaning or import. For example, the user may see an indicator, but forget what it means.

3. A user may fail to perceive a particular stimulus. He/she may not notice an icon for example, but see other items perfectly well.

4. A user may not perceive any of the surrounding stimuli, being for example, preoccupied with his/her thoughts and mind wandering.

Among the above problems the first one is rarely observed in extreme circumstances in critical conditions. Job descriptions, workflow and management decisions are expected to eliminate such scenarios. However, latter three are highly relevant to system design, information presentation, user interaction design and interruption mechanisms. It is therefore essential to establish a successful system-operator interaction with adequate and reliable information flow for a highly complex industrial plant.

An important fact of most industrial plants is that large amounts of staff work together. Some of them work in shifts whereas other work during day work hours. Therefore a collaborative awareness about plant is sought. This type of awareness about the status of the plant processes can be interrupted or totally lost, due to multiple causes such as lack of adequate collaboration or improper handover structure between shifts [7]. These problems can further be worsened through inadequate and mismanaged flow of information to user from the plant, unavailability of temporal and historical information from system monitoring displays and other factors.

It is therefore imperative for a good plant monitoring architecture, to be designed in accordance with expectations and needs of the staff and theoretical foundations of human machine interaction discipline.

The remainder of this paper is organized as follows. In Section 2 we discuss relevant properties of plant monitoring systems. Section 4 gives the results of the survey study which is given in Section 3. We provide results and discussion in Section 4 and Section 5 respectively. Conclusive remarks are given in Section 6.

II. PLANT MONITORING SYSTEMS

Complex industrial plants have numerous activities and processes that are independent of each other while being controlled by an army of staff and workers, where some work at different hours. Although these processes are independent processes and controlled by independent automation and staff, they are also interrelated through product flow, energy and other variables to some degree [8]. Therefore, upper level supervisory monitoring and control, focusing on the whole plant as a single entity is also a necessity. Typical staff types that are in touch with plant monitoring activities are as follows:

- Facility operation engineers, engineers, specialists
- Unit head operators, and unit operators
- Unit and facility operation chiefs, chiefs
- Unit and facility chief engineers, chief engineers, coordinators
- Managers and upper level staff

Considering needs of above different staff members of a plant, a hierarchical multiple level architecture for monitoring and control is necessary as shown in Figure 1.

Due to massive number of plant variables and separated units which are often installed by different vendors at different times, large scale industrial plants have at least four different monitoring levels about ongoing operations. At each level, as one moves up, level of detail is reduced and plant-wide abstraction with data combination is performed. This enables strategic thinking and better management decision making. On the other hand, monitoring tasks at lower levels require more fine grained, localized parametric access, visualization and control.

When an individual process of a plant is concerned, process level monitoring is implemented and staff is only trained and responsible from monitoring only the process that he/she is assigned to. Similarly, middle level engineers, unit chiefs, operation engineers are more interested in examining and monitoring the conditions of units that they are responsible for as a complete set. Upper management on the other hand, mush have access to all parameters in more abstract forms but only in detail when required. Moreover, security and authorization mechanisms must be implemented between different levels and members of all plant monitoring staff.

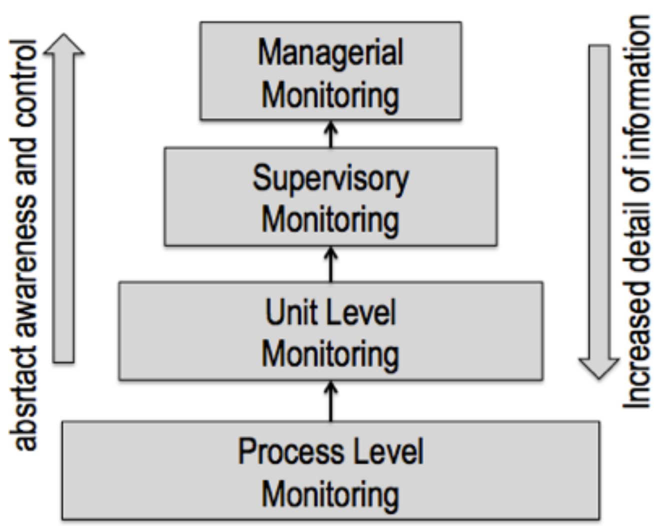


Figure 1. Multiple levels of monitoring in large scale plants.

For safety and successful integration, supervisory monitor level is placed at a level where chief engineers monitor individual units, processes or overall of the plant. At the top, managerial monitoring for decision process and implementation is required to steer the plant with respect to business goals and market conditions, such as accommodation to market conditions and adaptation to input supply levels. Therefore as greater overall control power is expected in managerial level of monitoring, lower details should be visible in order to enable situational awareness of the whole plant variables.

Unlike systems of 1990's or earlier, today's plant control computers are highly connected through plant-wide networks. These are usually number of individual networks such as separate ones for data collection, control and security operations. Typical components of large scale plant information system is shown in Figure 2. Vast number of sensors connected to individual machines, reactors, motors etc. allows plant engineers to collect valuable information remotely through plant-wide information network. However, this benefit poses three challenges:

a) Storing sensor and status data in a central database,
b) Using various collected data for information extraction and automation purposes,
c) Representation and visualization of this data for human consumption, monitoring and decision making.

Therefore, contemporary plant monitoring and control systems with data collection and storage facilities are required in order to keep up with this new kind of high volume data. Furthermore, presentation of such a high volume data in a proper way, so that it can be assimilated and used for decision making by different staff is a challenging task. A "user centric" plant monitoring and control strategy needs to be implemented to achieve high situational awareness about the current and historical status of the plant condition. The following section further discusses stages of plant monitoring and critical components in order to create a user centered modern plant monitoring system.

A. *Several Stages of Monitoring*

Li et. al in their study, suggests to study monitoring process in four stages [9]. These are briefly described as follows:

Detection: The first stage, detection, involves sensing, perception and discrimination of the current state of the process. Thus, early and accurate detection is critical to any successful human intervention. Existing research suggests that 30% of the human error failures occurred at the detection stage [10].

Analysis: The second stage is quite a complicated cognitive process. It usually involves interpreting current process state, reasoning possible causes of any unusual condition, projecting the future process state with or without a specific intervention, planning future actions or assessing the associated risks and competing prioritization of control tasks. These activities require adequate operational knowledge and experience. The research indicate that most operators mainly rely on trend displays, data overview displays and CCTV (Close-Circuit Television) to analyze process status.

Action: This step involves conducting necessary actions in order to meet predetermined goals of plant production based on the cognitive analysis of the data that is assimilated in previous stage.

Evaluation: The last stage, evaluation, basically determines whether the process has been stabilized or not by monitoring the feedback from control system. An operator needs to know what the current process state is, and whether it is moving in the right direction towards production goals, and when he/she can return control to automation if it exists. Usually, operators would target same displays and control system screens used at analysis stage for this evaluation task. There are several factors associated with these stages that may cause incorrect handling of plant operations among these the following can be given as major ones:

"Alarm Flooding" with too many alarms causing ignorance of alarms. Failure to mentally integrate distributed information on screens. Low trust in sensor readings and lack of early detection support on the interface and underlying technology. Lack of in-depth insight of critical process dynamics and lack of predication of future plant state [11]. As a result of above factors, an operator may fall into a situation where overall mental picture of process performance is absent.

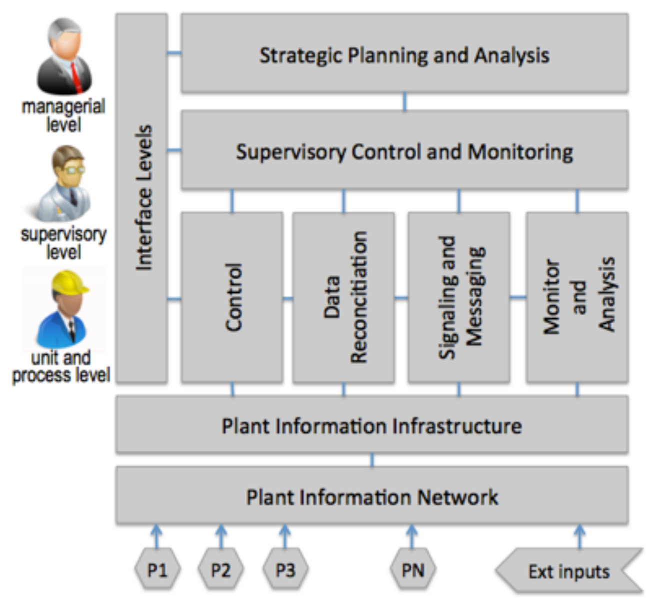


Figure 2. Primary monitoring components of large scale plant management information system.

B. *Proactive Monitoring*

Proactive monitoring is the term that describes monitoring paradigm, that enables operators to take actions before unwanted events occur. Through monitoring proactively, it is hoped that staff would notice problems while they are visible as small events, and can intervene before there are more drastic consequences, such as larger equipment failures, plant destabilization, human injury or loss of production efficiency [12].

A significant portion of proactive monitoring lies under “trend monitoring” and “trend analysis” techniques [13]. Reacting proactively means taking a more comprehensive look at human factors in order to sustain efficient and safe operation of a plant. Trend monitoring techniques rely on collection of significant amount of data and displaying current values along with and with respect to old values indicating trends to human operator [14][15]. Proactive monitoring therefore requires high situational awareness and understanding facts about a plant or unit in which a human monitoring staff is responsible. In order to provide this, such responsible staff should be able to ascertain,

- “What is the system doing now?”
- “Why is it doing that it is doing?”
- “What will it do in near future?”

We believe that well designed human machine interfaces along with adequately trained staff regarding underlying structure of the processes can achieve above three requirements. However, as systems include more automation behind controllers, the problem of “automation opacity” becomes an issue affecting awareness of control/monitor staff about current status of the processes. Abstraction techniques that incorporate and reflect principles of underlying automation (controller behavior) used for such processes, can help to overcome opacity issues and increase proactive monitoring success. Care must be taken when designing interfaces for these systems without obscuring underlying structure and creating false cognitive mental models. The survey study that is explained in Section 3 attempts to gather preliminary information on how information technology and graphical user interfaces can work together helping to achieve better level of proactive plant monitoring.

C. *Information Flow and Messaging*

Large industrial plants have number of people working together, sometimes on the same unit, plant, process and sometimes on different ones. Although most are relatively independent, there are sometimes relations between parameters where one staff is responsible from and other is not in particular. In such cases, monitoring staff usually inform each other via various different information channels. Common methods are calling via telephone, sending e-mail or SMS message, sending paper message or sometimes visiting in person.

Therefore, a contemporary plant monitoring system is necessary, enabling staff interaction through advanced messaging and notification mechanisms. This, in turn will help reduce overhead and redundant messaging and interactions through traditional channels, as well as possible errors. Further, a successful notification system will have learning capabilities, in such a way that certain types of events are automatically routed to responsible monitoring staff once they are initially addressed as a result of manual notification by peers. Messaging and notification architecture should allow labeling certain normal and abnormal cases for further notification processing.

It is also noteworthy to mention that temporal status of the plant units and processes should be carried through between shifts. As mentioned in earlier sections, using advanced trend analysis, display techniques and user message tagging to the trend display data, monitoring staff can have access to all information beyond their allocated time span during their work time.

III. SURVEY STUDY

Based on the literature that we have reviewed, it was found necessary to conduct a preliminary survey study, understanding what staff thinks about current and future monitoring structure. It was planned to make a survey study and then, to use the outcomes in design of the new plant-wide management information and plant monitoring system.

We have conducted a survey study prior designing the plant-wide monitoring for management information systems among the related staff. Based on the previous research, it was essential for the staff to accept the technology and methods in order to fully utilize in the work environment. 22 subjects selected randomly from petroleum plant staff working in various positions (except worker level) related to monitoring tasks are selected. They were interviewed and given a questionnaire regarding the plant-wide monitoring system. Both quantitative and qualitative results are obtained as an input to design of monitoring system.

A. *Questionnaire*

A total of 20 question questionnaire with an addition of open ended free discussion form was given to oil refinery staff. Questions were targeted to understand the attitude and perception of staff towards proactive and enhanced plant monitoring and management information system. Open ended and explanatory questions were also asked in order to gather as much information as possible. Main focus of the questionnaire was to identify how much staff time and effort was currently allocated to plant monitoring tasks and how difficult is to work with the existing structure. Understanding and attitude of staff perception towards plant monitoring were key issues in answering these questions.

The effect of possible new information architecture and staff perception towards this was among expected outcomes of the study. Furthermore, crucial parameters, the expected frequency of monitoring these parameters and additional information to understand daily activities of staff were asked as questions in the study. Prior to giving the questionnaire, full confidentiality were assured, and the purpose of the study was explained.

IV. RESULTS

Tabular results that were obtained from the questionnaires given in Table 1 and Table 2. In terms of a measure for engagement to monitoring activities, subjects indicated that they spend over 75% of their overall work time using computers where 17% of this time (mean value) is dedicated to monitoring activities. 22 parameters must be continuously monitored on average and 16 parameters must be checked daily. Higher number of parameters however (64) must be checked occasionally although variance between subjects was high on that parameter.

TABLE I. CATEGORICAL QUESTIONS REGARDING MONITORING TASKS

No	Question	mean (std) (median)
1	How often you look at the parameter that you are mostly interested in?	3.09 (0,68) (3)
2	Do you think you or other staff made errors in the past in monitoring refinery parameters?	4.35 (0,89) (4)
3	Do you think you spend too much time with software and methods in monitoring parameters?	3.75 (1,26) (4)
4	Do you think there are measurement devices in operation that have errors beyond acceptable limits?	4.55 (0,93) (5)
5	Do you prefer the variables that you follow be represented in graphical formats such as bar charts, pie charts, histograms etc?	4,61 (0,74) (5)
6	Do you believe that graphical representation is not as necessary and you can monitor parameters by looking at numbers?	1,55 (0,83) (1,5)
7	Do you share refinery parameter facts that pulls your attention with your colleagues other than your manager?	4,47 (0,77) (4)
8	Do you need printed material when making decisions regarding plant parameters?	2,69 (1,25) (3)
9	Do you think that better use of information technologies will improve refinery efficiency?	4,75 (0,66) (5)
10	Do you think is it technically beneficial to be able to check parameters that you are responsible by using mobile phone?	4,47 (0,70) (4)
11	Do you think that you know refinery processes well enough?	3,87 (1,21) (4)

Answers for 5-point likert scale (strongly disagree, disagree, neutral, agree, strongly Agree) questions indicated a clear positive attitude towards graphical plant monitoring systems and their benefits. Subjects believed that they or other colleagues made errors in monitoring tasks in the past (4,35). They also believe that measurement devices in the plant might have inaccurate readings (4,55), which is indeed reported in literature among causes of situational awareness problems.

They support the use of graphical content and use of mobile phones to monitor parameters will be beneficial along with more intelligent and proactive information presentation. Subjects felt comfortable with understanding the plant parameters and indicated that they rarely open up printed material in order to understand and solve issues. Open ended questions were related to individual parameters that they are mostly interested in and how they want them to be presented in detail. We have gathered quite helpful individual details, parameter conventions, ranges and suggestions from the survey and follow up interviews after filling the questionnaires. These are beyond the scope of this paper and will be included in the final design.

TABLE II. NUMERICAL QUESTIONS REGARDING MONITORING TASKS

No	Question	mean (std) (median)
1	What percent of your total time do you spend on computers in your job normally?	76 % (24) (90%)
2	What percent of your total computer usage time at work goes to monitoring refinery plant parameters only?	17 % (23) (7,5%)
3	How many plant parameters that you may want to follow continuously?	22 (38) (5)
4	How many plant parameters that you may look at once a day?	16 (24) (6)
5	How many plant parameters that you may look at occasionally?	64 (211) (6)

V. DISCUSSION

Modern plant control and monitoring systems are significantly different from systems of the past. With the help of new advanced information technologies, plant operators seem to have access to all parameters; yet, having access to all parameters causes information overload and failures. To make things worse, advanced process automation techniques that are implemented in modern process control systems may hide certain details from the operator inadvertently, causing automation opaqueness. Moreover, market pressure, environmental concerns and tighter profit margins push plants into operating ranges that are very narrow, which further makes controlling and monitoring more difficult and critical.

A contemporary solution to this should employ user centered design of plant monitoring. Traditional reactive approach should be replaced with proactive techniques. The use of mobile devices, tablets, standalone status displays are reported as beneficial. In order to proactively monitor the system, integration of task elements into a highly graphical and well designed interface, with messaging, notification and data reconciliation properties is suggested. Through this, monitoring staff can achieve higher level of situational awareness and work in parallel to the plant objectives.

There might be other critical dimensions regarding plant-wide monitoring and control in modern plants such as security and vulnerability to malicious software. It becomes extremely critical, when staff start using mobile devices, smartphones and other equipment to reach and control plant parameters. As traditional plant systems are not connected directly to internet and mobile devices and assume presence of certain physical access security, such modern access methods which might seem beneficial at first, require extreme security measures, which are beyond the scope of this paper.

VI. CONCLUSION

We conclude that as recent connectivity and information technology properties of large industrial plants make new types of plant-wide control and monitoring tasks possible, one needs to define how these tasks will be implemented successfully through modern interactive user interfaces and mobile technologies.

It appears that staff work shifts, hierarchical organization structure and common understanding of plant goals are among the essential factors that must be taken into account. We observed that in many situations considerable amount of time required to monitor certain parameters periodically. Therefore, parameter information that is graphical and easy to assimilate will be welcomed both in mobile and desktop settings by plant staff.

We believe that staff surveys, in large plants must be conducted prior to modification and new design of control and monitoring systems. Getting into contact and collecting feedback from plant staff during design stages were found beneficial and encouraging in many directions. We have recently used the results of this study in development of our new plant monitoring software. Initial feedback that we received about our new software were satisfactory. An evaluation of new software is also planned, after being used for about a year for further conclusions.

ACKNOWLEDGEMENTS

We thank TUBITAK, Turkish National Scientific Technical Research Organization for supporting this work through research grant no: 3130475.

REFERENCES

[1] Rogovin, M. (1979). “Three Mile Island: A report to the Commissioners and to the public”, Nuclear Regulatory Commission, Washington, DC (USA), 1979.

[2] Lau, Nathan, Greg A. Jamieson, Gyrd Skraaning jr, and Catherine M. Burns. "Providing Operator Support during Monitoring for Unanticipated Events through Ecological Interface Design." In Proc. of the 29th Annual Canadian Nuclear Society Conference. 2008.

[3] Munro, P. D., & Tilyard, P. A., “Back to the Future - Why Change Doesn’t Necessarily Mean Progress”, Conference Proceedings from the Tenth Mill Operators’ Conference 2009. Tenth Mill Operators’ Conference 2009, Adelaide, South Australia, 2009.

[4] Sarter, N. B., & Woods, D. D., “Situation awareness: A critical but ill-defined phenomenon”, The International Journal of Aviation Psychology, 1(1), pp45-57, 1991.

[5] Endsley, M. R., “Measurement of situation awareness in dynamic systems”, Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(1), pp65-84., 1995.

[6] Hopkin, V.,D, “Situational Awareness in Air Traffic Control”, Situational awareness in complex systems, Proceesings of the Center for Applied Human Factors in Aviation, Conference on Situational Awareness in Complex Systems, Orlando, FL, Ed. Gilson et. al, pp171-178,1994.

[7] Li, X., D. J. McKee, T. Horberry, and M. S. Powell. "The control room operator: The forgotten element in mineral process control." Minerals Engineering 24, no. 8 pp 894-902, 2011.

[8] Woods, David D. "Decomposing automation: Apparent simplicity, real complexity." Automation and human performance: Theory and applications pp3-17,1996.

[9] Li, X, Malcolm S. Powell, and Tim Horberry. "Human Factors in Control Room Operations in Mineral Processing Elevating Control From Reactive to Proactive." Journal of Cognitive Engineering and Decision Making 6.1, pp88-111, 2012.

[10] Hollifield, B., Oliver, D., Nimmo, I., Habibi, E., “The High Performance HMI Handbook: A Comprehensive Guide to Designing, Implementing and Maintaining Effective HMIs for Industrial Plant Operations”, PAS, 2008. Burns, Catherine M.

[11] Hollnagel, E., “Time and time again”, Theoretical Issues in Ergonomics Science, 3(2), pp143-158, 2008.

[12] "Towards proactive monitoring in the petrochemical industry." Safety Science 44, no. 1 pp27-36, 2006.

[13] Yin, Shanqing, Angela Tan, and Martin Helander. "Proactive Process Control Monitoring using Trends." In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 52, no. 24, pp. 2003-2007. SAGE Publications, 2008.

[14] Cook, Maia B., Harvey S. Smallman, and Cory A. Rieth. "Increasing the effective span of control: Advanced graphics for proactive, trend-based monitoring." IIE Transactions on Occupational Ergonomics and Human Factors, 2014.

[15] Charbonnier, Sylvie, Carlos Garcia-Beltan, Catherine Cadet, and Sylviane Gentil. "Trends extraction and analysis for complex system monitoring and decision support." Engineering Applications of Artificial Intelligence 18, no. 1 pp21-36,2005.

Web-based Immersive Panoramic Display Systems

for Mining Applications and Beyond

Dr Tomasz Bednarz
Digital Productivity Flagship
CSIRO
Dutton Park, Australia
e-mail: tomasz.bednarz@csiro.au

Dr Eleonora Widzyk-Capehart
Advance Mining Technology Center
Universidad de Chile
Santiago, Chile
e-mail: eleonora.widzykcapehart@amtc.cl

***Abstract*—The mining industry is interested in novel visualization systems to improve operational efficiency. Technologies to enhance the operator's experience are advancing but there is a lack of evidence supporting the extent to which these emerging technologies positively affect user experience and performance. In this paper, we describe initiative of web based immersive panoramic display system that could be used in mining context. This idea represents a step towards new platforms that will increase the efficiency and safety of the mining operations by video monitoring with annotations of information coming from analytics engines.**

Keywords-Panorama; VR; AR; WebGL; 3-D Annotations.

I. PAST WORK

In the past, we have developed a Panoramic Display System that was effectively used in scientific and industrial environments to visualize 360-degree panoramas in real-time (15FPS) [1]-[3]. The video was constructed out of six streams of images captured by PointGrey LadyBug3 camera [4]. The camera was located in a remote operating location, and stitched video buffers were sent to a location where 3-m hemispherical dome was installed. The stitched images were displayed on the virtual sphere and transformed to a fisheye projection in order to have proper mapping proportions and viewing perspectives, see Figures 1 and 2.

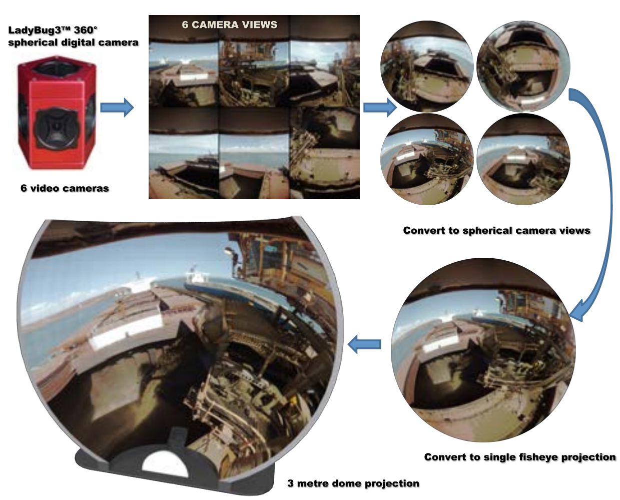


Figure 1. Panoramic Display System – Image Processing Workflow.

Figure 2. Virtual Mining Centre, CSIRO in Pullenvale.

II. WEB BASED PANORAMIC SYSTEMS

The new prototype is currently under development for displaying immersive video environments using technologies supported or natively built into modern web browsers (WebGL [5], WebCL [6], WebRTC [7]). Based on our previous experience, we intend to develop and use the capabilities of 360-degree immersive video in web-based environments, but with playback inside of a Head Mounted Display (HMD), such as, an Oculus Rift [8] (see Figure 3), 3-m hemispherical dome, or in the browser window itself.

Figure 3. Left: panoramic camera. Right: Oculus Rift VR.

The equipment to be used to produce 360-degree videos, is six or twelve GoPro 3+ cameras mounted on 3D printed mounts, 360-HEROS [9]. The videos are to be played back,

 ISBN: 978-1-61208-382-7

inside the web browser (Figure 4), in full-screen mode. WebGL based code is used for rending. Additional functionality will be provided to enable video streaming using standard web technology, WebRTC, for multi-users chat and data communication in web browsers. The intention is to have multiple users existing and observing the same virtual environments. Discussions and measurements of objects in the environment could also occur through placement of 3-D annotations.

III. Conclusions

The present paper described new prototype for displaying robust immersive video environments using modern web technologies natively built into web browsers. Such system would have many potential applications; some of them are listed below:

- 3D movies for training purposes,
- Immersive Tele-conference systems,
- Manufacturing,
- Remote supervision,
- Forensics.

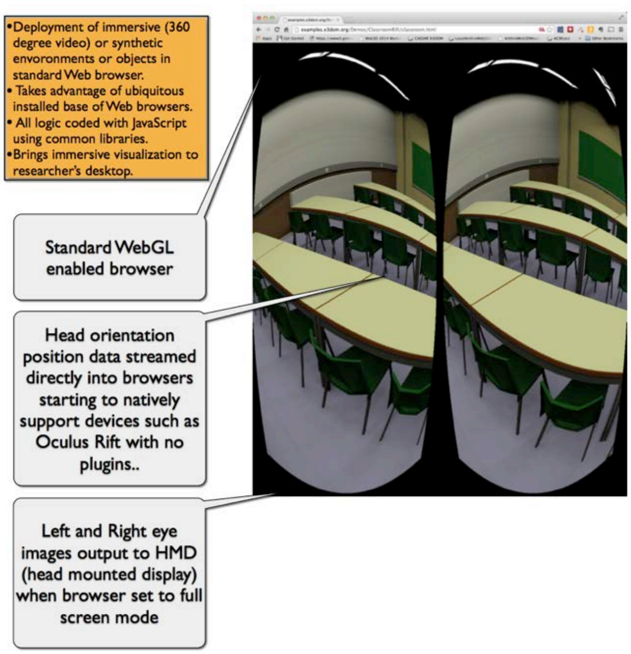


Figure 4. Video playback occuring in web browsers.

References

[1] T. Bednarz, C. James, C. Caris, K. Haustein, M. Adcock, and C. Gunn, "Applications of networked virtual reslity for tele-operation and tele-assistabce systems in the mining industry", VRCAI'11 Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry, 2011, pp. 459-462, doi: 10.1145/2087756.2087845

[2] T. Bednarz, C. Caris, and O. Bass, "Human-computer interaction experiments in an immersive virtual reality environment for e-learning applications", Engineering Education: an Australian Perspective, 2011, pp. 261-269.

[3] C. James, K. Haustein, T. Bednarz, L. Alem, and C. Caris, A. Castleden. "Remote operation of mining equipement using Panoramic Display Systems: exploring the sense of presence", The Ergonomic Open Journal, Vol. 4, 2011, pp. 93-102, doi: 10.2174/1875934301104010093

[4] PointGrey's Ladybug 3 Camera. [Online]. Available from: www.ptgrey.com/products/ladybug3/ladybug3_360_video_camera.asp 2014.10.20

[5] WebGL. [Online]. Available from: www.khronos.org/webgl 2014.10.20

[6] WebCL. [Online]. Available from: www.khronos.org/webcl 2014.10.20

[7] WebRTC (Web Real-Time Communications). [Online]. Available from: www.webrtc.org 2014.10.20

[8] Oculus Rift. [Online]. Available from: www.oculus.com 2014.10.20

[9] 360 HEROS. [Online]. Available from: www.360heros.com 2014.10.20

Combining Image Databases for Affective Image Classification

Hye-Rin Kim
Dept. of Computer Science Yonsei University
Seoul, Republic of Korea
Email: hrkim@cs.yonsei.ac.kr

In-Kwon Lee
Dept. of Computer Science Yonsei University
Seoul, Republic of Korea
Email: iklee@yonsei.ac.kr

***Abstract*—Affective image classification has attracted much attention in recent years. However, the production of more exact classifiers depends on the quality of the sample database. In this study, we analyzed various existing databases used for affective image classification and we tried to improve the quality of the learning data by combining existing databases in several different ways. We found that existing image databases cannot cover the overall range of the arousal-valence plane. Thus, to obtain a wider distribution of emotion labels from images, we conducted a crowd-sourcing-based user study with Amazon Mechanical Turk. We aimed to construct several different versions of affective image classifiers by using different combinations of existing databases, instead of using one. We used low-level features in our classification experiments to explore the discriminatory properties of emotion categories. We report the results of intermediate comparisons using different combinations of databases to evaluate the performance of this approach.**

Keywords–image emotion; emotion-based classification

I. EMOTION-BASED CLASSIFICATION

A. Image collections

Recently, many researchers have reported studies of emotion extraction from images. Several key issues influence the affective classification of images. In particular, it is necessary to obtain ground-truth emotion labels for images. However, obtaining high quality emotion-based images is not easy because of human subjectivity and there are no standard models of emotions. In general, researchers have conducted large-scale user studies to obtain emotion information with two types of emotion models: categorical and continuous models. Categorical models give a discrete value to an emotion using a word, such as happy, sad, or gloomy. By contrast, continuous models represent specific emotions as coordinates in a multidimensional space (a two-dimensional plane is usually preferred, which is called the arousal-valence plane) and we used this type of model in our experiments.

International Affective Picture System (IAPS) is a database of pictures that are used to elicit a range of emotions, which Lang et al. [1] employed in experimental studies of affective image classification. Mikels et al. [2] introduced a subset of the IAPS database for the categorization of images, which we used in our research to obtain the arousal and valence values of the pictures.

Geneva Affective PicturE Database (GAPED) contains 730 images with emotional values [3]. GAPED has four specific types of negative contents, including spiders, snakes, and negative scenes. The positive pictures mainly comprise images of human and animal babies, and nature scenes. The pictures are rated according to their arousal, valence, and congruence values.

The Nencki Affective Picture System (NAPS) [4], is another affective image database, which comprises 1,356 realistic, high-quality photographs with five subject categories (people, faces, animals, objects, and landscapes). The images were given affective arousal and valence ratings by 204 participants, who were mostly European.

Obtaining emotion information using crowd-sourcing Machajdik et al. [5] obtained emotion information based on categorical labels. Furthermore, the range of arousal-valence values is highly limited in other databases, as shown in Figure 1(a). Therefore, we collected arousal and valence values for the images in Machajdik et al.'s database based on a large-scale user survey. A total of 199 subjects were recruited to participate in the survey using Amazon Mechanical Turk and the subjects provided 6787 responses. We collected at least six responses for each image and each subject provided an average of 33 responses. Figure 1(b) shows the distribution of the emotion labels obtained in the survey, which demonstrates that the combined database was more evenly distributed in the arousal-valence plane compared with the original database.

B. Image Features

In this study, we applied most of the features used in previous studies, which are mainly related to color and texture. In addition, we used a new feature called color harmony (f31, f32 in TableI), which is based on color perception theory. Recently, several statistical studies have proposed methods for computing the harmony between colors. We employed one of these methods [9] to compute the harmony between

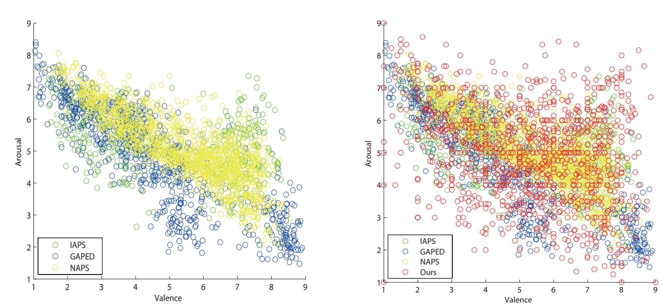


Figure 1. (a) Arousal-Valence distribution of images using three existing databases. (b) our user-study results are added (red dots)

 ISBN: 978-1-61208-382-7

TABLE I. OVERVIEW OF FEATURES IN OUR METHOD.

Feature	Description	character	Feature	Description	character
f1, f2, f3	The histogram of hue, saturation and value of image	color	f21, f22, f23	Average saturation for the first, second and third largest segment	color
f4, f5, f6	Average of hue, saturation and value of image	color	f24, f25, f26	Average value for the first, second and third largest segment	color
f7	The hue section that used in image over threshold	color	f27	Color descriptor in [6]	color
f8	The number of hue sections that used in image over threshold	color	f28	Color consistancy in [7]	color
f9, f10, f11	Activity, Weight and Heat of image [8]	color	f29	The existance of basic color	color
f12, f13	Mean and standard deviation of the magnitude of Gabor filtered image	texture	f30	The number of used colors for each basic colors	color
f14, f15, f16, f17	Energy, Entropy, Contrast, Homogeneity of gray scale image	texture	f31	Average color harmony of the most used ten colors	color
f18, f19, f20	Average hue for the first, second and third largest segment	color	f32	Color harmony between two colors among the ten representative colors	color

TABLE II. CLASSIFICATION PERFORMANCE USING VARIOUS COMBINATIONS OF DATABASES.

Database	5 fold cross validation	No. of images
GAPED	0.80	730
GAPED+NAPS	0.68	2086
GAPED+IAPS	0.64	1119
NAPS	0.60	1356
NAPS+IAPS	0.59	1745
IAPS + Machajdik + GAPED+ NAPS	0.54	3561
NAPS + Machajdik	0.54	2044
GAPED + Machajdik	0.54	1816

representative colors in image. For each image, we extracted 10 representative colors using k-means clustering and we then computed the harmony among all of the colors. The features used in this study are listed in Table I.

II. Classification

Given a set of features, we aimed to construct an appropriate classifier to estimate the emotion in a given image. We used the public library A Library for Support Vector Machines [10] to compute the nonlinear hyperplanes for class separation. To evaluate the classification performance, we divided the emotion space into four classes where the point (5, 5) was at the center of the arousal and valence axes. Based on the ratings in the database, all of the images were labeled according to one of the four classes for training. We performed a 5 fold cross-validation because we lacked a ground-truth database. The classifier was trained using various combinations of databases. Table II shows the classification performance based on 4 four categories in for each combination. The results show that the GAPED database recorded the best performance in with our scheme so far.

III. Conclusion

In this study, we compared the affective classification performance of different combinations of existing image databases, where we included the results of a user study to compensate for the lack of data. The main contributions of our study can be summarized as follows: 1) We performed a crowd-sourcing-based user survey to collect emotion information for a large set of images; 2) We evaluated emotion-based image databases using various combinations of categories. There is no research for affecitve classification using the combination of various databases. Therefore, we tried to find a research using GAPED database which recorded the best performance in our scheme, but couldn't find it. Statistically, the accuracy for catogorical affective classification is less than 80%. We leave the exact comparison with other methods for future work. We will also construct a more appropriate regression-based model to estimate the arousal and valence coordinates for images. In addition to low-level features, we may consider the use of high-level semantics to obtain better performance, which are employed widely in aesthetics as new features.

Acknowledgment

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2014.

References

[1] P. Lang, M. Bradley, and B. Cuthbert, "International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. Technical Report A-6," The Center for Research in Psychophysiology, University of Florida, 2005.

[2] J. A. Mikels, B. L. Fredrickson, G. R. Larkin, C. M. Lindberg, S. J. Maglio, and P. A. Reuter-Lorenz, "Emotional category data on images from the International Affective Picture System," Behav Res Methods., vol. 37, 2005, pp. 626–630.

[3] E. S. Dan-Glauser and K. R. Scherer, "The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance," Behav Res Methods., vol. 43, 2011, pp. 268–277.

[4] A. Marchewka, L. Zurawski, K. Jednorog, and A. Grabowska, "The Nencki Affective Picture System (NAPS). Introduction to a novel standardized wide range high quality realistic pictures database," Behav Res Methods., vol. 46, 2014, pp. 596–610.

[5] J. Machajdik and A. Hanbury, "Affective Image Classification Using Features Inspired by Psychology and Art Theory," in Proceedings of the International Conference on Multimedia, ser. MM '10. ACM, 2010, pp. 83–92.

[6] J. van de Weijer and C. Schmid, "Coloring Local Feature Extraction," vol. 3952, 2006, pp. 334–348.

[7] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-Based Color Constancy," Image Processing, IEEE Transactions on, vol. 16, no. 9, 2007, pp. 2207–2214.

[8] M. Solli and R. Lenz, "Color Based Bags-of-Emotions," in Computer Analysis of Images and Patterns, vol. 5702. Springer Berlin Heidelberg, 2009, pp. 573–580.

[9] L.-C. Ou and M. R. Luo, "A Colour Harmony Model for Two-Colour Combinations," Color Research and Applications, vol. 31, no. 3, 2006, pp. 191–204.

[10] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," vol. 2, 2011, pp. 27:1–27:27, software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm [retrieved: December, 2014].

 ISBN: 978-1-61208-382-7

Automatic Creation of a HLA Simulation Infrastructure for Simulation-Based UI Evaluation in Rapid UI Prototyping Processes

Bertram Wortelen and Christian van Göns

OFFIS – Institute for Information Technology
26121 Oldenburg, Germany
Email: {wortelen|vangoens}@offis.de

***Abstract*—Evaluating user interfaces with virtual user models is a means for rapid prototyping. Setting up a simulation environment for virtual user models often requires high effort due to the heterogeneous simulation tools. Furthermore, the frequent reconfigurations of the simulation due to the rapid changes of the user interface prototypes impose a high amount of workload upon the user. In particular, the manual reconfiguration of the communication between the simulation components is very complex and error prone. Small changes to the user interface often result in changes in the communication of several components. Our solution is the automatic generation of the communication data description for all simulation components. This paper presents the implemented solution and illustrates it with two scenarios from the maritime domain. These scenarios deal with collision avoidance strategies and new concepts for route exchanges between ships and vessel traffic service centres. The automated generation process facilitates handling the emerging changes, which are required in the complex simulation configurations. The evaluation of how well this supports the rapid prototyping process in these scenarios is not addressed in this paper, but is the topic of ongoing research.**

***Keywords*–rapid prototyping; virtual user models; co-simulation; user interface evaluation.**

I. Introduction

The dynamic simulation of virtual user models (VUMs) in interaction with the user interfaces (UIs) is a means to evaluate UI designs. Examples can be found in various domains. In the automotive domain Salvucci [1] uses a virtual driver model to predict the distractive effects of in-vehicle interfaces on the driving performance. Wortelen et al. [2] use a virtual driver model to simulate attention distribution among in-vehicle interfaces and other information sources. In usability research on desktop applications, the Cogtool approach focusses on the prediction of execution times by simulating VUMs [3]. Virtual seafarer models have been used in the maritime domain by Sobiech et al. [4] to test an adaptive bridge system. Lüdtke et al. [5] predicted pilot errors in commercial aircraft cockpits on a virtual pilot model.

The effort to set up a simulation environment for VUMs and UIs can be quite high. Cogtool [3] and ACT-CV [6] are two independent applications that strongly reduce this effort, but are limited to the UIs of desktop application. Both tools enable VUMs created with the cognitive architecture ACT–R to interact with a real UI (ACT-CV) or a UI mock-up (Cogtool) to predict for example task execution times. However, the UI of desktop application is often restricted to one screen, keyboard and mouse. In more complex environments like cars, cockpits, ship bridges or control rooms, it requires far more manual tasks to connect the VUMs to the UIs. Main reasons for this are the very different physical UIs and the heterogeneous tools used for simulation. In order to simulate a realistic scenario in a car or on a ship bridge, a simulation of the environment is required. Another reason can be found in the cooperative nature of many of the above mentioned scenarios, where a task is not executed by just one but multiple human operators, like in aircraft cockpits or on ship bridges. Thus, multiple VUMs might be required.

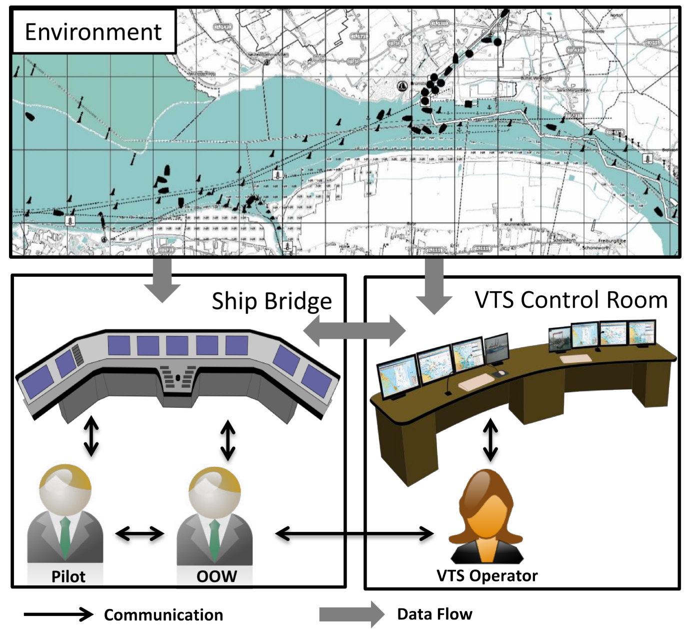


Figure 1. The use case scenario: A VTS operator cooperates with a ship master on a ship bridge, e.g., to inform him about a course conflict and negotiate a new course for the ship.

Puch et al. [7] provide a software framework that aims to reduce the technical challenges of connecting the heterogeneous tools involved in the simulation. They implemented parts of the IEEE 1516 standard on the High Level Architecture (HLA) framework [8], and tailored it to the needs of virtual user model simulations. This paper builds upon their work and shows how parts of their simulation framework can be automatically configured based on a model of a UI prototype.

The work is demonstrated with two case study scenarios from the maritime domain. The involved components can be seen in Figure 1. These scenarios involve a UI system used

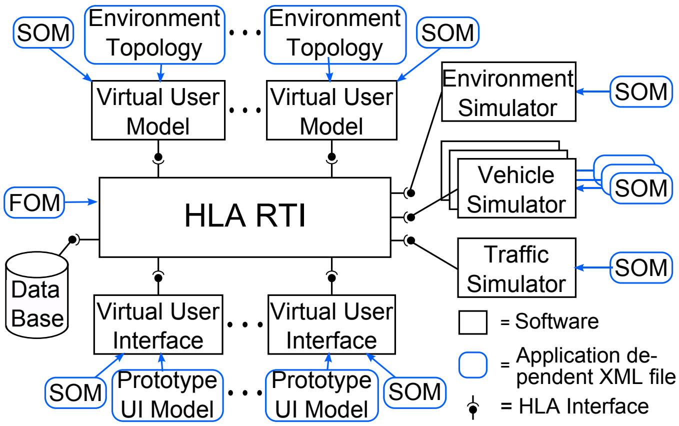


Figure 2. General HLA setup for virtual user simulations in scenarios from the transportation domain.

on a ship bridge operated by two users: The ship master and the pilot. It furthermore contains a Vessel Traffic Service (VTS) control room as a second UI system. A VTS operator is interacting with this UI, while communicating with the pilot and master on the ship bridge.

The remainder of this paper starts with a description of the typical settings for which the presented approach applies and a description of the components involved in the simulations. Afterwards in Section II, it is shown which parts of the simulation configuration can be automatically created. In the same section the algorithm for the automatic creation of the simulation configuration is described. Finally, the approach is illustrated based on maritime case study scenarios.

II. Components of the Simulation

This work builds upon the HLA simulation software provided by Puch et al. [7]. The general setting that is considered can be seen in Figure 2. It shows the different kinds of components that are connected to the HLA simulation.

The core of the simulation is the HLA Runtime Infrastructure (RTI), to which all other components (*federates*) are connected. The HLA standard defines an Object Model Template (OMT), which describes the data that is exchanged between the federates. Each federate defines the data objects it sends or receives in a Simulation Object Model (SOM). The SOM is specified in a dedicated XML file for each federate (see Figure 2). The Federation Object Model (FOM) is a superset of all SOMs and defines all objects that can be exchanged in the simulation.

In cooperative scenarios there can be multiple users interacting with multiple user interfaces. A user interface might even be shared by multiple users, e.g., in aircraft cockpits or on ship bridges. Besides the user and interface simulation the dynamics of the environment is an important aspect of the simulation. In scenarios from the transportation domain, this typically includes the behaviour of the vehicles steered by the user models, the surrounding traffic and further aspects of the environment like weather conditions, road topology or any kind of events.

In the following the data exchange between the environment simulation and the simulation of the UIs is not addressed. It is expected that the data exchange between these components

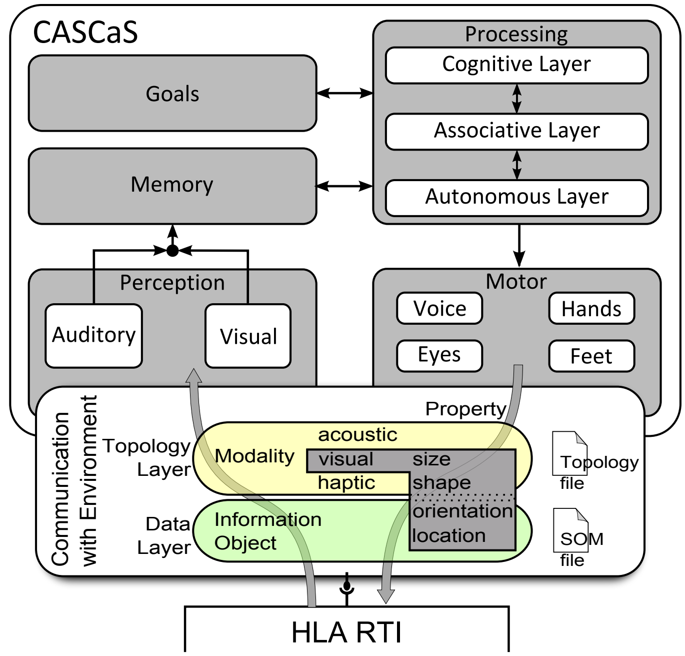


Figure 3. High-Level structure of the Cognitive Architecture CASCaS. At the bottom the two layers of the communication with the environment are shown.

changes infrequently in the scenarios we intend to support. The focus of design is the UI of the workstations that the human operators interact with. Changes in the UI design lead to changes in the communication between the workstation UI and the operator, but not necessarily between the workstation and the environment. In order to analyse multiple UI design proposals with VUMs, it should be easy to change the communication structure between VUM and UI (marked as black arrows in Figure 1).

A. Virtual User Models

There are several ways to create virtual user models. Cognitive Architectures are a good approach to create VUMs, which are capable of simulating human behaviour. A cognitive architecture consists of an integrated set of general models about different aspects of human cognition, like memory or attention processes. These models are defined independently of the task that the user is working on. Human behaviour is shaped by these task-independent aspects. However, the current task is also an important driver of human behaviour.

Cognitive architectures load a formal model of the user's task procedures, typically as a set of production rules. A cognitive architecture interprets and applies these rules based on its internal state and the information that it perceives from the environment. Figure 3 shows the structure of the main components of the Cognitive Architecture for Safety Critical Task Simulation (CASCaS) [9]. This is the architecture used in the case study to create the virtual user models, i. e., cognitive seafarer models. Like many other cognitive architectures it contains components to simulate perceptual and motoric processes. Information perceived from the environment and knowledge stored in a dedicated memory component is processed in an central information processing component. The aim of CASCaS is to simulate goal-directed human behaviour.

 ISBN: 978-1-61208-382-7

Therefore, a goal component is defined that simulates how humans divide their attention between multiple task goals.

In order to interact with the UI, the cognitive architecture needs to know which information objects are displayed on the UI and which actions are available on the UI. Furthermore, the information modality of the information and actions needs to be specified. Is information displayed as text, as graphical icon, as voice message or acoustic signal? Is an action executed on a touch screen, a wheel or lever? This kind of information defines the *environment* that the VUM can interact with.

The data exchange with the environment of a VUM created with CASCaS differentiates between the two aspects of information mentioned above (information objects and information modality). The information modality also includes the 3-dimensional location of the information objects in order to properly simulate eye or hand movements and different visual fields (foveal and peripheral).

The data exchange with the environment is organized in two layers (see bottom of Figure 3). The information itself is transported on the Data Layer, while the modality is defined in the Topology Layer. In the setting described in this paper the technology used for implementing the Data Layer is HLA.

On the Topology Layer, the modality is defined in a separate topology file. Based on the modality further properties can be defined in the topology to specify in more detail, how the information is presented to the VUM, like size and location of an information object. It is differentiated between static and dynamic properties. If the property does not change over time it is statically declared and defined in the topology. If it changes over time, it is declared in the topology, but the actual value of the property is dynamically received via the Data Layer.

This is illustrated on the case study scenario. Consider the current heading of a ship as an information object `heading`. It can be visualized on a display in a text field at a fixed position. In that case the modality properties do not change over time. Alternatively, the heading information can also be visualized on an electronic sea chart as a ship icon that is oriented in the direction of the current heading. The location of the icon moves with the movements of the ship. In that case the location and orientation properties have to be received dynamically via the Data Layer.

Besides the declaration in the topology file all information communicated via the Data Layer needs to be specified in the SOM files to inform the HLA software about the exchange (see top of Figure 2).

B. Virtual User Interface

A VUM does not interact directly with the physical interface. That is to say, it does not move levers or type texts. These interactions are simulated on a virtual user interface (VUI) and forwarded to the underlying control logic of the technical system or to a software mock-up.

A virtual user interface should be tailored to the cognitive architecture that is used for the evaluation of the user interface. Different cognitive architectures are able to interact with different kinds of information. For example, some architectures focus on the simulation of reaction to salient information like a flashing light in its peripheral view [10]. Others are able to use properties of visual objects like colour, orientation or shape to guide visual attention in search tasks [11]. Thus, a VUI only needs to consider those aspects that the cognitive architecture can process in a meaningful way. It mainly defines the level of abstraction of the VUI. Keeping the description language of a VUI simple makes it easier to change a UI prototype during the design phase. This again facilitates rapid evaluation of different design concepts.

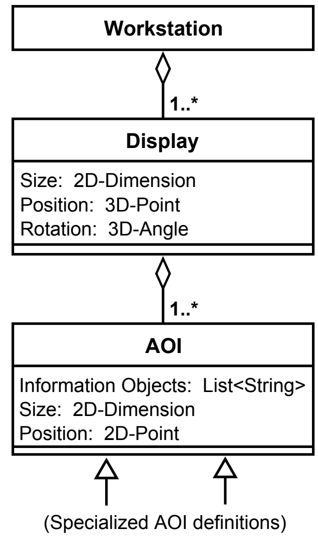


Figure 4. UML model of a workstation. It consists of a set of displays and AOIs located on it. An AOI specifies the information objects it presents.

For the simulation of the virtual user interfaces a framework presented by Sobiech et al. [4] is used. It defines UI models in an XML-based format that only considers the properties size, location and orientation for visually communicated information. Each VUI describes the layout of an entire workstation that may contain multiple displays. For each display multiple areas of interest (AOIs) are defined, which again contain multiple information objects (see Figure 4).

Each information object in an AOI is specified by its name. Within the VUI framework this name is linked to an information source. This source can be a simulated sensor, e.g., from a high-fidelity vehicle simulation or environment simulation, or it can be a real-world data record form a database (see Figure 2). This representation is often sufficient for a large part of the workstation. However, it might be beneficial to represent some AOIs in more detail - especially complex and dynamic AOIs like radar screens or sea charts. In such a case a specialized AOI class needs to be derived from the AOI class (Figure 4), that also defines the simulation of the dynamic properties.

III. Creation of Data and Topology Layer Definitions

Using the HLA software developed by Puch et al. [7], CASCaS has been integrated in simulations in the automotive [12] and maritime [4] domain. The VUI framework presented by Sobiech et al. [4] was used to prototype interfaces on ship bridges. Both tools facilitate setting up a working simulation environment. However, there is still a task that has been shown to be prone to errors and needs to be done repetitively in rapid prototyping scenarios. If a UI prototype is changed and re-evaluated, the information objects exchanged between VUI and

 ISBN: 978-1-61208-382-7

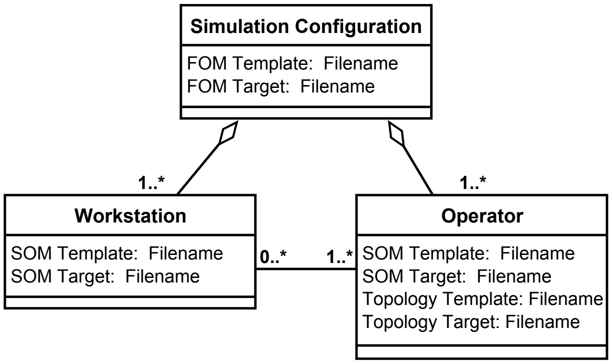


Figure 5. UML model of the simulation configuration.

VUM typically also change. This requires changes in several documents (see Figure 2):

- The SOM definitions of VUI and VUM federates
- The FOM definition of the federation in order to reflect the changes in the SOM definitions
- The topology definitions of the VUMs

Changes in the VUI design typically also lead to changes in the procedure of the VUM, because users interact in different ways with different UI elements. The question, how a change in a VUI design can efficiently be reflected in the VUM, is focus of ongoing research and is not considered in this paper.

Updating the above listed documents is a simple but error prone task. It involves manual edits in several files. Our goal is to support this task and ease the rapid evaluation of different UI designs. The approach taken here is to define all aspects of the entire simulation configuration that are relevant for the communication between VUIs and VUMs in a simple way in a single repository.

The model of the simulation configuration can be seen in Figure 5. All aspects of the communication between VUIs and VUMs in the object model files and topology files can be created automatically from the model of the simulation configuration. Any changes in the configuration lead to an automatic modification of all involved simulation documents.

A configuration consists of one or more operators and workstations. An operator can be assigned to multiple workstations. A workstation can be operated by one or more operators. These relations are defined in the simulation configuration model. The data displayed on the workstation and the commands available on the workstation define, what is exchanged between VUI and VUM. This kind of information is specified for each workstation in its workstation model (Figure 4).

The data exchange between the simulation components is not limited to exchange between VUMs and VUIs. VUIs and VUMs also interact with the simulation of the environment and the technical system controlled by the VUM via the VUI. The object models of this kind of data exchange must also be included in the SOM, FOM and topology definitions. The data exchange with the environment typically does not change when the UI design is changed. Therefore, in rapid prototyping scenarios we do not expect many changes for these object models, if any. We assume that the data exchange between environment and VUI/VUM is already specified in the

```
OP ← All operators in the simulation configuration
WS ← All workstations in the simulation configuration
ws.aois ← All AOIs in the workstation definition ws.

function CreateObjectModelDefinitions(s)
    for all ws ∈ WS do
        for all aoi ∈ ws.aois do
            Insert aoi definition in SOM target of ws;
            Insert aoi definition in FOM target;
        end for

        OP_ws ← List of operators, that have access to ws;

        for all op ∈ OP_ws do
            for all aoi ∈ ws.aois do
                Insert aoi definition in SOM target of op;
                Insert aoi definition in topology target of op;
            end for
        end for
    end for

    for all op1 ∈ OP do
        for all op2 ∈ OP do
            if op1 ≠ op2 then
                Insert object models for voice
                communication from op1 to op2 in
                    1) topology target of op1
                    2) topology target of op2
                    3) SOM target of op1
                    4) SOM target of op2
                    5) FOM target
            end if
        end for
    end for
end function
```

Figure 6. Creation of all SOM/FOM and topology files.

topology and SOM/FOM files. Dibbern et al. [13] presented a method that enables the automatic generation of SOM and FOM specifications related to the data exchange to the environment simulation.

The VUI/VUM related object models need to be integrated into the existing topology and SOM/FOM files. The simulation configuration model therefore references two types of files for topologies and SOM/FOM definitions: Templates and Targets (see Figure 5). The template files contain the environment related object models. The target files are created automatically by inserting the VUI/VUM related object models. How this is done is described as pseudo-code in Figure 6.

The algorithm iterates over all workstations defined in the simulation configuration. For all aois defined on any display of a workstation, the information objects communicated via the AOI are defined in the SOM target file of the workstation. The same is done in the target file for the federation object model (FOM). Furthermore, the algorithm iterates over all operators, that have access to the workstation and inserts the definitions for the communicated information objects into the SOM files of the respective operator. It also generates the layout and modality definitions in the topology file of the operator.

```xml
<Configuration>
  <!-- Workstation definitions -->
  <Workstation name="Bridge">
    <Display filename="conning.xml"/>
    <Display filename="ecdis.xml"/>
    <Display filename="radar.xml"/>
    <Operator name="Master"/>
    <Operator name="Pilot"/>
  </Workstation>
  <Workstation name="VTS">
    <Display filename="vts.xml"/>
    <Operator name="VTS"/>
  </Workstation>
  <!-- Global fom file-->
  <FOM template="TmpFom.xml" target="GenFom.xml"/>
  <!-- Som File for VUI-->
  <SOM template="TmpSomVUI.xml"target="GenSomVUI.xml"/>
  <!-- Operator definitions. -->
  <Operator name="Master">
    <SOM template="TmpSomBridge.xml" target="GenSomM...
    <Topology template="TmpTopBridge.xtop" target="G...
  </Operator>
  <Operator name="Pilot">
    <SOM template="TmpSomBridge.xml" target="GenSomP...
    <Topology template="TmpTopBridge.xtop" target="G...
  </Operator>
  <Operator name="VTS">
    <SOM template="TmpSomVTS.xml" target="GenSomVTS.x...
    <Topology template="TmpTopVTS.xtop" target="GenTo...
  </Operator>
</Configuration>
```

Figure 7. Configuration file with three operators and two workstations.

Finally the communication between each VUM is defined. Users do not only interact with the technical system, but also with other users. Therefore, information objects are defined in the SOM targets of each operator. These information objects describe the information exchange via voice communication between each other operator. In the topology target these information objects are defined as voice actions (speaker) or as acoustic voice information (listener).

IV. Case Study: Maritime Systems

The described approach is tested in the context of eMIR (eMaritime Integrated Reference Platform) [14]. eMIR is a reference platform that provides a means for testing and demonstrating new maritime safety systems or eNavigation systems. eMIR combines a physical test bed in the south east German Bight with a simulation-based software test bed.

For the case study a pure simulation-based setting is used, in which the environment, traffic, the ship dynamics, user interfaces and all operators are simulated. Simulation components for the environment, maritime traffic and the ship dynamics are available in the eMIR simulation test bed. In this section only the VUIs and VUMs are described. Two scenarios are planned in which new user interface concepts will be evaluated.

1. Scenario: **Route exchange.** This scenario deals with a new system for exchanging routes between ships and vessel traffic services (VTS) [15]. It considers the ship master and the VTS operator as human operators, who interact with two different workstations (the bridge system and the VTS system).

2. Scenario: **Collision avoidance.** It deals with cooperative tasks in a collision avoidance scenarios. For this scenario two operators (pilot and ship master) are considered, who interact with one workstation (the bridge system).

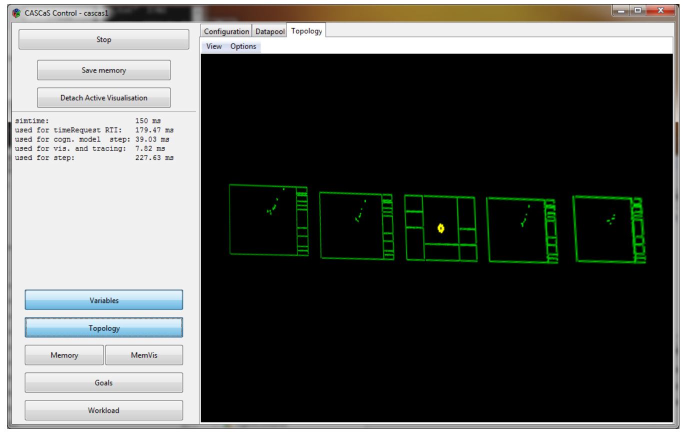


Figure 8. Representation of dynamic topology in CASCaS.

Both scenarios share common components (ship master and bridge system). The automatic generation of the simulation configuration eases the reuse of these components. Even a scenario involving all operators and UIs can be defined. The corresponding configuration file is shown in Figure 7.

All three operators are modelled with the cognitive architecture CASCaS. A formal description of the operators' task procedures is used as input for CASCaS to simulate the operators' behaviour. We aim at building up a knowledge base of these procedures. The current focus is on the procedures of the ship master. The procedures are obtained from different sources. The first step was a literature review. The high level structure of the seafarer's tasks is based on the Operator Function Model (OFM) of navigation tasks defined by Sanquist et al. [16]. The OFM is a framework for representing how a human operator might decompose control functions to meet system objectives and ensure system safety [17]. The OFM is not on a level of detail that allows dynamic simulation. Field observations on ship bridges and VTS control rooms, as well as interviews with the operators, are conducted to gain knowledge about details of certain task procedures. This is work-in-progress.

For the simulation of the UIs of the VTS control room and the ship bridge the VUI framework presented in [4] is used. Our case study scenarios only focus on UI concepts for the support of navigation tasks. Thus, the VUIs do not model the entire bridge system, but only those parts relevant for navigation. This is mainly the Electronic Chart Display and Information System (ECIDS) and the Radar system. The same applies for the operator procedures. These also focus on the relevant navigation tasks.

For the bridge system, Figure 8 shows how CASCaS internally represents the topology layer. What can be seen are the information of the Topology Layer in its current state for all visual information objects of the bridge system. The wireframes show the outlines of all AOIs from all displays defined in the topology of the virtual ship master model in CASCaS. Represented are two ECIDS displays, two Radar displays and one conning display. This topology has been created by the algorithm in Figure 6 based on its VUI model and is used for the virtual model of the pilot and the ship master. The VUI only describes those aspects of the UI, that can be interpreted by CASCaS and are relevant for the investigated tasks.

After all VUMs have been sufficiently defined, the next step will be to evaluate how well the simulation-based approach is able to support rapid prototyping in such complex and safety-critical domains.

V. Conclusion and Future Work

In this paper, a simple algorithm was presented that creates the configuration files for the Data Layer and the Topology Layer required for co-simulations of virtual user models and virtual user interfaces. HLA is used as simulation infrastructure. CASCaS is used as architecture for the virtual user models. The VUI framework presented by Sobiech et al. [4] is used for the virtual user interfaces.

Even though, there is no standard for connecting cognitive models with environment or user interface simulations, the information that needs to be interchanged is essentially the same, regardless of the specific cognitive architecture or user interface description language that is being used. The main difference when using other cognitive architectures like ACT–R [18], or other user interfaces description languages like UsiXML [19] is the representation of this information. Thus we believe that the general approach presented in this paper can also be used in similar settings. However, there definitely are differences between different cognitive architectures and different user interface description languages. Therefore using a different one, would require a re-implementation of the generation algorithm shown in Figure 6.

The automatic generation of the files for the Data and Topology Layer lessens the effort for changing UI concepts in a simulation with virtual user models. This is a prerequisite for efficiently doing rapid UI prototyping. The next steps in our research will be to use this algorithm in actual UI prototyping processes for the described maritime scenarios. We will explore whether it is sufficient for fast and efficient simulation-based evaluation of UIs in rapid prototyping processes.

Furthermore, we believe that our approach also supports the development of normative task procedures in the same way that it supports the development of UIs. New procedure concepts can be evaluated by using the VUMs to simulate the procedures and analysing the simulated behaviour. This is planned to be focus of a subsequent research step.

Acknowledgment

This research has been performed with support from the Cosinus project (http://www.emaritime.de/projects/cosinus/), funded by the *Bundesministerium für Wirtschaft und Technologie (BMWI)* in the framework programme *Next-Generation Maritime Technologies (2011-2015)*, and with support from the EU FP7 project CASCADe, GA N.: 314352. Any contents herein are from the authors and do not necessarily reflect the views of the European Commission.

References

[1] D. D. Salvucci, “Rapid prototyping and evaluation of in-vehicle interfaces,” ACM Transactions on Computer-Human Interaction, vol. 16, no. 2, Juni 2009, pp. 9:1–9:33.

[2] B. Wortelen, M. Baumann, and A. Lüdtke, “Dynamic simulation and prediction of drivers’ attention distribution,” Transportation Research Part F: Traffic Psychology and Behaviour, vol. 21, 2013, pp. 278–294.

[3] B. E. John, “Using predictive human performance models to inspire and support ui design recommendations,” in Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, ser. CHI ’11. New York, NY, USA: ACM, 2011, pp. 983–986.

[4] C. Sobiech, M. Eilers, C. Denker, A. Lüdtke, P. Allen, G. Randall, and D. Javaux, “Simulation of socio-technical systems for human-centred ship bridge design,” in Proceedings of International Conference on Human Factors in Ship Design & Operation 2014. 8-9 Northumberland Street, London: The Royal Institution of Naval Architects, February 2014, pp. 115–122.

[5] A. Lüdtke, J.-P. Osterloh, T. Mioch, F. Rister, and R. Looije, “Cognitive modelling of pilot errors and error recovery in flight management tasks,” in Proceedings of the 7th Working Conference on Human Error, Safety and Systems Development Systems Development (HESSD), ser. LNCS 5962, M. W. Jean Vanderdonckt, Philippe Palanque, Ed., vol. 1. Springer, 2009, pp. 54–67.

[6] M. Halbrügge, “Bridging the gap between cognitive models and the outer world,” in Grundlagen und Anwendungen der Mensch-Technik-Interaktion. 10. Berliner Werkstatt Mensch-Maschine-Systeme, E. Brandenburg, L. Doria, A. Gross, T. Günzler, and H. Smieszek, Eds. Berlin: Universittsverlag der TU Berlin, October 2013, pp. 205–210.

[7] S. Puch, M. Fränzle, J.-P. Osterloh, and C. Läsche, “Rapid virtual-human-in-the-loop simulation with the high level architecture,” in Proceedings of Summer Computer Simulation Conference (SCSC) 2012. Curran Associates, Inc., July 2012, pp. 44–50.

[8] 1516-2010 IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA), IEEE Computer Society Std., August 2010.

[9] A. Lüdtke, J.-P. Osterloh, L. Weber, and B. Wortelen, “Modeling pilot and driver behavior for human error simulation,” in Digital Human Modeling, ser. Lecture Notes in Computer Science, V. G. Duffy, Ed. Springer, Berlin, 2009, vol. 5620/2009, pp. 403–412.

[10] R. Rukšėnas, J. Back, P. Curzon, and A. Blandford, “Verification-guided modelling of salience and cognitive load,” Formal Aspects of Computing, vol. 21, no. 6, 2009, pp. 541–569.

[11] J. M. Wolfe, “Guided search 4.0: Current progress with a model of visual search,” in Integrated Models of Cognitive Systems, 1st ed., W. D. Gray, Ed. New York: Oxford: Oxford University Press, April 2007, ch. 8, pp. 99–119.

[12] S. Puch, B. Wortelen, M. Fränzle, and T. Peikenkamp, “Evaluation of drivers interaction with assistant systems using criticality driven guided simulation,” in Digital Human Modeling and applications in Health, Safety, Ergonomics and Risk Management, ser. LNCS, V. G. Duffy, Ed., vol. 1, no. 8025. Springer, 2013, pp. 108–117.

[13] C. Dibbern, A. Hahn, and S. Schweigert, “Interoperability in co-simulations of maritime systems,” in Proceedings of 28th European Conference on Modelling and Simulation, Brescia, Italy, May 2014.

[14] eMIR, “eMaritime Integrated Reference Platform,” retrieved: 2015.01.08. [Online]. Available: http://www.emaritime.de

[15] A. Bolles, A. Hahn, M. Braun, S. Rohde, M. Gluch, and K. Benedict, “Cosinus – cooperative vessel guidance for nautical safety,” in Proceedings of the International Symposium Information on Ships, ISIS 2014, Hamburg, Germany, 04.-05. September 2014.

[16] T. F. Sanquist, J. D. Lee, and A. M. Rothblum, “Cognitive analysis of navigation tasks: A tool for training assessment and equipment design,” U.S. Department of Transportation – United States Coast Guard, Interim Report, April 1994, retrieved: 2015.01.08. [Online]. Available: http://www.dtic.mil/dtic/tr/fulltext/u2/a284392.pdf

[17] E. Palmer and C. M. Mitchell, “Operator function modeling: Cognitive task analysis, modeling and inteligent aiding in supervisory control systems,” Center for Human-Machine Systems Research, School of Industrial & Systems Engineering, Georgia Institute of Technology, Tech. Rep., September 1990, NASA Ames Grant NAG 2-413. Retrieved: 2015.01.08. [Online]. Available: https://archive.org/details/nasa_techdoc_19910002067

[18] J. R. Anderson, How can the Human Mind Occur in the Physical Universe, ser. Oxford Series on Cognitive Models and Architectures. Oxford University Press, August 2009.

[19] UsiXML 2010, Proceedings of the 1st International Workshop on USer Interface eXtensible Markup Language, 2010, Retrieved: 2015.01.08. [Online]. Available: http://www.usixml.eu/sites/default/files/FaureVanderdonckt-UsiXML2010.pdf

Sentiment Classification for Chinese Microblog

Wen-Hsing Lai, Chang-Hsun Li
Department of Computer and Communication Engineering
National Kaohsiung First University of Science and Technology
Kaohsiung, Taiwan
e-mail: lwh@nkfust.edu.tw

***Abstract*—A sentiment classification method for Chinese microblog is presented. For short sentence microblog, it is very challenging because the information of emotion is very limited. First, an emotion lexicon is built from training corpus. A simple measure – difference ratio is used to choose words from lexicon as features for classification. Support vector machine and voting on counts and accumulated difference ratio are jointly combined as classification method. The experimental results show that our recognition rate is better than the popular method using collocation strength. Our recognition improvement is about 2.06% in testing. Therefore, the difference ratio measure we used and the tactic in constructing the lexicon are proved very effective.**

Keywords- sentiment classification; emotion; support vector machine; microblog; emoticon

I. INTRODUCTION

We can analyze one's emotional state, like happiness, anger, and sadness, by observing a person's body reaction at the time. Imagine applying it in data processing of computer. By means of detecting the characteristics of the computer data, we can foretell the emotional state the information conveyed. Music, video, text are the three computer information data that are often analyzed. Among them, text sentiment analysis has lots of applications. Through the analysis of internet users' comments about news, books, or products, we could know their feelings and do the proper reaction or promotion. The major task, sentiment classification, is to identify users' opinions or emotions as positive or negative. The applications could be very useful. For example, Tao [1] did a sentiment analysis of the news comments, while Wu [2] try to figure out the emotion of the public shareholders by an Internet-forum-based stock market analysis method.

Microblog is a popular broadcast medium. Different from a traditional blog, its content is typically smaller such as short sentences, individual images, or video links. Users can post their texts or comments and use emoticons to express their feelings. The emoticons they used can be considered as a sentiment tag and make microblog a convenient and easy-to-get corpus for short sentence sentiment classification studies. Because, in sentiment classification studies, we generally need sentiment annotated corpus for model training (if we use supervised pattern recognition methods) and for testing the system recognition rate, using the emoticons instead of annotating sentiment corpus by human is very convenient. Therefore, by using emoticons , it can save lots of time of collecting and annotating corpus that are generally very labor intensive. Since the material is short sentence, it is very challenging to recognize its sentiment by utilizing so limited information. However, emoticons are only used in corpus annotating; the objective of our study is to classify emotion based on sentences without emoticons.

Sentiment classification normally needs the help of an emotion lexicon or dictionary that brings together positive and negative words, which can be used as the features in sentiment classification. A good lexicon can help improve the recognition rate a lot. However, one major obstacle to sentiment classification is the lack of a complete sentiment dictionary for many languages. How to overcome the obstacle is very important. Wu [3] combines multi-dictionary and commonsense knowledgebase by gathering nine kinds of sentiment dictionaries as sentiment concept seed, then through sentiment spreading activation from common sense network (ConceptNet) to get more sentiment concepts. On the other hand, Yang [4] built an emotion lexicon automatically from weblog corpora. Considering the difficulty of collecting dictionaries and domain relevance, we will construct an emotional lexicon automatically from corpus.

Further, machine learning or pattern recognition methods are commonly used in sentiment classification. For example, fuzzy association rules is proposed to analyze melancholiac patients' emotion from their daily text messages [5]. Support vector machine (SVM) is also a popular machine learning method and is adopted in emotion classification of weblog [4][6]. Since SVM has accomplished very promising performance in many pattern recognition fields, we also take on this model combined with a voting method.

This paper is organized as follows. The next section will introduce the source of our corpus. Then, the process of establishing the emotion lexicon is presented in the third section. The fourth section describes our sentiment classification method. Experimental results will be shown in the following section. Finally, conclusions and future works are discussed.

II. CORPUS

Collecting and annotating an emotion corpus is very labor-intensive. How to build a source with large amounts of

Figure 1. Plurk.

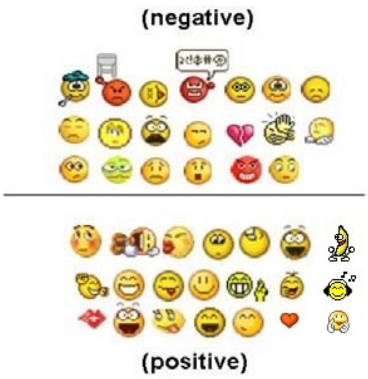


Figure 2. Positive and negative emoticons.

了不起的概念啊。
(Great concept.)
(a)

現在是怎麼了，這種犯人這樣就可以放過了。
(Now what, let such prisoner get away with this.)
(b)

Figure 3. Examples of (a) positive and (b) negative sentences.

diverse text and annotated sentiment is a very tough job. Fortunately, internet users have developed visual cues,generally known as smileys or emoticons, to express their emotion. The text on Internet with emoticons becomes a convenient and useful source for building the emotion corpus [7].

We, therefore, select Plurk [8] as the source. Plurk is a net for microblog, as shown in Figure 1, where users can publish their text, which is limited to 210 Chinese characters or English alphabets. Short sentences from political and news source with emoticons are collected.

Then, we need to determine our emotion model. Plutchik's [9] 2D and 3D emotion model is very popular. Thayer [10] also developed a two-dimensional emotion space with four quadrants. One axis is valence from negative to positive. The other axis is arousal from silent to energetic. In this paper, for simplification and consideration of corpus size, we focus on only positive and negative sentiment.

41 emoticons, including 21 positive emoticons and 20 negative emoticons, as shown in Figure 2, are adopted. Sentences with positive emoticons are annotated as positive and sentences with negative emoticons are annotated as negative. Totally, there are 8520 sentences in our training emotion corpus, comprising 4578 positive sentences and 3942 negative sentences. Testing corpus contains 4229 sentences, involving 2470 positive sentences and 1759 negative sentences. Examples of positive and negative sentences are shown in Figure 3.

III. EMOTION LEXICON

An emotion lexicon is automatically built for sentiment classification from training corpus. Word segmentation is firstly applied. Since two-syllable, three-syllable, and four-syllable words are the most frequently appeared, only the three types of words are collected.

Then, the numbers of words are counted by using a word counting program [11] to get the word count that appeared in the positive and negative sentences, namely N_p and N_n. If one word appears in positive and negative sentences just as frequently, we assume this word does not have discrimination in emotion and discard it. Such method of setting a threshold to discard words without discrimination can reduce wrong classification and increase system performance [12]. However, word application is always context dependent. Thus, the word applications in themselves are highly revealing of the context in which they are being used. So, in our future studies, we will reconsider those words discarded and divide them into further categories - like incidental constructs (not really possessing any significance), and context-revealing constructs (possibly adjectives which reveal the positive-negative weighting purely through application context). The remaining words will be classified according to their word counts in positive and negative sentences, that is, they will be classified into positive emotion lexicon if they appear more frequently in positive sentences, and be classified into negative emotion lexicon if they appear more frequently in negative sentences. That is,

$$If \begin{cases} N_p = N_n & Discard \\ N_p > N_n & Positive \\ N_p < N_n & Negative \end{cases} . \qquad (1)$$

The final step of building an emotion lexicon is word sorting according to their significance. There are several measures to determine the word significance. Pointwise mutual information (PMI), which measures the strength of the association of two samples, is widely used in relation extraction, word collocation, and word sense disambiguation in natural language processing. A variation of pointwise mutual information, which measures the collocation strength $co(e,w)$ between an emotion e and a word w, is defined as [4] :

$$co(e,w) = c(e,w) \times \log \frac{P(e,w)}{P(e)P(w)}. \qquad (2)$$

c(e,w) is total co-occurrences of *e* and *w*. Normalized collocation strength is also proposed and used in [12].

Instead, in our experiment, we use a very simple measure – difference ratio (DR), which is defined as

$$\frac{(N_p - N_n)^2}{N}, \qquad (3)$$

where N is the sum of N_p and N_n. When the total count N is the same and the difference of N_p and N_n is higher, the difference ratio is larger and the word is more significant. If the word appears in positive and negative sentences nearly equally, the difference ratio is low and the word is insignificant. If the difference of N_p and N_n is the same, the word appears less frequently got larger difference ratio and is more significant.

Words are sorted according to difference ratio and we got an emotion lexicon arranged with significance from high to low. The lexicon contains a total of 10476 words, including 5954 positive words and 4522 negative words. The procedure of building an emotion lexicon is presented in Figure 4.

From the emotion lexicon generated, we found out that there are some words unexpected, like people's name, or place's name. They are originally neutral, but classified into positive or negative emotion by learning. From Yu's study [13], we know that emotion lexicon generally includes two kind of words. One is *domain dependent word* and the other is *domain independent word*. *Domain independent word* is word with emotion linguistically and *domain dependent word* is from learning in accordance with the training field. Therefore, those people's names, or place's names are *domain dependent words* from learning. Since our corpus contains political and news contents, it causes politicians' and place's names appear in our emotion lexicon as emotion words.

IV. Sentiment Classification

There are two stages in our sentiment classification of a testing sentence. The first stage is using SVM to do the classification. Those sentences that SVM cannot classify will pass to the second stage using voting method.

In SVM, the features we used are 100 top words from emotion lexicon with highest difference ratio. If one specific word appears in sentence, the value of feature is set to 1, otherwise, it is 0.

If all the 100 words do not show in the sentence and SVM fail to classify, then we use the whole emotion lexicon to vote. If we cannot find any word of the sentence in lexicon, then the sentence is unrecognizable. Otherwise, we count the numbers of positive and negative words in lexicon appear in the sentence. The larger number determines the sentiment class. That is, if the number of positive words in sentence is greater than the number of negative words, then the sentence is set to positive, and vice versa. If the votes are the same, we accumulate the difference ratio of positive and negative words in the sentence. When the sum of the

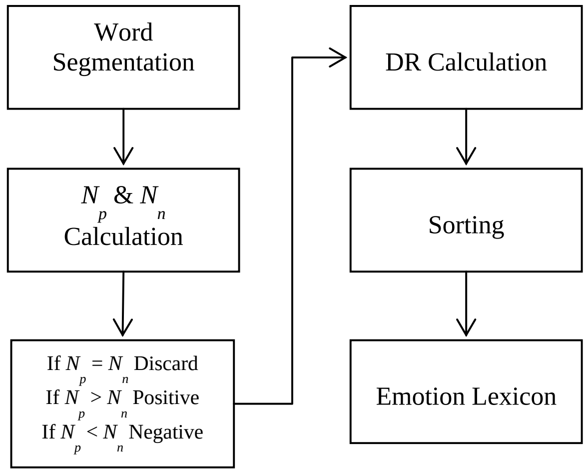


Figure 4. The procedure of constructing an emotion lexicon.

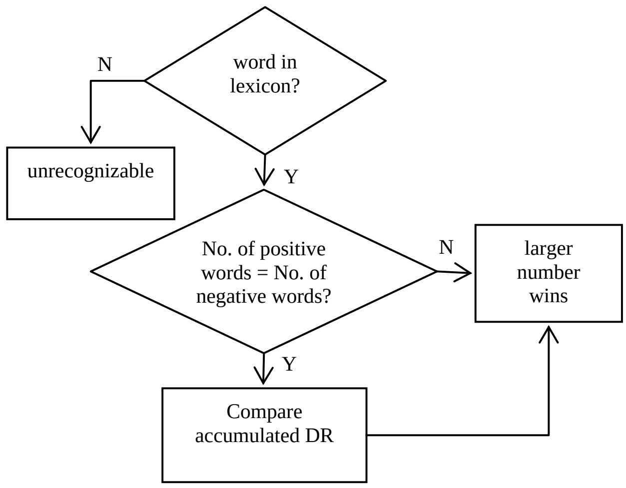


Figure 5. The voting procedure.

太過分了！一想到牠們這麼痛苦的死去，就好難過。

(a)

太|過分|了！一|想到|牠們|這麼|痛苦|的|死去，就|好|難過。

(b)

太|過分(N)(too much)|了！一|想到|牠們|這麼|痛苦(N)(pain)|的|死去，就|好|難過(N)(sad)。

(c)

Figure 6. An example of (a) a testing sentence "It's too much! As I thought They died such painfully, I am so sad." (b) Word segmentation. "|" is used as word boundaries. (c) Finding positive (P) or negative (N) words.

difference ratio of positive words is larger, then the sentence is set as positive sentiment, otherwise, it is negative. The voting procedure is shown in Figure 5. An example is shown in Figure 6. First, we have a testing sentence as in Figure 6(a). Then, we applied word segmentation to get the result as in Figure 6(b). Word boundaries are marked by using "|". Assume the SVM fails, then, we check the number of positive and negative words to vote as the procedure in Figure 5. The positive and negative words in lexicon are marked as in Figure 6 (c).

V. EXPERIMENTAL RESULTS

In the SVM stage, the software LIBSVM [14] is used. In training, 5105 sentences out of 8520 sentences are recognizable. That is, there are 3415 unrecognizable sentences that do not contain any of the 100 feature words. In the 5105 sentences, the recognition rate is 59.92%. In testing, 1476 sentences out of 4229 sentences are recognizable, and 2753 sentences are unrecognizable. In the 1476 sentences, the recognition rate is 71.68%.

Those unrecognizable sentences are sent to the second voting stage. In the second stage, 2212 sentences out of 2753 sentences are recognizable. That is, the rest 541 sentences do not contain any word in emotion lexicon. In the 2212 sentences, the recognition rate is 62.79%. Among them, 1871 sentences are decided by vote counts directly, 341 sentences are decided by further comparing the accumulated difference ratio. In the vote counts part, the recognition rate is 64.03%, and in the accumulated difference ratio part, the recognition rate is 56.01%.

Totally, in testing, 3688 sentences out of 4229 sentences are recognizable. In the 3688 sentences, 2447 sentences can get correct sentiment answer and the recognition rate is 66.35%.

By observing the recognition results, most of the expression of the sentences is appropriate. But some are not. The reasons are, firstly, the emoticons annotated are not always very accurate, and, secondly, some blogger would say things with irony (use positive words to express negative feeling), which makes the system hard to recognize.

Because there are few studies focus on short sentence sentiment classification, we will compare our method with Yang's method [4], which is a very effective emotion classification model applied in weblog articles. We will simulate Yang's method - cooperating SVM and Yang's method 3 on our short sentence corpus from microblog and classify the sentiment into positive and negative, because his method 3 outperforms his method 1 and 2 in most cases. Yang's method is applied in another context, so, to be fair, we use Yang's lexicon creation method to create lexicon from our corpus in simulating Yang's method.

First, Yang uses collocation strength to build an emotion lexicon. All collocations (word-emotion pairs) are listed in a descending order of collocation strength. For a specific word, if the collocation strength of positive emotion is larger, then the word is classified as positive emotion and vice versa. The lexicon by using Yang's method from our training corpus is 11316 words including 5892 positive words and 5424 negative words. Since our method discard the words which appear in positive and negative sentences just as frequently, our lexicon is smaller.

TABLE I. COMPARISON OF RECOGNITION RATE AND THE NUMBER OF RECOGNIZABLE SENTENCES IN BRACKETS OF OUR AND YANG'S METHOD.

Method	SVM		Voting		Total
	training	testing	counts	DR/co	
Ours	59.92% (5105)	71.68% (1476)	64.03% (1871)	56.01% (341)	66.35% (3688)
Yang's	62.63% (5336)	64.26% (1950)	62.92% (1502)	54.36% (298)	62.93% (3750)

To be compared with our method, top 100 words in the lexicon by Yang's method are used as features in SVM. The major difference of Yang's sentiment classification procedure and ours is they use accumulated collocation strength to decide the sentiment when the positive and negative votes are the same.

In the simulation of Yang's method, 5336 sentences out of 8520 sentences are recognizable, and the recognition rate is 62.63% in SVM training. In SVM testing, 1950 sentences out of 4229 sentences are recognizable. In the 1950 sentences, the recognition rate is 64.26%.

In the second voting stage, 1800 sentences are recognizable. In the 1800 sentences, the recognition rate is 61.5%. Among them, 1502 sentences are decided by vote counts directly, 298 sentences are decided by further comparing the accumulated collocation strength. In the vote counts part, the recognition rate is 62.92%, and in the accumulated collocation strength part, the recognition rate is 54.36%.

Totally, in testing by Yang's method, 3750 sentences out of 4229 sentences are recognizable. In the 3750 sentences, 2360 sentences can get correct sentiment answer and the recognition rate is 62.93%.

To make it clearer, all the experimental results are summarized in Table 1. Comparing our method with Yang's method, since we discard some words that we think do not have discrimination in emotion, which led to smaller lexicon, and causes the number of sentences that we can process is fewer. But our recognition rate is better than Yang's method using collocation strength. In other words, we both use 8529 sentences for training, and in 4229 testing sentences, our method can hit correct sentiment in 2447 sentences, 66.35% of 3688 sentences, comparing with 2360 sentences, 62.93% of 3750 sentences, by using Yang's method. The difference is 87 sentences. Our recognition improvement is about 2.06% of 4229 testing sentences.

From the results compared to Yang's method, we can conclude that discarding words which appear in positive and negative sentences just as frequently and adopting a measure considering the difference between N_p and N_n and the total word frequency N can help improve the recognition rate.

VI. Conclusions and Future Works

This paper presents a sentiment classification method for Chinese microblog Plurk. It is a very challenging job because the material is short sentence and the information of emotion is very limited.

A very simple measure, i.e., difference ratio, is chosen for lexicon building and feature selection. SVM and voting are combined as classification method. The experimental results show that our recognition rate is better than Yang's method using collocation strength. In using the same training and testing sentences, our method can get correct sentiment in 2447 sentences, while Yang's method can only get correct sentiment in 2360 sentences. The difference is 87 sentences. Our recognition improvement is about 2.06% of 4229 testing sentences. Therefore, the difference ratio measure we used and the tactic in constructing the lexicon are proved very effective.

Our method can help intelligent HCI (Human-Computer Interaction) systems sense, i.e., detect and interpret the user's emotional states automatically, and then respond appropriately, which will make the HCI systems more natural, efficacious, persuasive and trustworthy. Furthermore, the emotional data may be used in many applications, including music or consumer products recommendation, training plan or education improvement, and even health care.

However, the size of our corpus and emotion lexicon is still not enough and needs to be extended in the future. Besides, words collocations with strong sentiment orientation are important for the text sentiment analysis [15]. Semantic models can help to identify the sentiment orientations of collocations and improve our sentiment classification in the future.

References

[1] F. M. Tao, J. Gao, T. J. Wang, and K. Zhou, "Topic Oriented Sentimental Feature Selection Method for News Comments," Journal of Chinese Information Processing, Vol. 24, No. 3, May, 2010, pp. 37-43.

[2] J. Wu, Y. X. Chen, and D. M. Liu, "Internet-forum-based Stock Market Analysis Method," Computer Engineering, Vol. 38, No. 13, July 2012, pp. 254-256.

[3] H. H. Wu, "Sentiment Analysis Using Multi-dictionary and Commonsense Knowledgebase," Master Thesis, Department of Computer Science and Information Engineering, National Taiwan University, Taiwan, July 2011.

[4] C. H. Yang, K. H. Y. Lin, and H. H. Chen, "Building Emotion Lexicon from Weblog Corpora" Proceedings of 45th Annual Meeting of Association for Computational Linguistics, poster, June 23rd-30th, 2007, Prague, Czech Republic, pp. 133-136.

[5] H. W. Chiu, "Using Fuzzy FP-tree Model to Discover Associations Among Melancholiac Patient's Daily Text Messages," Master Thesis, Department of Electrical Engineering, National Taipei University of Technology, Taiwan, July 2010.

[6] C. H. Yang and H. H. Chen, "A Study of Emotion Classification Using Blog Articles," Proceedings of the 18nd Conference on Computational Linguistics and Speech Processing (ROCLING 2006), 2006, pp. 253-269.

[7] J. Read, "Using Emotions to Reduce Dependency in Machine Learning Techniques for Sentiment Classification," Proceedings of the ACL Student Research Workshop, 2005, pp. 43-48.

[8] Plurk, http://www.plurk.com/ [retrieved: January, 2015].

[9] R. Plutchik, The Emotions, University Press of America, 1991.

[10] R. E. Thayer, The Biopsychology of Mood and Arousal, Oxford University Press, 1989.

[11] Chinese Word Segmentation System, Academia Sinica, Taiwan, http://ckipsvr.iis.sinica.edu.tw/ [retrieved: January, 2015].

[12] Y. T. Sun, C. L. Chen, C. C. Liu, C. L. Liu, and V. W. Soo, "Sentiment Classification of Short Chinese Sentences," Proceedings of the 22nd Conference on Computational Linguistics and Speech Processing (ROCLING 2010), Puli, Nantou, Taiwan, September 2010, pp. 184-198.

[13] H. C. Yu, K. T. H. Huang, and H. H. Chen, "Domain Dependent Word Polarity Analysis for Sentiment Classification," Computational Linguistics and Chinese Language Processing, Vol. 17, No. 4, December 2012, pp. 33-48.

[14] C. C. Chang and C. J. Lin, "LIBSVM -- A Library for Support Vector Machines," http://www.csie.ntu.edu.tw/~cjlin/libsvm/ [retrieved: January, 2015].

[15] S. Wang and A. Yang, "A Method of Collocation Orientation Identification Based on Hybrid Language Information," Journal of Chinese Information Processing, Vol. 24, No. 3, May, 2010, pp. 69-74.

 ISBN: 978-1-61208-382-7

Two Dimensional Shapes for Emotional Interfaces: Assessing the Influence of Angles, Curvature, Symmetry and Movement

Daniel Pacheco
Laboratory of Synthetic Perceptive, Emotive and Cognitive Systems. (SPECS) Universitat Pompeu Fabra. Roc Boronat, 138. 08018 Barcelona, Spain.
Email: dapachec@gmail.com

Sylvain Le Groux
Department of Psychology Stanford University 450 Serra Mall Stanford, CA 94305

Paul F.M.J. Verschure
Laboratory of Synthetic Perceptive, Emotive and Cognitive Systems. (SPECS) Universitat Pompeu Fabra. Institució Catalana de Recerca i Estudis Avançats (ICREA) Barcelona, Spain

***Abstract*—Recent investigations aiming to identify which are the most influential parameters of graphical representations on human emotion have presented mixed results. In this study, we manipulated four emotionally relevant geometric and kinematic characteristics of non symbolic bidimensional shapes and animations, and evaluated their specific influence in the affective state of human observers. The controlled modification of basic geometric and cinematic features of such shapes (i.e., angles, curvature, symmetry and motion) led to the generation of a variety of forms and animations that elicited significantly different self-reported affective states in the axes of valence and arousal. Curved shapes evoked more positive and less arousing emotional states than edgy shapes, while figures translating slowly were perceived as less arousing and more positive than those translating fast. In addition, we found significant interactions between angles and curvature factors both in the valence and the arousal scales. Our results constitute a direct proof of the efficacy of abstract, non-symbolic shapes and animations to evoke emotion in a parameterized way, and can be generalized for the development of real-time, emotionally aware user interfaces.**

Keywords–Affective Computing; Emotional interfaces; Graphical User Interfaces; Emotional Design; Expressive Interfaces.

I. INTRODUCTION

In the recent years, several efforts have been made in the field of Human Computer Interaction (HCI) to design and implement computer systems that can recognize and express emotion. A number of models have been developed to interpret physiological measurements [1], or behavioural records of body and facial expression in real time [2], and today, reliable ways of monitoring users emotions are being incorporated in commercial systems, such as Microsoft Kinect. On the other hand, models for the expression of emotion — which are usually based on psychological research — have been shown to coherently convey emotion to humans by manipulating human-like or anthropomorphic emotional stimuli. Specifically in the field of Computer Graphics, models for the synthetic expression of emotion using Computer Generated Imagery (CGI) have traditionally involved the use of so called avatars, i.e., virtual characters that simulate human facial expression [3][4], or body movement [5][6], to express a particular emotion explicitly (for a review see [7]).

However, it is well known that humans not only respond to human-like or symbolic emotional stimuli. In the literature on music and emotion, for instance, it has been shown that musical parameters such as tempo, pitch or tonality may profoundly affect a person's affective state [8][9][10]. Since most of the interactive systems with which we interact today present bidimensional user interfaces that are non-symbolic / non-anthropomorphic, it is relevant to identify what are the most important graphical parameters of emotion in simple forms and animations that can be used to generate such interfaces. Furthermore, the advent of new communication technologies allows today for the real time generation of highly parameterized CGI which offers great possibilities for the design and implementation of emotionally aware user interfaces. How can we identify the geometrical and cinematic properties of 2D shapes and animations that have more impact on emotion, and make use of this knowledge in the design of affective HCI systems?

Based on literature presented in Section II, we defined an experimental setup to investigate this question by assessing the influence of four specific graphical parameters of shapes and animations on emotion: angles, curvature, symmetry and speed of movement (Section III). Our results show that it is possible to experimentally induce emotional states in controlled environments by parametrically tampering these geometric and kinematic characteristics. The specific impact of each one of them is discussed in Section IV. In Section V, we present our conclusions and future work.

II. SHAPE, MOVEMENT AND EMOTION

A. Shape

The relationship between non symbolic graphic features and emotion has been studied from different perspectives. One approach mostly adopted in the Image Retrieval field has been to study global image features and assess their effectiveness in conveying specific emotions. Several parameters of color (i.e., hue, brightness or saturation [11]), and textures (i.e., coarseness, contrast and directionality [12]), have been shown to influence the affective states of human observers. Similar parameters have been identified in saliency-based visual models (i.e., colour, intensity, orientation and symmetry [13][14]) which can predict human fixations, although, to our knowledge, the relationship between saliency and emotion models has not been studied.

A problem with the study of complex images and global features is the great amount of dimensions in which such stimuli can be parameterized. Real images usually involve semantic components that can be highly influential on human emotion. Traditional databases for emotion induction, such as the International Affective Picture System (IAPS) [15] do not differentiate between symbolic and non symbolic emotion determinants, which makes them not suitable to assess the specific role of these two key elements. Some efforts, however, have been made to achieve this goal [16].

A different approach has been taken in the field of visual perception, where the synthetic generation of visual stimuli has been adopted in early studies. The seminal work of Attneave with bidimensional abstract shapes already showed that the parametric variation in basic geometric characteristics of such figures — number of turns in the contour, symmetry and angles — evoked completely different subjective judgements about their “complexity" [17][18].

A complexity scale was also used to rate shapes in [19] and [20], although the concept was defined in different ways — reproduction performance for the former and difficulty in providing a verbal description of an image for the latter. The variables that were considered more influential on the perceived complexity of a shape in both studies were orientation, repetitiveness and variance in interior angles.

Can such parameters be influential on emotion? The literature on the topic is sparse. Some studies have shown that curved shapes are better in portraying emotion than shapes composed by straight lines [21], and that features such as angles ratio, curvature and symmetry [22][23], can predict the emotion induced by specific 3D shapes. Other studies have proposed that the perceived emotion of abstract figures is determined by internal dynamics such as the subjectively judged “instability" of a figure. The intensity of emotions that can be ascribed to the figures is correlated with their perceived instability, which is defined by the figure orientation with respect to a predefined ground [24].

B. Movement

The relevance of movement in the expression of emotion has been highlighted in several studies, most of which have focused in human — or anthropomorphic — body motion. It has been argued that exaggerated corporal motion enables the recognition of the intensity of the affective states that can be attributed to body postures, and that parameters that define a specific body configuration can be correlated with the emotion attributed to it [6]. Moreover, it has been shown that speed and spatial amplitude play a fundamental role in emotional perception of human-like body movement [25]. Several studies on human body motion using point lights show that perceivers are able to infer emotions reliably and easily basing their judgments solely on the dynamic patterns of actions [26][27].

Abstract motion has also been studied, but in the same way that abstract figures have captured less attention than figurative graphical representations, the influence of non-articulated motion on human perception is less understood than its symbolic counterpart. Early studies already showed that emotional descriptions can be attributed to abstract figures that move in a non articulated, non anthropomorphic way [28][29]. In the same line, it has been suggested that the perception of animacy of a shape can be predicted by the magnitude of the speed change and the change of direction measured in angles [30][31] among other parameters. What are the features of abstract movement that most influence emotion? Density (animated notches of the contour), strength (amplitude in the deformation or translation), and speed were described as the most determinant factors of the emotion elicited by abstract shapes [25]. On the other hand, speed in the motion of abstract patterns has been shown to be highly influential on human emotion and behaviour in a mixed reality setup [32].

III. Methods

A. Description

The methodology and overall design of our experiment was based on a previous study conducted in our laboratory [33].

Different visual samples varying among two visual determinants of emotion (shape and movement) were presented to the participants. The morphological features that were considered — lines/curves ratio, acute/obtuse angles ratio, and symmetry — were extracted from [22] and [23], and adapted to our specific setup. Such parameters were the only geometrical features of shapes studied with respect to emotion that we found in the literature. Each one of these geometrical parameters was tested at three different levels of movement: low, medium and high (L/M/H). In total, 81 animations were rendered (i.e., 27 shapes at 3 different levels of movement each). The parameters that were manipulated are described in the following paragraphs.

1) Lines/Curves ratio: The Lines/Curves ratio (LCR) was calculated according to the following formula:

$$\text{LCR} = \frac{\text{Lines}}{\text{Curves} + \text{Lines}} \tag{1}$$

A low LCR was considered under $\frac{1}{3}$, medium between $\frac{1}{3}$ and $\frac{2}{3}$ and high between $\frac{2}{3}$ and 1.

2) Acute/Obtuse angles ratio: The Acute/Obtuse angles ratio (AOAR) was calculated according to the following formula:

$$\text{AOAR} = \frac{\text{Acute angles}}{\text{Acute angles} + \text{Obtuse angles}} \tag{2}$$

A low AOAR was considered under $\frac{1}{3}$, a medium AOAR between $\frac{1}{3}$ and $\frac{2}{3}$ and a high AOAR between $\frac{2}{3}$ and 1. AOAR was also considered for curved shapes by counting the angles formed by correspondent tangents of adjacent curves. Reflex angles (between 180 and 360 degrees), were not considered in the equation.

3) Simmetry: Symmetry (SYM) was considered in 2 axes. A symmetrical figure among two axes was given a high level of symmetry; a symmetrical figure among one axe was considered at a medium level of symmetry, and a non symmetric figure was given a low level of symmetry.

4) Movement: All shapes were rendered in three different levels of movement (MOV): low, medium, high. In the low level, the image of the shape was rendered statically. In the medium and high level, movement was produced by translating the shape in pseudo-random directions, determined by a perlin noise algorithm. In the medium level, the range of translation was set to 5 VR units and the speed of translation of the algorithm set to 1. In the high level of movement, the range of movement was also 5, but the speed of translation was 5 times faster than the medium level.

5) Examples: A circle has a high symmetry, low lines curves ratio, and low acute obtuse angles ratio. An example of a figure that has high symmetry, high LCR and high AOAR is a star (for more examples see Figure 1).

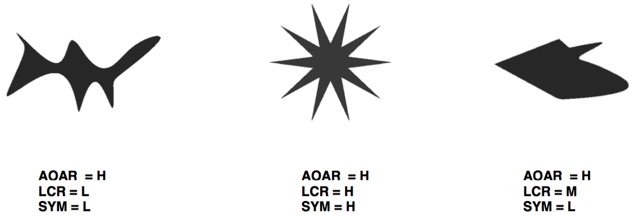


Figure 1. Abstract shapes used in our experiment. Left: high AOAR, low SYM and low LCR. Center: high AOAR, high LCR, high SYM. Right: high AOAR, medium LCR and low SYM.

B. Experimental Procedure

Our experimental design included four independent variables (LCR, AOAR, SYM, MOV) with three levels each (Low, Medium, High), and two dependent variables (arousal, valence) — the participants ratings in a nine points self-assessment scale (the self-assessment manikin [34], based in Rusell's circumplex model of emotions [35]).

Stimuli were presented in a randomized order. After eight seconds of exposure, participants were asked to rate the shape in the SAM scale while the stimulus remained visible. Precise instructions were given to them in order to assess their emotional states as it was at the moment of exposure. After the self-assessment has been achieved, an in-between period of eight seconds with no visual stimulation (black screen) preceded the next trial. Exposure time was made two seconds longer than the original self assessment study conducted with images [34] to allow a good exposure to moving images (animations). An application was developed for the generation and rendering of the stimuli and the online recording of the participant's responses using the Unity3D Game Engine.

C. Participants

A total of 12 university students (5 women, MAge = 28.333, range = 22-39) participated in the experiment. All of them had normal or corrected-to-normal vision.

IV. RESULTS

A. Graphic representations

We carried a Two-Way Repeated Measure Multivariate Analysis of Variance (MANOVA), followed by univariate analysis and post hocs. Kolmogorov-Smirnov and Shapiro Wilk tests showed that the data was not normally distributed in valence and arousal scales; therefore, we run Kruskal-Wallis tests in the univariate analysis to verify the results.

1) Valence: In the valence scale, shapes composed mostly by curves were perceived significantly more positively than shapes composed by similar numbers of lines and curves. Shapes composed mostly by curves were also perceived more positively than shapes composed mostly by lines, although this difference was not significant. The analysis showed a significant multivariate effect for LCR $F(2, 9)$ ($p < 0.05$). Follow-up univariate analysis revealed an effect of LCR on valence $F(1.322,26,717) = 6.882$ ($p < 0.05$). Mauchly tests indicated that the assumption of sphericity was not met. Hence we corrected the F-ratios with Greenhouse-Geisser. A post-hoc pairwise comparison with Bonferroni correction showed a significant mean difference of 0.4527 between Low and Medium LCR on the valence scale (Figure 2).

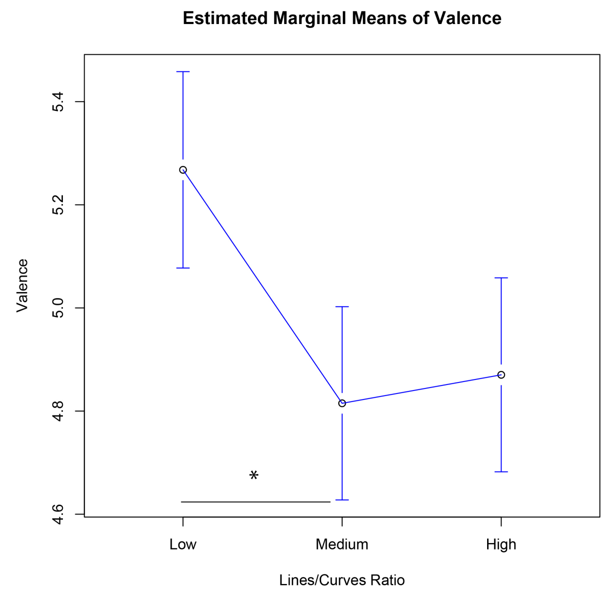


Figure 2. A figure composed mostly with curves is perceived significantly more positive than a figure composed by curves and lines.

Besides, a significant interaction was found between AOAR and LCR for valence $F(4, 40) = 6,444$ ($p < 0.05$), suggesting than shapes composed mostly by acute angles — e.g., a star — are perceived more positive when they are also composed by straight lines. The analysis on valence also revealed a significant interaction effect in the multivariate analysis for the AOAR*SYM interaction $F(4,7)$ ($p < 0.05$), and for the LCR*SYM interaction $F(4, 7)$ ($p < 0.05$). When a shape is symmetrical in two axes, the presence of acute angles is perceived more positive. On the other hand, when the figure is asymmetrical, it will be perceived as more positive if it is composed mostly by curves.

2) Arousal: In the arousal Scale, the only significant multivariate effect was found for the interaction between LCR and AOAR $F(4, 7)$ ($p < 0.05$), and univariate analysis confirmed this result $F(4, 40) = 4.694$ ($p < 0.05$). When a shape has High AOAR, it tends to be perceived as more arousing if it also has a high LCR.

B. Movement

1) Valence: No significant multivariate effect was found for movement on valence. However, influence of movement on valence was reported by univariate analysis $F(1.633,143.265) = 5.969$ ($p < 0.05$) (Figure 3).

A post-hoc pairwise comparison with Bonferroni correction showed a significant mean difference of -1.1023 between Medium and High. Therefore, figures that moved slowly were perceived as significantly more positive that figures translating fast.

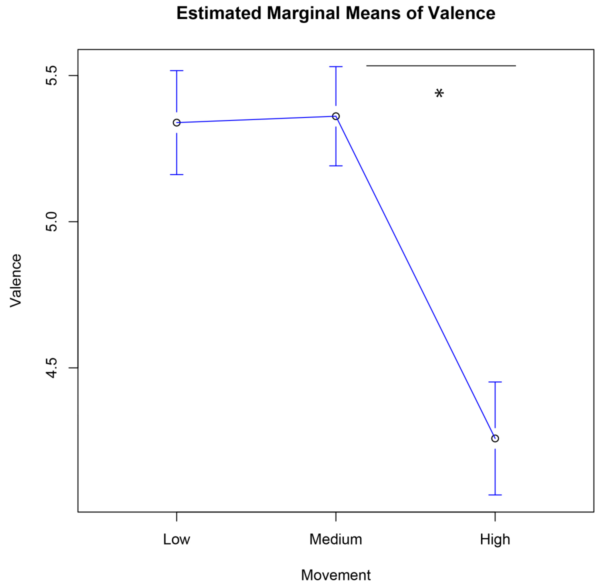


Figure 3. Different kinds of movement elicited different reports in the valence scale. Fast moving shapes were considered significantly less positive than shapes that moved slow and non-moving shapes.

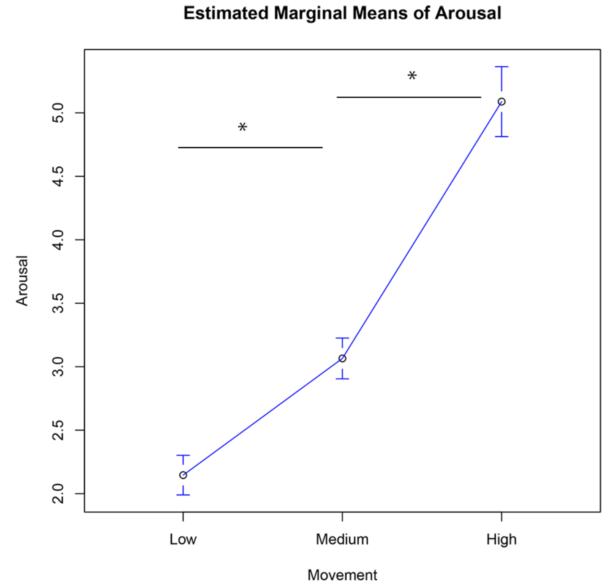


Figure 4. Movement significanlty affected the participants reports in the arousal scale. Fast moving shapes were considered significantly more arousing than shapes that moved slow and non-moving shapes.

2) Arousal: Movement was found to be highly influential in on arousal. Moving shapes were perceived significantly more arousing than non-moving shapes, and fast-moving figures were reported to be significantly more arousing than slowly-moving figures (Figure 4). The analysis showed a significant multivariate effect for movement F(2, 9) ($p < 0.05$) on arousal. Mauchly tests indicated that the assumption of sphericity was met. Therefore, we did not correct the F-ratios for the following ANOVAS. Follow-up univariate analysis revealed an effect of movement F(1.112, 1232.994) = 20.715 ($p < 0.05$) on arousal. A post-hoc pairwise comparison with Bonferroni correction showed a significant mean difference of -0.9195 between Low and Medium movement, of -2.0231 between Medium and High movement, and of -2.9427 between Low and High movement (Figure 4). In our design, Low was defined as absence of movement, and Medium and High were defined as different levels of speed in random translation.

V. Conclusions

Abstract shapes and animations varying among four emotionally relevant graphical parameters (i.e., proportion of lines and curves, proportion of acute and obtuse angles, symmetry and movement), were presented to the participants of our experiment and the correspondent emotional responses were recorded using self-reports based in the circumplex model of affect. Our results show that the manipulation of such low level graphical parameters evoked different affective states in our participants, and that some of them (i.e., LCR and movement) had a specific influence in the valence and arousal scales. In some cases, the interaction between parameters was significant (e.g., for curvature and angles in both the valence and arousal scales). Our results are a contribution to the understanding of the role that basic geometric characteristics of abstract, bidimensional shapes play on human emotion, and may be useful for developers and designers wishing to develop and implement emotional user interfaces.

We speculate that the manipulation of these parameters in real time and in less controlled experimental conditions will coherently influence the affective states of humans observers in a similar manner than observed in our experiment, and plan to reformulate our experimental design in order to include the real time generation of the stimuli from the identified parameters. We also plan to include more objective measurements of emotion such physiological records — i.e., Electrodermal Activity, Hear Rate, Respiration —, as a complement to the self-assessment responses. Such physiological data could be used in a second stage as an input to the system, allowing for the controlled generation of the stimuli depending on the users emotional states, and the development of more sophisticated, emotionally aware user interfaces.

VI. Acknowledgements

This research received funding from the European Community's Seventh Framework Programme (FP7-ICT-2009-5) under grant agreement n. 258749 [CEEDS].

References

[1] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 31, no. 1, 2009, pp. 39–58.

[2] M. Thrasher, M. D. Van der Zwaag, N. Bianchi-Berthouze, and J. H. Westerink, "Mood recognition based on upper body posture and movement features," in Affective Computing and Intelligent Interaction. Springer, 2011, pp. 377–386.

[3] P. Weyers, A. Mühlberger, C. Hefele, and P. Pauli, "Electromyographic responses to static and dynamic avatar emotional facial expressions," Psychophysiology, vol. 43, no. 5, 2006, pp. 450–453.

[4] A. Tinwell, M. Grimshaw, D. A. Nabi, and A. Williams, "Facial expression of emotion and perception of the uncanny valley in virtual characters," Computers in Human Behavior, vol. 27, no. 2, 2011, pp. 741–749.

[5] M. Fabri, D. J. Moore, and D. J. Hobbs, "The emotional avatar: non-verbal communication between inhabitants of collaborative virtual environments," in Gesture-Based Communication in Human-Computer Interaction. Springer, 1999, pp. 269–273.

[6] M. Inderbitzin, A. Valjamae, J. M. B. Calvo, P. F. Verschure, and U. Bernardet, "Expression of emotional states during locomotion based on canonical parameters," in Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on. IEEE, 2011, pp. 809–814.

[7] V. Vinayagamoorthy, M. Gillies, A. Steed, E. Tanguy, X. Pan, C. Loscos, M. Slater et al., "Building expression into virtual characters," 2006.

[8] L.-L. Balkwill and W. F. Thompson, "A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues," Music perception, 1999, pp. 43–64.

[9] P. N. Juslin and J. A. Sloboda, Music and emotion: Theory and research. Oxford University Press, 2001.

[10] A. Goldstein, "Thrills in response to music and other stimuli." Physiological Psychology, 1980.

[11] P. Valdez and A. Mehrabian, "Effects of color on emotions." Journal of Experimental Psychology: General, vol. 123, no. 4, 1994, p. 394.

[12] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," Systems, Man and Cybernetics, IEEE Transactions on, vol. 8, no. 6, 1978, pp. 460–473.

[13] G. Kootstra, A. Nederveen, and B. De Boer, "Paying attention to symmetry," in Proceedings of the British Machine Vision Conference (BMVC2008). The British Machine Vision Association and Society for Pattern Recognition, 2008, pp. 1115–1125.

[14] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 20, no. 11, 1998, pp. 1254–1259.

[15] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International affective picture system (iaps): Technical manual and affective ratings," 1999.

[16] N. Liu, E. Dellandréa, B. Tellez, and L. Chen, "Associating textual features with visual ones to improve affective image classification," in Affective Computing and Intelligent Interaction. Springer, 2011, pp. 195–204.

[17] F. Attneave, "Some informational aspects of visual perception." Psychological review, vol. 61, no. 3, 1954, p. 183.

[18] F. Attneave, "Physical determinants of the judged complexity of shapes." Journal of Experimental Psychology, vol. 53, no. 4, 1957, p. 221.

[19] C. M. Mavrides and D. Brown, "Discrimination and reproduction of patterns: Feature measures and constraint redundancy as predictors," Perception & Psychophysics, vol. 6, no. 5, 1969, pp. 276–280.

[20] C. Heaps and S. Handel, "Similarity and features of natural textures." Journal of Experimental Psychology: Human Perception and Performance, vol. 25, no. 2, 1999, p. 299.

[21] R. Hiraga, "Emotion recognition in polygons and curved shapes," in Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on. IEEE, 2011, pp. 3286–3291.

[22] S. Achiche and S. Ahmed, "Mapping shape geometry and emotions using fuzzy logic," in Proceedings of IDETC/CIE, 2008.

[23] S. Achiche and S. Ahmed-Kristensen, "Genetic fuzzy modeling of user perception of three-dimensional shapes," Artificial Intelligence for Engineering Design, Analysis and Manufacturing, vol. 25, no. 1, 2011, pp. 93–107.

[24] M. Pavlova and A. Sokolov, "Perceived dynamics of static images enables emotional attribution," Perception-London, vol. 34, no. 9, 2005, p. 1107.

[25] K. Amaya, A. Bruderlin, and T. Calvert, "Emotion from motion," in Graphics interface, vol. 96. Citeseer, 1996, pp. 222–229.

[26] A. P. Atkinson, W. H. Dittrich, A. J. Gemmell, A. W. Young et al., "Emotion perception from dynamic and static body expressions in point-light and full-light displays," PERCEPTION-LONDON, vol. 33, 2004, pp. 717–746.

[27] W. H. Dittrich, T. Troscianko, S. E. Lea, and D. Morgan, "Perception of emotion from dynamic point-light displays represented in dance," Perception-London, vol. 25, no. 6, 1996, pp. 727–738.

[28] F. Heider and M. Simmel, "An experimental study of apparent behavior," The American Journal of Psychology, vol. 57, no. 2, 1944, pp. 243–259.

[29] A. Michotte, "The perception of causality." 1963.

[30] P. D. Tremoulet, J. Feldman et al., "Perception of animacy from the motion of a single object," PERCEPTION-LONDON-, vol. 29, no. 8, 2000, pp. 943–952.

[31] B. J. Scholl and P. D. Tremoulet, "Perceptual causality and animacy," Trends in cognitive sciences, vol. 4, no. 8, 2000, pp. 299–309.

[32] A. Betella, M. Inderbitzin, U. Bernardet, and P. F. Verschure, "Non-anthropomorphic expression of affective states through parametrized abstract motifs," in Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on. IEEE, 2013, pp. 435–441.

[33] S. Le Groux and P. F. Verschure, "Subjective emotional responses to musical structure, expression and timbre features: A synthetic approach," 9th International Symposium on Computer Music Modelling and Retrieval (CMMR), 2012.

[34] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," Journal of behavior therapy and experimental psychiatry, vol. 25, no. 1, 1994, pp. 49–59.

[35] J. A. Russell, "A circumplex model of affect." Journal of personality and social psychology, vol. 39, no. 6, 1980, p. 1161.

You Do Not Miss Advice from Mentor during Presentation: Recognizing Vibrating Rhythms

Ali Mehmood Khan
Universität Bremen, TZi
Bremen, Germany
akhan@tzi.de

Michael Lawo
Universität Bremen, TZi
Bremen, Germany
mlawo@tzi.de

Abstract— **The haptic technology takes into consideration the human sense of touch and gives a new dimension to the way people communicate. In this work, we examined how a mentor could advise speakers to control their voice pitch and volume like speak slowly, speak fast, speak loudly and speak softly using haptic feedback during their speech. We made four vibrating rhythms and conducted a user study on different people in order to know whether they can feel the difference among four vibrating patterns or not. We used mobile's phone vibrating motor, for this purpose, which was meant to give a haptic feedback to participants. Our results show that participants were able to judge the vibrating rhythms with high accuracy.**

Keywords- Actuator, Wearable computing, Haptic feedback.

I. Introduction

Giving a speech or presentation to the audience is not easy, especially if you do not have lot of experience. People sometimes lose control of their voice and pitch; they speak fast, loudly, slowly and softly. The issue is that they do not realize it while giving a speech. It is not possible for a mentor/friend to interrupt the speaker during a presentation and advise him to control voice and pitch. Therefore, we need a system that gives the mentor the possibility to seamlessly advise the speaker during his speech.

We want to figure out how to convey advise to a speaker since he is giving a speech to the audience and it is not a good idea to interrupt him in front of all or show some body gestures which would be seen by others as well. However, rehearsals always help to make a better presentation but they do not solve this problem. Bottom line is that we need a system which can help a speaker during his speech so that he can control his voice and pitch.

Studies conducted before revealed that by using haptic vibration patterns (the so called tactons) is an an approach which can be used for communication [11]. It is not necessary to limit ourselves to understanding and perceiving single vibration patterns from or devices when there is a big possibility of understanding much more complex messages. Tactons are defined as structured tactile messages. For a wide range of distinction between them, several encoding dimensions have been studied. The research made in this area revealed that with the right combination and number of patterns, the recognition rates are not that different [11].

The way in which the information is encoded and how we distinguish the meaning of the message can prove challenging sometimes especially when we have large number of possible tactoons [12]. The limitations are given by the incapacity of distinguishing between the patterns, if they are not clear. As for parameters, it is common to use frequency, amplitude, roughness and rhythm.

Another important aspect is the spatial localization of the vibrating device. The human skin does not have the same sensibility on its entire surface to perceive information. This is due to the innervations density that varies across the average area of two square meters of skin. In order to avoid confusion of some parameters, it is necessary to consider that the skin is not able to distinguish between frequency and intensity [6]. Also, studies have shown that several areas of the body are more perceptive than others. In order to get a good perception of the pattern, the source of the vibration has to be situated close to the elbow or the spine [13]. For less sensitive parts of the body, not much research has been conducted.

In order to successfully transmit information using vibration patterns, all these aspects are needed to be taken into consideration. In this research, we want to:

- Achieve a good overall recognition rate (more than 80% of correct patterns recognition).
- Have user acceptance of the device and the idea behind.
- Figure out parameters that affect general recognition rate score
- Discover if the neck area can be an alternative for the spatial localization.

We decided on the following points regarding the number of patterns and the spatial localization. We designed four patterns using different lengths and time pauses between the vibrations (i.e., rhythm for the messages including more than one vibration). A rhythm was realized with this combination. Four messages meaning speak louder, speak quieter, speak slower, speak faster were encoded using this way. In case of an extended vocabulary, it is recommended to add at least one more variable parameter (intensity for instance). In one of the studies [7], the results showed that a combination of rhythm and intensity offers a higher recognition rate than using rhythm and roughness.

As for the location, the side part of the neck was used, since there are no studies that reveal if the neck is perceptive or not. A reduced number of patterns were designed (based

 ISBN: 978-1-61208-382-7

on rhythm). The main goal behind the study is to see how perceptive this area is, how the users accept the device and if people are able to act according to the information received, when engaged in a speaking activity.

We developed a system, using a connected actuator to mobile phone via Bluetooth/Wi-Fi, which was controlled by a mentor for advising speakers if they wanted to change their voice and pitch (as shown in Figure 1).

In the next section, the related work will be discussed. Hypothesis and research question will be discussed in Section III. Experimental methodology will be discussed in Section IV. Results and analysis will be discussed in Section V. Participants' feedback will be discussed in Section VI, and conclusion and future work will be presented in the last section.

II. RELATED WORK

Integrating vibrating sensors with mobile phones are not a new research field. A lot of research has been done in this area [1]-[5]. Similarly, researchers have already designed different vibrating patterns in order to give a haptic feedback and it was proven that people can distinguish different vibrating patterns [5]-[9].

Research conducted by Feige [8], where participants were asked to keep mobile phones in their pockets, shows results far from being satisfactory to make it possible for users to distinguish different vibration patterns. That is why a prototype in a form of wristband with embedded vibromotor was built and used. Five patterns with different rhythms were designed during an iterative process, which included pre-tests on each iteration until patterns showed to be discernible. Fourteen persons aged 19 to 46 (avg: 26.2, sd: 8.4) participated in the experiment and were randomly assigned to either the experimental or the control group which were of equal size. The difference between groups was in the environment where they received tactile patterns: control group - in a neutral room environment, experimental - in a mobile street environment where they had to walk according to route randomly selected for each participant. The results show that the environment does not makes any difference in the overall recognition score (about 93% (sd: 11.5) in the control group and even 94% (sd: 16.2) in the experimental group).

Another research conducted by Lorna M. Brown and T. Kaaresoja was devoted to investigate if a regular mobile phone can be used as the source for vibrations to represent tactile patterns [7]. Nine patterns with different roughness (from smooth to rough and very rough) and intensity (low-, mid- and high-level, controlled by vibromotor frequency) were used in experiment [9]. An important fact that needs to be mentioned is that the amplitude-modulated waveforms (used to create rough nesses) could not be reproduced on the phone motor. An approximation of roughness was created by using different speeds of on-off pulses. As for positioning, participants were holding a phone in non-dominant hand.

"A series of experiments was conducted to evaluate the effectiveness with which a tactile display mounted on either the forearm or the back can be used to communicate simple instructions and commands"[6]. The results revealed that the back is more perceptive (up to 98% correct pattern recognitions, lowest - 75%), while for the forearm, the correct pattern recognition rate was only between 30% and 96% depending on specific patterns. When participants were engaged in different activities, like physical or cognitive tasks, the results of accuracy of identification were up to 92% (for back). The amount of patterns used during these experiments varied from 7 to 8 and some of the patterns were adjusted during procedures. The big difference with many other haptic experiments and with one we have conducted is in the type of vibrotactile devices and patterns used. Here, mounted tactile displays constructed from several (9 for forearm and 16 for back) vibro motors were used, so tactile patterns were encoded not with a roughness or intensity, but with a sequence of active motors. This combination made it possible to give sophisticated navigation commands with application to military forces and with a minimum amount of tactile patterns training required. Even though the solution can be considered highly scalable and effective, it is also a sophisticated and expensive one.

Figure 1: Actuator is connected to mobile phone via Bluetooth/Wi-Fi

In our research, we want to develop a system where a mentor should be able to send advice to speakers that they need to control their voice and pitch. Speakers will be wearing an actuator which would be connected to a mentor's mobile phone via Bluetooth or Wi-Fi as shown in Figure 1.

III. HYPOTHESIS AND RESEARCH QUESTION

The vibrating rhythm generated by an actuator, worn by a speaker, controlled by a mentor which is connected to the mentor's mobile device via Bluetooh/Wi-Fi advises the speaker to change his voice and pitch during a speech.

We want to prove that it is possible to determine the user's wish regarding receiving the incoming mobile calls by using some contextual information like surrounding temperature, noise and light intensity, user location like "indoors" or "outdoors", user's physical position like "sitting", "horizontal" and "vertical" and location of the mobile device like "pocket", "bag", "table" and "hand".

IV. EXPERIMENTAL SETUP

This section describes the approach we have adopted; it introduces the experimental set-ups, including the software system for testing and the gathered data in detail.

A mobile phone Huawei U8800 Pro running Android v.2.3.5 operating system was used as the source for vibrations. The device was attached to the participant's neck using an elastic band as shown in Figure 2. For data transmission, a Bluetooth Logical link control and adaptation protocol (L2CAP) was used.

Figure 2: Phone is attached to participant's neck with elastic band.

In order to solve our research question, we implemented an application for the mobile phone which generates four types of vibrating rhythms as shown in Figure 3.

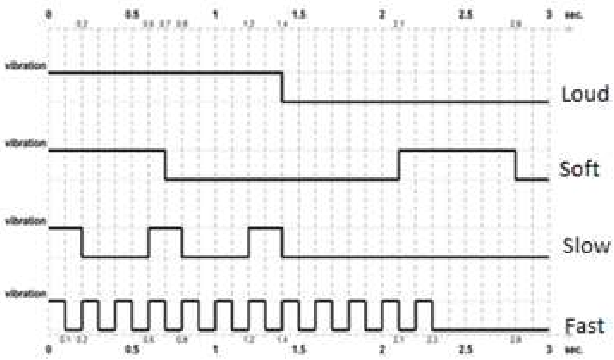


Figure 3: Different vibrating rhythms

We recruited 10 participants (7 males, 3 females), age ranged from 21 to 32 (mean: 25.2, SD: 4.11) and ranged in BM (body mass index) [10] was from 19.0 to 30.4 (mean: 22.99, SD: 3.38) because we also wanted to know whether corpulent people can feel vibrating rhythms or not.

Firstly, the purpose of the experiment was explained to participants and each participant was familiarized with the vibration patterns by holding the mobile device in his hand. After this short introduction, each participant had an unlimited amount of time to explore the four vibrating patterns in order to learn them. During training session, the device was attached to one side of the neck using the elastic band. The software user interface allowed them to select the desired pattern that was transmitted to the device as many times as he or she needed to feel it.

Later, the participants had to go through with the testing session in order to know whether they are familiar enough with the vibrating patterns. A pre-experiment was done and participants were asked to determine patterns that were sent in a random order (the default is 20 times and it can be configured). The developed software has a testing mode which shows the percentage of correct answers after completion of this part. Each participant was required to achieve a minimum of 75%; otherwise, he had to go through all processes again in order to get better trained.

Finally, participants were asked to speak about any topic for few minutes (around 7 minutes). During the speech, each participant received a total of 12 vibrating patterns (each vibrating pattern repeated 3 times). There was an interval of at least 30 seconds between two instructions. Participants were observed throughout the experiment by an observer.

V. Results and Analysis

While conducting the experiments, we were expecting two major drawbacks, which can be categorized into the following groups:

A. Confusion

When the participant recognizes the pattern, but feels unsure or unpleasant with a vibration and this affects his or her speech accordingly.

B. Interruption

When the participant's speech was interrupted for a while and then they continued.

These two are the most common reactions on vibrations and they can occur with correct pattern recognition (i.e., following the given hint), ignoring of the pattern (pattern could have been recognized but ignored or have not been recognized and ignored) or when pattern is recognized incorrectly (i.e., participant interpreted it as another pattern).

We obtained data from 10 participants (as shown in Table 1). Each participant received 12 patterns, bringing the overall total to 120 patterns; out of these, 107 patterns were interpreted correctly, i.e., 89.17% correct interpretations, taking a fact into account that 20% had two trainings, 10% had four trainings and rest 70% had three trainings before experiment.

Overall, most of the participants were able to act accordingly to a received pattern without showing any of the mentioned emotions and reactions. So, the speech was neither interrupted, nor was there any confusion or repeated speech in 80 correct interpretations, that is 74.766% of all 107 correct pattern recognitions or 66.667% of all 120 patterns being totally sent during experiment. The statistics for problems with the speech is the following: 14 speech interruptions, 11 confusions and 2 repeated speeches, all among 107 correct pattern interpretations.

TABLE I. OVERALL RECOGNITION FOR ALL PATTERNS

	Loud %	Soft %	Slow %	Fast %	Total Avg(%)
Correct interpretation/ % of correct interpretation	28/30	23/30	27/30	29/30	107/120
	93.33	76.67	90	96.67	89.17
Incorrect interpretation/ % of incorrect interpretation	0/30	4/30	1/30	0/30	5/120
	0	13.33	3.33	0	4.17
Ignored, no interpretation/ % of ignored interpretation	2/30	3/30	2/30	1/30	8/120
	6.67	10	6.67	3.33	6.66

Our results show that vibrating patterns were recognized with an accuracy of 89.17%. “Fast” pattern was recognized with an accuracy of 96.67% whereas “Soft” pattern was recognized with an accuracy of 76.67%. There were only 5 patterns out of 120 which were interpreted incorrectly, whereas 8 patterns out of 120 were ignored by participants, which means that participants were not able to feel anything.

.

VI. PARTICIPANTS’ FEEDBACK

We could infer that the neck can be a good place for usage of vibrotactile device. However, we must pay a lot of attention to the social acceptance, as well as user acceptance and usability issues, because if the device and system are helpful but rejected by public it will not have any success. This is a very important aspect in wearable computing domain. This is why after each experiment we asked participants for their opinions and feedback and here is a summary of what we were able to find out:

- Sometimes the speech was interrupted because of vibration. Since the difference between the patterns was often associated with the number of vibrations, it can be possible that the participants waited until the device stopped vibrating and counted the vibrations.
- Since mobile phone was a source of vibrations and its vibromotor made noise, one participant stated that he associated the vibrated noise with the patterns instead of vibrated rhythm (haptic feedback).
- The device was considered too big for the neck area and it created a discomfort for the participants (in order to have user acceptance, the device needs to be smaller and lighter).
- Several participants mentioned that there is a chance that others could notice that you are receiving some information or hints in some way. Also, 30% of participants stated that they can probably show better results after additional training. Finally, only one (10%) participant stated that placing a vibrating device on the neck is suitable for him and does not make a big disturbance. While most of the participants wanted to take the device off as soon as possible because it felt very unpleasant and disturbing, one of the participant totally not accepted the device and stated that he “felt like a dog” because of the elastic band and big vibrating device located on the neck.

VII. CONCLUSION AND FUTURE WORK

This paper explains how to give instructions to speakers during their speeches, this paper describes the importance of the problem, discusses the methodology, results and analysis.

Currently, we are using mobile phone's vibrating motor instead of using any separate actuator (vibrating motor). Mobile phone’s vibrating motor. We generated four different vibrating rhythms by using mobile phone's vibrating motor; those rhythms were differentiated by participants with high accuracy. Results have shown that one actuator can be enough for solving this problem. This prototype is a proof of concept and our results show that one actuator can solve this problem by using the haptic feedback. Results also show that people can use this system independent of gender, BMI (body mass index) and age group.

We also got a good feedback from our participants for this proposed system because they think that this system would be helpful to those people who want to give a good presentation. Later, we will embed an actuator either in a necklace or in a tie in order to give a haptic feedback to speakers.

ACKNOWLEDGEMENT

This research was funded by USEFIL (www.usefil.eu)

REFERENCES

[1] A. M. Khan, An empirical study on mobile phone usage. BCS-HCI 2011, Newcastle, UK . Proceedings, pp. 120-125

[2] D. Siewiorek, et al, A context-aware mobile phone. In Seventh IEEE Interna-tional Symposium on Wearable Computers, 2003. Proceedings, pp. 248-249

[3] A. Schmidt, et al., Advanced interaction in context. In Handheld and Ubiquitous Computing: First International Symposium, Huc'99, Karlsruhe, Germany, September 27-29, 1999, Pro-ceedings, pp. 89-101. Springer.

[4] R. Ballagas, F. Memon, R. Reiners, and J. Borchers. 2006, iStuff Mobile: prototyping in-teractions for mobile phones in interactive spaces. In Proc. PERMID, Workshop on Perva-sive Mobile Interaction Devices at PERVASIVE.

[5] A. M. Khan 2012, “You don't miss a phone call: Recognizing vibrating rhythms”, HCI 2012; Baltimore, USA. DOI: 10.2316/P.2012.772-034

[6] L.A. Jones, J. Kunkel, and E. Piateski, "Vibrotactile Pattern Recognition on the Arm and Back," Perception, vol. 38, pp. 52-68, 2009.

[7] L. M. Brown, T. Kaaresoja, “Feel who is talking: using Tactons for Mobile Phone Alerts.”, CHI '06. Proceedings, pp. 604-609

[8] S. Feige, “Can you feel it? Using Vibration Rhythms to Communicate Information in Mobile Con-texts.”, INTERACT’09. Proceedings, pp. 120-125

[9] L. M. Brown, S. A. Brewster and H.C. Purchase, “A First Investigation into the Effectiveness of Tactons”, in Proc. World Haptics 2005, IEEE Press (2005), pp. 167-176.

[10] Body Mass Index - http://www.nhlbisupport.com/bmi/ (15th December, 2014).

[11] Brewster, S.A. and Brown, L.M. Tactons: Structured Tactile Messages for Non-Visual Information Display. In Proceedings of Australasian User Interface Conference 2004 (Dunedin, New Zealand), Austalian Computer Society, pp15-23.

[12] D. Ternes and K. E. MacLean, “Designing Large Sets of Haptic Icons with Rhythm”, University of British Columbia Vancouver, BC, Canada V6T 1Y8.

[13] R. W. Cholewiak, A. A. Collins, "Vibrotactile localization on the arm: Effects of place, space and age", Perception and Psychophysics 65, 7 (2003), 1058-1077.

The Effect of Touch-key Size and Shape on the Usability of Flight Deck MCDU

Lijing Wang, Qiyan Cao, Jiaming Chang
School of Aeronautic Science and Engineering
Beihang University
Beijing, China
wanglijing505@126.com, caoqiyan@foxmail.com,
michael1044518077@gmail.com

Chaoyi Zhao
Human Factors Laboratory
China National Institute of Standardization
Beijing, China
zhaochy@cnis.gov.cn

***Abstract*—This paper focuses on the effect of touch-key size and shape on the usability of the multifunction control display unit (MCDU) of the flight deck. A total of thirty subjects participated in the trials on the touchscreen-MCDU to perform the task of preflight preparation. The sizes of the touch-key were 7mm, 10mm, 13mm, 16mm, 19mm and 22mm; the touch-keys were divided into two shapes: rectangle and square ones. The completion time of the task, the error rate and participants' subjective ratings were collected as the indicators of the usability, and Analysis of Variance (ANOVA) was done to test the data. Results showed that both touch-key size and shape affected the usability. The usability of the touchscreen-MCDU increased as the touch-key size increased up to a certain size (19 mm in this study), at which they reached asymptotes. The square touch-keys provided a better usability than the rectangle ones when the width was the same. However, when the width reached 19mm, the usability stayed the same for both shapes.**

Keywords-touch screen; flight deck; MCDU; touch-key size; interface.

I. INTRODUCTION

Touch screen has been widely used in mobile phones, tablets, laptops, etc., all due to its intuition, convenience and adjustability. People are used to using touch screen instead of traditional keyboard as the input method in their daily life. Despite its spreading use in the daily life, touch screen can also be introduced into other fields.

As the function of the flight deck grows with the need of airlines, the complexity of the operating environment and avionics systems also advances at a rapid pace. The complexity of the flight deck brings more workload for the pilots and raises the risk of pilots' operations. The idea of introducing touch screen into the flight deck might simplify the flight deck interface and help pilots perform better [1]. Rockwell Collins has unveiled its touch-control single primary flight display [2]. Garmin G3000 also implemented the touch screen technology on the flight deck [3]. The USA Joint Strike Fighter F-35 used two touch screen displays to replace the traditional displays on the flight deck [4]. At the 2011 Pairs Air Show, Thales first exhibited its next generation touch-screen cockpit concept [1]. The implement of touch screen into the flight deck might be a trend for the next generation flight deck. MCDU is one of the most important input devices on the flight deck. The traditional keyboard of MCDU has always been made up of buttons and switches ever since its first implement into the flight deck. However, touch screen has earned its way into the design of MCDU. Thales [1] and Barco [5] both brought the inspiration of touch screen into their new concept of the MCDU. Though the concept of bringing touch-screen MCDU into the flight deck has come up, few studies focused on the human factors issue. This paper will focus on the human factors study of the touch-screen MCDU on the flight deck.

With the widely use of touch screen, quite a number of researchers studied touch screen interfaces. The studies about touch screen included: target sizes [6][7], gestures [8][9], extra muscle fatigue [10][11], touch screen for older or disabled users [12], etc. Touch screen target size is one of the most important element of the interface. Among the studies, touch screen target size has always been popular, but the appropriate sizes for different touch screen devices varied. Colle & Hiszem [13] found that the 20 mm was an appropriate key size for touch screen numeric keypad. Kim [14] suggested in his study that the appropriated touch-key size of the In-Vehicle Information System should be 17.5mm. Parhi [6] recommend that the target size should be at least 9.2 mm for single-target tasks and 9.6 mm for multi-target tasks for one-handed thumb use of mobile handheld devices equipped with a touch-sensitive screen. Schedlbauer [15] did further study based on the experiment of Colle & Hiszem [13], and ended with the conclusion of 12mm to be the appropriate size. Though, studies about touch screen have been researched for almost two decades, nevertheless, there were few studies about the target size of the MCDU on the flight deck.

This paper will describe an experiment about the target size of the MCDU on the flight deck, investigate the difference of the usability for different sizes and propose an appropriate target size for the touch screen MCDU on the flight deck. In Section 2, the method of how to perform the experiment will be discussed. Section 3 shows the usability results of the experiment from three perspectives: time, error rate and subjective ratings. The issue to be considered in the experiment and the conclusion of the experiment is described in Section 4 and 5, respectively.

 ISBN: 978-1-61208-382-7

II. METHOD

A. Subjects

A total of thirty subjects participated in the human factors experiment. They were selected from the Flying College of Beihang University, male, who had learnt flying skills for at least one and a half years and understood how to use the real MCDU on the flight deck. All of them were willingly to participate in this experiment. Their ages ranged from 18 to 24 years old (mean = 21.3, SD =2.76). They had normal vision and no motor impairments.

B. Apparatus

According to the procedure of Crane [16] proposed in her study, a PC experiment was carried out before the final experiment on the flight deck. The experiment of this paper was conducted on a laptop and the further experiment would be planted in a flight deck simulation.

The experiment was conducted with a Lenovo Flex 2 laptop, which was equipped with the Windows 8 operation system and a touch screen. The size of the touch screen was 145×310mm, and the resolution was 1366×768 pixels.

The simulated MCDU software was projected onto the Lenovo laptop. The simulated MCDU software was similar to the real A320 MCDU and can perform the "initial A" function of the preflight preparations (see Figure 1). The soft keys to be pressed on the MCDU contained two different types: 1) the alphabet and digit keys of square sizes and 2) the "R*" and "L*" keys of rectangle sizes. The width of the keys was the same.

Figure 1. The simulated touch screen MCDU software

C. Experiment Design

The independent variables of this experiment are the touch-key size and shape of the MCDU. Two kinds of touch-keys with the same width were used in the interface: square ones and rectangle ones. The ratio of width/height of the rectangle keys designed here was 1.5. The width started at 7mm, and was incremented by 3mm. A preliminary research was conducted to determine the largest touch-key size. As said in the introduction, the appropriate target size suggested by other researchers ranged from 10mm to 20mm [6][7][13][14][15]. Taking the previous studies and their conclusions into account, the largest size of the simulated MCDU was determined to be 22mm. The final set of the touch-key size included: 7mm, 10mm, 13mm, 16mm, 19mm and 22mm. Figure 2 is part of the interface. It showed the six set of square touch-key sizes for the touch screen MCDU. Figure 3 showed the comparison of the square and rectangle touch-keys on the same interface.

D. Experiment Tasks

The task of preflight preparation was chosen to be the experiment task, which consisted of eight sub-tasks. First, subjects pressed the key of "INI" to start the task. As soon as the key of "INI" was pressed, the software started to record the time. Second, enter the ICAO code of the airport of departure and landing (ZYTL/ZBTJ), press "R1" and then "L6" to return. Third, the ICAO code of the alternate airport (ZBSH) should be entered, and press "L2" to confirm the entry. Fourth, flight number "CSN6125" should be entered and press "L3" to confirm that. Fifth, cost index "35" would be entered, and "L5" would confirm the entry. Then, enter cruise altitude "FL310" and press "L6" to confirm it. After this, subjects pressed "R3" to confirm the status. In the end, the key of "F-PLN" was pressed to finish the trial, meanwhile, the time record stopped. In addition, the key turned blue whenever the key was pressed.

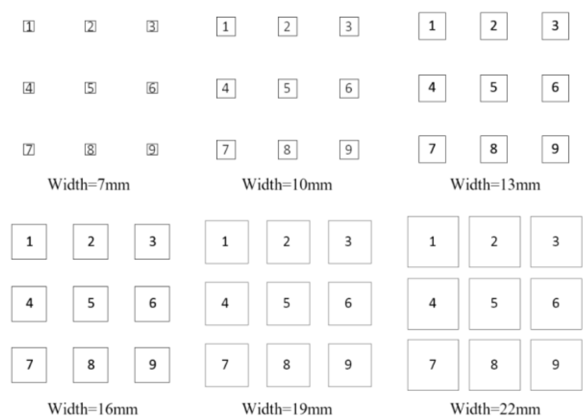


Figure 2. Six set of square touch-key sizes on the touch screen MCDU

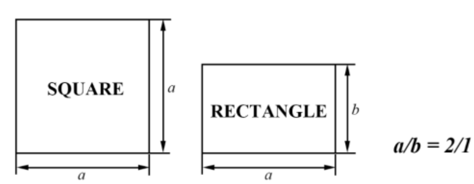


Figure 3. Comparision of the square and rectangle touch-key sizes on same interface

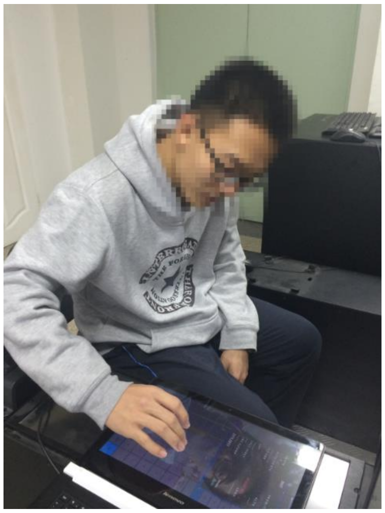


Figure 4. The positon of subjects performing the experiment

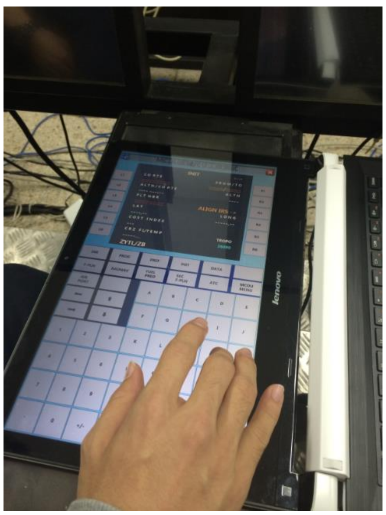


Figure 5. Subject performing the task on the touch screen MCDU

When performing the tasks, the laptop was put on the right side of the subjects to imitate the positon of the real MCDU on the flight deck (see Figure 4). Figure 5 showed the situation of one subject performing the task on the touch screen MCDU.

E. Experiment Procedures

The experiment consisted of four steps. First, the demographic information, such as age, eyesight and flying skills experience, was collected. Second, experimenters explained the purpose and procedures of this experiment to each subject. Written instructions of the experimental objectives and procedures were also given to help them understand. Third, after the subjects fully understood the experiment, they practiced on the touch-screen MCDU to become familiar with the interface and memorize the operation tasks. Fourth, in the main experiment, the subjects were asked to operate the given tasks. The subjects were told to finish the task as quickly and accurately as possible. There was no rule on which finger they used, however only one hand can be used. Subjects completed the operation task, which was already described in Section D. Experiment tasks. For each of the 6 different sizes, the same operation would be performed. To avoid the influence of fatigue and familiarity with the operation tasks, the order of the sizes randomly showed up. After each of the six tasks, the subjects were asked to fill a usability questionnaire to evaluate the usability of the interface. A Likert scale questionnaire was used, ranging from 1-9, where 1 meant the interface was too hard to operate and 9 meant the interface was perfect. In the end, the subjects were asked to give comments on sizes of the touch-key.

III. RESULTS

The results of this experiment included two kinds of data, the objective one and the subjective one. The objective data collected in this study included TIME and ERROR. TIME means the time it took to finish the task, i.e., the time from the point of pressing the “INI” key to the point of pressing the “F-PLN” key. ERROR means the number of errors occurred during the task. The subjective data was the ease of use and satisfactory of the subjects for each size.

The interface we designed had keys of two different shapes: rectangle and square. The error rate for both two shapes was also calculated.

A. Time

For completion time, one trial with extremely abnormal data was eliminated from the data set. The trail data was eliminated because it showed a completely different trend with the other data. The time for operating the task on the interface with the largest size was almost two times longer than the smallest size. The interview after experiment showed that the subject who performed this data was disturbed when he completed the task. Then, the largest size and the mean of the remaining trials were calculated for each size. The mean completion time of all the subjects was calculated first, and the ANOVA test was performed. The result of the ANOVA showed that Time was significantly affected($F(5,84)=5.925$, $p<0.002$). No surprisingly, as the size grew, participants were able to finish the task within a shorter time period. However, except for the smallest size, the completion time slowly decreased in comparison with the increase of the key size (see Figure 3). There is no significant difference between the size of 19mm and 22mm ($F(1, 18) = 0.075$, $p>0.5$).

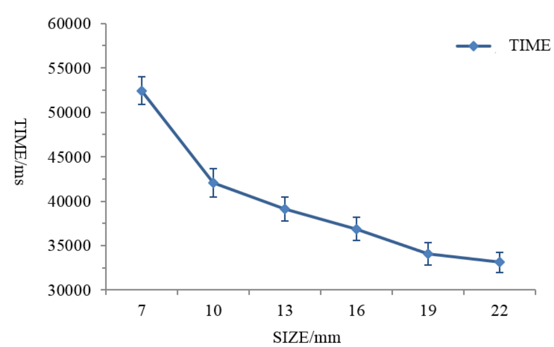


Figure 6. Mean compeltion time of different touch-key sizes

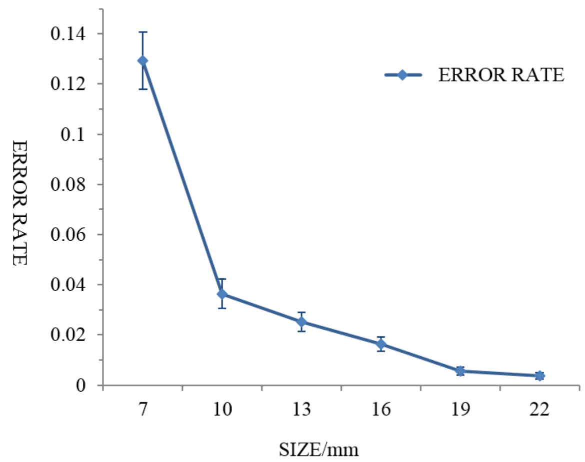


Figure 7. Mean error rate of different touch-key sizes

B. *Error Rate*

Error rate was the ratio of the number of errors occurred during the task to the total number of operations of the task. The mean error rate of all the subjects was calculated first, and the ANOVA test was performed. The error rate of the task is shown in Figure 8. Error rate was significantly affected by size ($F(5,84)= 17.94$, $p<0.001$). Errors declined as size increased. The smallest size (7mm) made an extremely high error rate which was three times compared to the 10mm size. The error rate of next three sizes showed a slow decline as the size increased. And there was no significant difference between the size of 19mm and 22mm ($F(5, 84)= 1.119$, $p>0.1$).

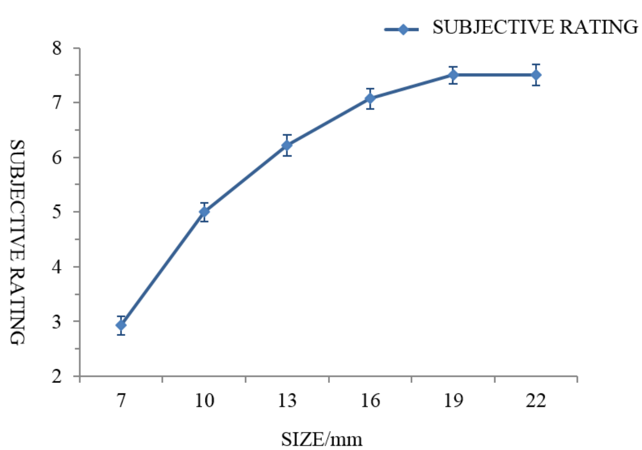


Figure 8. Mean subjective ratings of different touch-key sizes

C. *Subjective Ratings*

The mean subjective ratings of all the subjects was calculated first, and the ANOVA test was performed. The subjective ratings of the task are shown in Figure 9. The subjective ratings was also significantly affected by size ($F(5,84)=29.18$, $p<0.001$). The subjective ratings increased as the size increased. The bigger the key was, the better subjective ratings were given by the subjects. There was no significant difference between the size of 19mm and 22mm.

D. *Error Rate Comparision between the Retangle and Square*

The error rate of the rectangle and the square keys was also calculated. The error rate of the two shapes is shown in Figure 9.

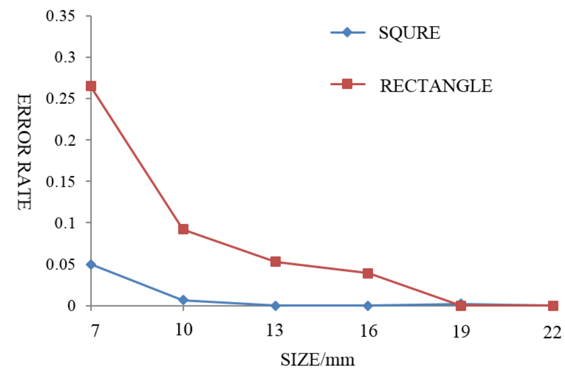


Figure 9. Error rate comparison between two different touch-key shapes(retangle and square touch-keys)

The ratio of height/width for the rectangle keys used in the experiment was less than 1. Under these circumstances, the error rate of the rectangle keys was higher than the square ones when the width was smaller than 16mm. However, with the larger size, 19mm and 22mm, subjects seldomly made mistakes. So the trend of the two lines coincided.

IV. Discussion

Overall, the usability of MCDU of the flight deck increased as the touch-key size increased. When the touch-key size was larger than 19mm, the completion time, error rate and subjective ratings did not differ much.

From the data we obtained from the experiment, some subjects gave higher values to the size of 19mm than the size of 22mm. For the ones who rated this way, an interview was done to determine the reason. Some subjects assumed the reason for this was that the spacing between the touch-keys decreased as the width of the touch-keys increased, which caused a cluster in the interface. As a result of the cluster, subject had a feeling that they would accidently tap on the adjacent touch-keys. Others claimed that the 22mm touch-keys were too large to perform for them under the circumstance. Studies about these questions would be done in the future.

Meanwhile, the error rate of the two different touch-key shapes was also examined. The result showed that error rate differed between different touch-key shapes. The error rate of the square keys was lower than the rectangle ones. However, the error rate reached an asymptote at the width of 19mm. The reason for this might be the difference in the tapping area. The square keys had larger area when the width was the same, correspondingly, the error rate was lower.

 ISBN: 978-1-61208-382-7

In this study, we focused on the factors of the key size and shape related to usability; in addition, other design factors, such as the key spacing and the location of the keys, should be investigated.

V. CONCLUSION

In this study, the effect of touch-key sizes and shapes on the usability of MCDU of the flight deck was examined. The usability of the MCUD on the flight deck increased with increased touch-key sizes. However, the usability reached an asymptote beyond certain touch-key sizes. When the ratio of the height/width was less than 1, the usability of the square touch-keys was better than the rectangle ones with the same width. However, the usability reached an asymptote when the width reached 19mm. Moreover, the touch-key sizes proposed by this study cannot be directly applied to the design of MCDU touch-key sizes because the experiment was only a PC simulated experiment. Therefore, a number of further studies should be performed to determine the size of the real MCDU of the flight deck.

REFERENCES

[1] THALES, "The cockpit of the future by Thales Innovation for future civil and military avionics platforms", Le Bourget, June 20-26, 2011.

[2] Rockwell Collins Press, Rockwell Collins unveils touch-control flight displays to enable natural, eyes-forward operation, [Online]. Available from: http://www.rockwellcollins.com/sitecore/content/Data/News/2011_Cal_Yr/CS/FY11CSNR50-touchcontrolled-flight-display.aspx, July 24, 2011[retrieved: July, 2014].

[3] S. Pope, Garmin G3000 brings touch screen tech to flight deck, Avionics Magazine, October 20, 2009.

[4] F-35 Cockpit, [Online]. Available from: http://www.darkgovernment.com/news/f-35-cockpit/[retrieved: May, 2012].

[5] Barco, "Touch the future: concept of a multi-touch cockpit". July 20,2011.

[6] P. Parhi, A. K. Karlson, and B. B. Benderson, "Target size study for one-handed thumb use on small touchscreen devices," Proceedings of 8th International Conference on Human Computer Interaction with Mobile Devices and Services, Espoo, Finland, 2006, pp. 203-210.

[7] Y. Park, S. H. Han, J. Park, and Y. Cho, "Touch key design for target selection on a mobile phone," In: the Proceedings of the 10th Mobile HCI conference.Amsterdam, Netherland. 2008, pp. 423 – 426.

[8] E. Choi, S. Kwon, D. Lee, and M. K. Chung, "Towards successful user interaction with systems: Focusing on user-derived gestures for smart home systems," Applied Ergonomics, 45(4): 2014, pp. 1196-1207.

[9] R. Leitao and P. A. Silva, "Target and spacing sizes for smartphone user interfaces for older aldults: design patterns based on an evaluation with users," In: Proceedings of the 19th Conference on Pattern Languages of Programs (PLoP 2012), 2012.

[10] K. B. Chen, A. B. Savage, and A. O. Chourasia, D. A. Wiegmann, M. E. Sesto, "Touch screen performance by individuals with and without motor control disabilities," 2013, pp. 297-302.

[11] C. Irwin and M. Sesto, "Performance and touch characteristics of disabled and non-disabled participants during a reciprocal tapping task using touch screen technology," Appl. Ergon. 43, 2012, pp. 1038-1043.

[12] J. Zhao, T. Plocher, and L. Kiff, "Touch screen user interfaces for older adults: button size and spacing." Universal Acess in Human Computer Interaction. Coping with Diversity. Springer Berlin Heidelberg, 2007, pp. 933-941.

[13] H. A. Colle and K. J. Hiszem, "Standing at a kiosk: Effects of key size and spacing on touch screen numberic keypad performance and user preference," Ergonomics, vol. 47, no. 13, 2004, pp. 1406-1423.

[14] H. Kim, S. Kwon, J. Heo, H. Lee, and M. K. Chung, "The effect of touch-key size on the usability of In-Vehicle Information Systems and driving safety during simulated driving," Applied Ergonomics, vol. 45, May. 2014, pp. 379-388, doi: org/10.1016/j.apergo.2013.05.006.

[15] M. Schedlbauer, "Effects of key size and sapcing on the completion time and accuracy of input tasks on soft keypads using trackball and touch input," Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, 2007, pp. 429-433.

[16] J. M. Crane, E. S. Bang, and M. C.Hartel, "Standardizing interactive display functions on the 777 flight deck," R. SAE Technical Paper, 1994.

ACKNOWLEDGEMENTS

The work presented in this paper has been carried out as part of the Chinese project the National Key Technology R&D Program of the Ministry of Science and Technology (2014BAK01B03), led by China National Institute of Standardization.

A Literature Review: Form Factors and Sensor Types of Wearable Devices

Dong Yeong Jeong, Sung H. Han, Joohwan Park, Hyun K. Kim, Heekyung Moon, Bora Kang
Department of Industrial Management & Engineering
POSTECH
Pohang, South Korea
{comnet924, shan, pkjhwan, emokubi, gomsak, purple31}@postech.ac.kr

***Abstract*— Wearable devices provide a new way to recognize the users context with high accuracy. Selecting suitable form factors and sensors are important to recognize users' contexts. In this study, the form factors and sensor types of released, prototype, and concept products were explored. A total of 175 literatures were collected and analyzed in terms of sensor and form factor. Thirty sensors were collected and classified according to measurands. Twenty-three form factors were listed by nine applicable body parts.**

Keywords-wearable device; form factor; sensor; body part.

I. INTRODUCTION

Wearable devices can provide new ways of sensing users' contexts [1]. Since they can be worn, wearable devices tend to have a higher context recognition rate than existing personal mobile devices [2]. For that reason, wearable devices can provide new functions based on context recognition to improve the usability and efficiency of services [3].

There are several body parts on which suitable wearable products can be worn. The wearable products can be classified into several "form factors." These form factors are such wearable devices as necklaces and glasses. Form factors and sensors are used to develop wearable devices [4][5]. However, it is difficult to find papers discussing current form factors and sensors of wearable devices.

In this paper, sensors and form factors of wearable devices are investigated based on a literature survey. Sensors are classified according to the measurands. Form factors and sensor types are listed according to their applicable body parts.

II. METHOD

Information on the released, prototype, and concept products was collected from journal papers, proceeding papers, news, and patents. A total of 175 literatures published between 2001 and 2013 were collected.

Sensors were classified according to their measurands, based on the study of Richard [6]. Image/vision and gas sensors were added to the sensors found in the literatures.

Form factors and sensors were listed according to their wearability on the body parts. The body parts are classified according to the kinesiological classification: head, neck, shoulder girdle, Trunk, arm, forearm, hand, pelvis, thigh, leg, and foot [7].

III. RESULT

A. *Sensor classification*

The sensors are classified into nine types (Table I).

- Photo sensor: sensor which detects characteristics of visible, infrared, and ultraviolet light.
- Image/vision sensor: sensor which identifies the visual pattern, shape, location, and movement from the image and video information.
- Electro-magnetic radiation sensor: sensor which detects the electro-magnetic waves.
- Electrical activity sensor: sensor which detects the electric properties of devices and the body.
- Magnetic field sensor: sensor which detects the magnetic field through the property of solid matter and voltage change.
- Gas sensor: sensor which recognizes the property of gas, and it measures the components, concentration, and pressure of the gas.

TABLE I. SENSOR CLASSIFICATION

Sensor Type	Sensor
Photo sensor	• Free space optical communication • Proximity sensor • Pulse oximetry sensor • Photoelectric sensor • RGB/illuminance sensor
Image/Vision sensor	• Motion detection sensor • 3D depth camera • Infrared camera • Eye-tracking sensor • Vision/Image recognition
Radiation sensor	• GPS sensor • Near field communication module
Electric sensor	• EEG/EMG/ECG/EOG sensor • GSR sensor • Capacitive sensor • AC current sensor
Magnetic sensor	• Magnetic field variation sensor • Geomagnetic sensor
Gas sensor	• Gas component analysis sensor • Atmospheric pressure sensor • Humidity sensor
Acoustic sensor	• Microphone • Phonomyography
Mechanical sensor	• Pressure sensor • Gyroscope • E-textile • Acceleration sensor • Physical button
Thermal sensor	• Body temperature sensor • Heat flux sensor • Temperature sensor

 ISBN: 978-1-61208-382-7

- Acoustic sensor: sensor which detects the sound waves of the dial tone, voice, and ultrasonic waves.
- Mechanical sensor: sensor which recognizes the physical movement and other mechanical properties of the human and the device.
- Thermal sensor: sensor which measures the size and flow of temperature.

B. Form factors and sensors with body parts

Form factors and sensor types were listed by their applicable body parts (Table II). There is no form factor found to be applicable to neck, shoulder girdle, and pelvis.

The biggest number of form factors was related to the head. The smallest number of form factors was related to the arm, thigh, and leg. Module/clip, patch, clothes were used for various body parts.

Mechanical sensors were combined with most form factors.

TABLE II. FORM FACTORS AND SENSOR TYPE WITH BODY PART

Body Part	Form Factor	Sensor Type
Head	Hat	Image/Vision, acoustic, mechanical
	Helmet	Image/vision, gas, mechanical
	Glasses	Photo, image/vision, acoustic, mechanical, electro-magnetic radiation, magnetic field
	Headset	Mechanical, image/vision, electrical activity, acoustic, photo, electro-magnetic radiation
	Earphone	Photo, electrical activity, mechanical
	Ear Wrap	Electro-magnetic radiation
Trunk	Necklace	Image/vision, mechanical, acoustic, electro-magnetic radiation, electrical activity
	Chest Band	Electro-magnetic radiation, electrical activity
	Belt	Photo, electro-magnetic radiation, magnetic field, gas, acoustic, mechanical
	Bag	Gas
Arm	Arm Band	Electrical activity, magnetic field, mechanical, thermal
Forearm	Wristwatch	Photo, image/vision, acoustic, thermal, electro-magnetic radiation, mechanical, electrical activity, magnetic field
	Bracelet	Photo, image/vision, acoustic, thermal, electro-magnetic radiation, mechanical, electrical activity, magnetic field
Hand	Gloves	Magnetic field, mechanical
	Stick	Photo, electro-magnetic radiation, gas, mechanical, thermal
	Ring	Photo, image/vision, electrical activity, acoustic, mechanical
Thigh	Robot	Mechanical
Leg	Ankle Band	Magnetic field, mechanical
Foot	Shoes	Magnetic field, mechanical
	Socks	Electro-magnetic radiation, electrical activity
Whole body	Module/Clip	Image/vision, mechanical, electrical activity, electro-magnetic radiation
	Patch	Photo, electrical activity, magnetic field, mechanical, thermal
	Clothes	Photo, image/vision, electro-magnetic radiation, magnetic field, acoustic, mechanical

IV. DISCUSSION

Most form factors are related to terminal body parts: head [8], hand [9], forearm [10], foot [11], and leg [12]. It is easy to wear things on terminal body parts. There are also various signals that can be detected on terminal body parts easily. For example, the head is a useful location to detect the gaze direction, pulse, breath. The hand is useful to detect the pulse and the signal of user's activity. The foot is useful to sense body pressure and walking pattern.

Modules/clips form factors have been used to detect non-physiologic signals that can be detected on any body parts [13]. The existing modules/clips can be transformed to another form factors using additional accessories [14]. Modules/clips [13], patches [15], and clothes were used when the wearable devices needed to capture signals in several body parts simultaneously [16].

V. CONCLUSION

This paper classified the sensors according to the measurands, and listed form factors and sensors according to various applicable body parts. Although, form factors were applicable to several body parts, most form factors were found on the terminal body parts. This paper helps the designers of wearable devices to understand the various form factors and sensor types.

REFERENCES

[1] M. D. Rienzo, F. Rizzo, G. Parati, G. Brambilla, M. Ferratini, and P. Castiglioni, "MagIC system: A new textile-based wearable device for biological signal monitoring. Applicability in daily life and clinical setting," In Engineering in Medicine and Biology Society, 2005, pp. 7167-7169.

[2] J. Farringdon, A. J. Moore, N. Tilbury, J. Church, and P. D. Biemond, "Wearable sensor badge and sensor jacket for context awareness," In Wearable Computers, October. 1999, pp. 107-113.

[3] K. Hung, Y. T. Zhang, and B. Tai, "Wearable medical devices for tele-home healthcare," In Engineering in Medicine and Biology Society, Vol. 2, September. 2004, pp. 5384-5387.

[4] C. Loclair, S. Gustafson, and P. Baudisch, "PinchWatch: a wearable device for one-handed microinteractions," In Proc. MobileHCI, 2010.

[5] J. Scheirer, R. Fernandez, and R. W. Picard, "Expression glasses: a wearable device for facial expression recognition," In CHI'99 Extended Abstracts on Human Factors in Computing Systems, 1999, pp. 262-263.

[6] W. M. Richard, "A sensor classification scheme," In IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol.2. UFFC-34, 1987, pp. 124-126.

[7] J. E. Muscolino, Kinesiology, 2nd ed., Mosby, pp. 6, April 1955.

[8] M. Funk, A. Schmidt, and L. E. Holmquist, "Antonius: A mobile search engine for the physical world," In Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication, September. 2013, pp. 179-182.

[9] C. Amma, M. Georgi, and T. Schultz, "Airwriting: Hands-free mobile text input by spotting and continuous recognition of 3D-space handwriting with inertial sensors," In Wearable Computers (ISWC), 2012 16th International Symposium on, June. 2012, pp. 52-59.

[10] M. D'Souza, T. Wark, M. Karunanithi, and M. Ros, "Evaluation of realtime people tracking for indoor environments using ubiquitous motion sensors and limited wireless network infrastructure," Pervasive and Mobile Computing, Vol. 9 (4), 2013, pp. 498-515.

[11] J. Healey, “Gsr sock: A new e-textile sensor prototype,” In Wearable Computers (ISWC), 2011 15th Annual International Symposium on, June. 2011, pp. 113-114.

[12] C. Strohrmann, H. Harms, G. Tröster, S. Hensler, and R. Müller, “Out of the lab and into the woods: kinematic analysis in running using wearable sensors,” In Proceedings of the 13th international conference on Ubiquitous computing, September. 2011, pp. 119-122.

[13] N. K. Taylor, E. Papadopoulou, M. Y. Lim, P. Skillen, F. R. Blackmun, and H. Williams, “Congestrian: monitoring pedestrian traffic and congestion,” In Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication, September. 2013, pp. 1355-1358.

[14] Fitbit one hompage. [Online]. Available from: http://www.fitbit.com

[15] T. Fahrni, M. Kuhn, P. Sommer, R. Wattenhofer, and S. Welten, “Sundroid: Solar radiation awareness with smartphones,” In Proceedings of the 13th international conference on Ubiquitous computing, September. 2011, pp. 365-374.

[16] R. Johnson, N. Bianchi-Berthouze, Y. Rogers, and J. van der Linden, “Embracing calibration in body sensing: using self-tweaking to enhance ownership and performance,” In Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, September. 2013, pp. 811-820.

Identifying Interaction Problems on Web Applications due to the Change of Input Modality

André Constantino da Silva, André Luis Correa Viana, Samuel Gomes de Lima
Instituto Federal de São Paulo – Campus Hortolândia
Hortolândia, São Paulo, Brazil
andre.constantino@ifsp.edu.br, andre_viana40@yahoo.com.br, samuelsmgl@gmail.com

***Abstract*—New equipments and software for providing different modes for user interaction emerged and became popular in the last decade; they are used in various devices, such as mobile and game consoles, which can display Web pages due to the increasing of the processing power in these devices. Since interaction with each mode has peculiarities, when a user is interacting with a mode not considered in the design time, she might have interaction problems. Here, we present our work about Web navigation with motion sensors; we chose Google Maps, Google Street View, and TelEduc, which is an e-learning environment, to evaluate the use of WiiMote, the motion sensor of Wii console. In this paper, we analyze data from the first case study (Google Maps and Google Street View) and preliminary results from the second one (TelEduc). The collected data confirm our previous findings; adaptations are necessary to users have a good experience when navigating through web pages using a motion sensor.**

Keywords-Multimodality; Usability Evaluation; Web Applications.

I. Introduction

The number of input hardware on desktop computers is increasing, with computers equipped with touchscreen devices, microphones and cameras. New computational devices with input and output hardware are been built; most of them are equipped with processing power enough to render Web pages, allowing users to navigate into Web pages using modalities not available in desktop computers. Websites/portals, even the ones that had been passed by usability testing, have interaction problems when changing the mode of interaction (cross-modality problem) [1]. Since usability is a key factor for system acceptance [2], low usability in a context can cause the user´s withdrawal from interacting with an application.

Soon, given the growing use of devices equipped with motion sensors, users equipped with these devices will access Web applications. Thus, it becomes necessary to study problems caused by the changing of interaction modality from keyboard+mouse to motion sensors. Our hypothesis is that for users to have a good experience in navigate through the Web applications using motion sensors some page adaptations are necessary. To suggest adaptations on the Web pages it is necessary to understand the interactions problems that happen due to modality changing. Our general goal is to have a better understanding of the impact due to the modality changing and to propose guidelines to design user interfaces with high usability for applications that can be accessed by many modalities.

Here, we present our work about Web navigation with motion sensors and some interaction problems that happen when users navigate through Web pages using motion sensors. We planned some case studies based on user tests, where volunteers used the WiiMote, the motion sensor controller of Wii console, to perform some tasks in Google Maps, Google Street View [3] and TelEduc [4]. Section II presents a literature review about the subject. Section III describes our methodology and data analysis. Section IV summarizes the results of our work so far.

II. Literature Review

Nielsen [2] defines the general acceptability of a system, composed by social acceptability and practical acceptability. One of the concepts that composes the practice acceptance is the “usefulness”, which refers to use the system to achieve a particular objective. This concept consists of the usefulness and usability. Usability is a combination of five elements: ease of learning, efficiency, ease of remember, probability of the user making few errors and user satisfaction. Since usability is indirect related with the acceptability of a system, we can affirm that the user experience in using a system is a factor that determines the user´s system acceptance.

Modality is the term used to define a mode in which a user´s input or system´s output is expressed. Nigay and Coutaz [5] define modality as an interaction method that an agent can use to reach a goal, and that modality can be specified in general terms as “speech” or in more specific terms as “using microphones”. Several modalities have become research topics in recent decades; among them, we can mention the voice, handwriting recognition, touch, and gestures [6]. Bernsen [7] says that there are no two equal modalities; each of them has its own strengths and weakness.

In our previous work, we cataloged some problems that occur when accessing the TelEduc, an e-learning environment, in touchscreen devices and can occur in other Web applications [1]. These problems were identified through analysis of results generated by inspection methods (methods that involve usability experts) and empirical tests (involve the participation of users), and the identified problems were divided into three categories: problems related with the change of the modality, problems related with the change of platform and problems not related with the change of modality or platform. In another previous work

 ISBN: 978-1-61208-382-7

[8], we presented some results about the use of WiiMote to interact with Google Maps and Google Street View from data collected by questionnaire. The applied opinion questionnaire was composed by 16 (sixteen) questions related to enjoyment, perceived response time of the movement, handling digital artifact activation of the user interface components, selection of options, and navigation in virtual space. Comparing the users opinion, there are a slight difference between navigate in Windows 8 and Linux Ubuntu. The main difference is trigger small buttons and the reason can be related with the used pointer's icon: in the Windows is a half-filled circle and in Linux is the standard mouse icon. The standard icon has a lower visibility, but a higher accuracy. Related to interaction problems due to the size of the interface elements and the high sensitivity of the control (performing pointer movement when pressing the button) points that the need to do some adjustments for a better interaction with this device on the Web. About the how easy was to trigger small buttons of the Google Maps' user interface, 3 volunteers said that it was easy most of the time, 2 volunteers said that it was easy half of the time, and 1 volunteer said that it was easy in few times. About how easy was to use the virtual keyboard to type a text, 3 volunteers said that it was easy every time, 2 volunteers said that it was most of the time, and 1 volunteer said that it was half of the time. Here, we present results from video analysis that support our previous findings.

III. METHODOLOGY

Our methodology is based on an empirical method, user tests in laboratory, to identify interaction problems. We planned two case studies, one for Google Maps and Google Street View, and another for TelEduc. We finished the first one case study, where six volunteers navigated via the applications. The second case study is in-progress; we just did the planning and a pre-test.

The main activities of the planning phase of both case studies were: (i) selection of the tools and features, and specification of tasks to be performed by volunteers, (ii) creation of the profile and opinion questionnaires, (iii) definition of hardware and software to be used, (iv) preparation of the environment, and (v) creation of the consent term.

In the execution phase, six volunteers performed the defined tasks, while the interaction was recorded; then, the specialists analyzed these videos. Before starting to interact with the application, the volunteers filled up the profile questionnaire, and after session test, the volunteers filled up the opinion questionnaire. The profile questionnaire was composed by 13 questions about age and genre, the previous experience in using the selected applications, asking about how long and frequency the volunteer had used a Wii console, and information about which sites, browsers, frequency of use of the Google Maps and Google Street View, and where the volunteer usually access Internet from.

A. Materials and Tools

The used hardware (Figure 1) was a motion sensitive controller of the Wii console, the WiiMote, connected through Bluetooth with a computer (a HP Touchsmart tx2-1040br). A 19 inches monitor was used to have a good readability in a distance required to use the motion controller, usually two steps back.

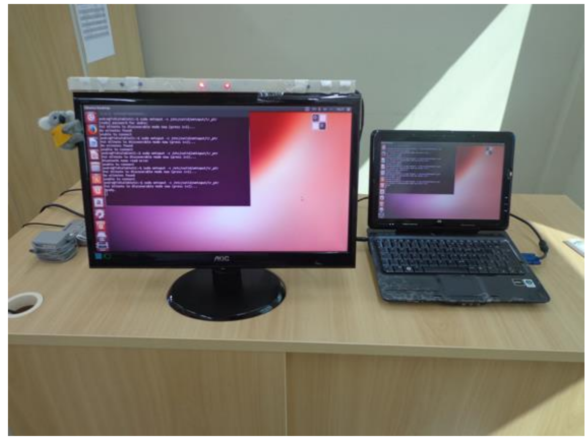


Figure 1. Equipements used in the case study: infrared sensor, 19" display and a computer.

The infrared bar was placed on the top of the monitor to have a higher accuracy in the recognition of movements. The infrared bar was built with a total of 8 infrared LEDs, but tests showed that there is a better accuracy of the movements when only two LEDs are connected; so, we used this configuration.

The mapping between input data from the Wii control to mouse actions enables to use this device to interact with a desktop computer. This mapping was done on Linux through WMInput application, available in the package CWiid version 0.6.00. We select this set of software (SO plus browser and input data receiver program) due to our previous experience [8] that showed fewer differences in navigate in Linux and Windows using the WiiMote. The main difference is related to the icon of the mouse pointer: in the WMInput (Linux) application, this is the standard one, while Win Remote application (Windows 8) uses a half-filled circle.

To supply the lack of a keyboard next to the user, we used a virtual keyboard fixed in the bottom of the screen occupying about two-thirds of the screen. In the Linux SO, we used the Onboard virtual keyboard and, in the Windows SO, we used the standard virtual keyboard. To navigate through the Internet, volunteers used the Google Chrome in both SOs.

B. Case Study 1: Navigation on Google Maps and Google Street View

Google Maps [3] is a Web application that provides information about geographical regions around the world, offering aerial and satellite views. Considering its functionalities, we defined two tasks to be performed by the volunteers: Task 1 - Find your home and point it; Task 2 – Show the route from your home to college. Google Street View is integrated with Google Maps and offers street views for some cities, function accessed by the use of *Pegman* icon in the Google Maps. Considering its functionalities, we wrote two tasks to be performed by the volunteers: Task 3 -

Join Google Street View and show the facade of your home; Task 4 - Show the route from your home to the nearest market.

Six volunteers who participated (three volunteers for each operation system) have on average 17 years old; one woman (volunteer 1) and five are males (volunteers 2, 3, 4, 5 and 6). All volunteers declare that does not have a Wii console, and only two volunteers have used sometime, although without frequently (volunteer 2 used for about 1 year, and volunteer 5 among 3-6 months). All volunteers said they access Internet, five of them reported using more than 10 hours per week, while one declared access it for about 5-10 hours a week. All volunteers declared access in their homes, and four declared that access at school. A volunteer reported using the Internet at friends and family home, and two volunteers declared access in places that offer free access through Wi-Fi or through 3G network. Three volunteers reported using Mozilla Firefox browser, while the other four ones reported using the Google Chrome browser. All volunteers reported had been used Google Maps; volunteer 5 informed that uses about once or twice a month and two other volunteers reported using more than three times per month (volunteers 2 and 3). About Street View, three volunteers reported had never been used, while one declared having used only once (volunteer 4), and two other volunteers reported using once every two months (volunteers 3 and 5).

About Task 1 - Find your home and point it, the users can use two different strategies or a combination of them: i) navigate in the map (volunteers 2, 3, and 5); or ii) use the keyboard to type the address (volunteers 1, 4, and 6). Analyzing the recorded video, volunteers had difficult to trigger the zoom buttons; this was observed mostly in interactions where volunteers used the first strategy due to the need to navigate and zooming into the map. Volunteer 2 tried 20 times to trigger one of the zooming buttons (zoom in or zoom out): 9 hits and 11 errors. To confirm that she was pointing out the correct place on map, volunteer 2 triggered the Satellite View feature. In this interaction, the user tried to trigger the feature 2 times (1 hint and 1 error). Volunteer 3 chose the same strategy; but, to trigger the zoom in feature, she did a combination of commands instead of only pressing the zoom in button in the user interface. The commands were press the left button, point to the zoom in option and press the right button. Due to the difficult in identify which buttons had been mapped to the mouse´s left click and right click, the volunteer tried 12 times to zoom in (9 hints and 3 errors due to press the wrong button). She did an error in more 2 times, when she tried to trigger the zoom in feature but when she pressed the WiiMote´s button, she moved the mouse cursor and the system recognized as image selection. Volunteer 5, who preferred to trigger the zoom buttons on the user interface, tried 11 times to trigger these buttons (8 hints and 3 errors).

About the interaction of the volunteers that preferred to type the address, volunteer 1 typed an address with 29 letters and triggered 37 buttons due to some wrong letters typed (4 wrong typed keys and more 4 on backspace button). Analyzing the wrong letters, we noticed proximity between the typed wrong letter and the correct one; e.g., the volunteer wanted to type the letter ‘a’ and button ‘s’ was pressed. This user also used the zoom buttons in their interaction, trying to trigger the zoom 8 times (3 hints and 5 errors).

Volunteer 4 tried to complete the task 1 by typing the address but, due to the difficult to complete it, she changed to the map navigation way. Using the typing address strategy, this volunteer typed 13 correct words of the desired address and tried to select one of the suggested addresses. Since she did not selected the desired address due to press the wrong button of the WiiMote (pressed the wrong button three times), she decided to complete the address, clicking on address bar and typing more 2 letters of the virtual keyboard. Finally, she pressed the ‘enter’ button of the virtual keyboard. The system showed the place of the typed address with an information window. The volunteer closed this window and the system returned to the previous position (before the address selection). The volunteer decided to change the address and typed one with 14 letters, all correctly typed. To adjust the view, the volunteer used the map navigation with zooming buttons. She moved the map three times, and tried to zoom 8 times, but pressed 3 times the wrong WiiMote´s button.

Volunteer 6 typed an address with 27 letters (no errors). The system suggested some streets based on previous interaction, and volunteer selected one. Due to the selected street is not the desirable one, the volunteer edited the address clicking in the text. In the first time, the volunteer selected a letter. She tried again just to put the text cursor in the correct place. After, the volunteer completed the address with 10 letters (the volunteer pressed 14 keys, due to press 2 wrong keys and need to erase them pressing the backspace button 2 times). Similar to volunteer 1, the wrong pressed keys were next to the desirable key: volunteer pressed ‘k’ key instead of ‘l’ key; and volunteer pressed ‘u’ key instead of ‘i’ key.

About Task 2 - show the route from your home to college, the volunteers needed to do map navigations using a combination of click and hold the left button over the map and moving the hand in the desire direction. Volunteer 1 did this combination to move the map 23 times, volunteer 2 did 8 times, volunteer 3 did 81 times (more 2 errors), volunteer 4 did 7 times, volunteer 5 did 25 times and volunteer 6 did 38 times. Some volunteers (1, 4, and 5) used the zoom features to perform the Task 2. Volunteer 1 tried to trigger the zoom feature 10 times (4 hits and 6 errors), volunteer 4 tried 4 times (4 hits) and volunteer 5 tried 4 times (3 hits and 1 error).

To perform Task 3 - Join Google Street View and show the facade of your home, volunteers needed to trigger the *Pegman* button, hold it and release it in the map. Before it, the volunteer could be some adjustment in the map view. After, volunteers needed to navigate to find the façade of her home. Volunteer 1 tried 16 times until trigger the *Pegman* function, volunteer 2 tried 4 times (3 errors), volunteer 3 tried 2 times (the first one was considered an error due to the volunteer released in the wrong place), volunteer 6 hit it in the first time.

To navigate in the Google Street View, volunteers 1 and 3 did not need to navigate, and volunteer 2 triggered 5 times (2 hits and 3 errors). Volunteer 6 hit the trigger in the first time (1 hit and 0 error). About the view position, volunteer 1 trigged 5 hits, volunteer 2 triggered 4 hits, volunteer 6 did 2 hits, and volunteer 3 did not need to change the position´s view.

To complete Task 4 - Show the route from your home to the nearest market, volunteers needed to navigate into the virtual world and change the position´s view. About the navigation, volunteers 1 and 3 used the direction pad of the WiiMote, pressing 1 and 41 times, respectively. Volunteer 2 tried to change the position 15 times (4 errors), and volunteer 6 tried 20 times (2 errors). The total of trying was 35 (29 hits and 6 errors), without count the use of direction pad. To change the position´s view, volunteer 1 hit 4 times, volunteer 2 hit 6 times, volunteer 3 hit 4 times and volunteer 6 hit 4 times. The total of hit was 18 times. No error was identified.

Table I summarizing the collected data. Volunteers 1, 2, 5, and 6 tried to trigger one of the zoom buttons (both have the same size, so we computed together) in Task 1, and volunteers 1, 4, and 5 in Task 2. They tried zooming 57 times, and performed 31 hits and 26 errors. Volunteers 1, 2, and 6 used the virtual keyboard in their interactions, pressed 80 keys: 68 hints, 6 errors and more 6 to erase the wrong typed letter. To trigger the *Pegman* button, volunteers tried 23 times (4 hits). Trigger the zoom buttons can be considered with low efficacy due to the half of the time the users do an error; but trigger the virtual keyboard´s key had a better efficacy. We believe the cause is the buttons´ dimensions. About map navigation, volunteers did 182 times the combination of pressing the button and moving the hand and made 2 errors.

TABLE I. HITS AND ERRORS IDENTIFIED IN VOLUNTEERS´ INTERACTIONS

	Command				
	Trigger zoom in and zoom out buttons	***Typing in the virtual keyboard***	***Trigger the Pegman button***	***Navigate into the virtual world***	***Change the position view***
Trials	57	80	23	41	29
Hits	31 (54.38%)	68 (85.00%)	4 (17.39%)	32 (78.05%)	29 (100%)
Errors	26 (45.62%)	6 + 6 (15.00%)	19 (82.61%)	9 (21.95%)	0

C. Case Study 2: Navigation on TelEduc

Online systems that support e-Learning through the Web are called e-learning environment systems or Virtual Learning Environments (VLE) or Learning Management Systems (LMS). An e-learning environment system is an application that uses the Web infrastructure to support teaching and learning activities, designed to support a variety of users and learning contexts. This environment is composed of tools that allow users to create content, communicate with other users, and manage the virtual space, e.g., chat, forums, portfolios, and repositories. One example is TelEduc, a Brazilian e-learning environment.

Due to the amount of tools available at TelEduc, we chose 6 tools (mail, agenda, pool, support material, portfolio and message board) and defined nine tasks to users perform: i) login, ii) read mail messages, iii) delete mail messages, iv) read the course´s agenda, v) vote in a pool, vi) watch a video, vii) post a portfolio item, viii) publish in the message board, and ix) logoff. Questions that composed the opinion questionnaire were about the difficult to: i) read text in the screen, ii) trigger icons such as navigator icon, iii) trigger the scroll bar, iv) the menu, v) the keys of the virtual keyboard, vi) put the keyboard focus on the text field, and vii) trigger small buttons, such as the minimize button.

We performed a pre-test to evaluate our questionnaire for the second case study. Analyzing the recorded interaction, we found difficult to trigger small buttons and links in words, to put the mouse pointer in a specific place and trigger the scroll bar due to the small size of its buttons and its width. Another problem happens when the user will press the button A: the mouse pointer moves. This problem was identified in another user test; but, in this case, the problem happen on trigger links: the user needs to try many times to reach her goal.

IV. FINAL CONSIDERATIONS

From our previous work [8], we can perceived that using motion sensor to navigate on Web page is fun and the users might feel motivated to adopt this interaction modality. In this paper, we presented the analysis of the recorded interactions, and identified that volunteers had difficulties to trigger small buttons and to typewrite text. This analysis confirms our hypothesis that adaptations are necessary to users have a good experience when navigating through web pages using a motion sensor. We believe that to minimize the number of interaction problems that happen due to the size of the user interface elements or the high sensitivity of the control (performing pointer movement when pressing the button) is necessary to do some adjustments, such as increasing the width and the height of the elements. We plan to investigate solutions for the identified problems and develop tools that aim to collect and analysis the generated data in similar case studies.

As future work, we plan to continue our study by analyzing the interaction using a Kinect movement sensor instead of WiiMote. To compare the data collect in the cases study, we plan to apply the same questionnaires and develop guidelines for design Web applications which users can access them with keyboard+mouse or motion sensors.

ACKNOWLEDGMENT

The authors thank the CNPq and IFSP for financial support.

REFERENCES

[1] A. C. da Silva, F. M. P. Freire, and H. V. da Rocha, "Identifying Cross-Platform and Cross-Modality Interaction Problems in e-Learning Environments," Proc. of 6th

International Conference on Advances in Computer-Human Interactions (ACHI 2013), IARIA, Feb. 2013, pp. 243-249.

[2] J. Nielsen, Usability Engineering, San Francisco, CA: Morgan Kaufmann Publishers Inc., 1993.

[3] Google Inc., Google Maps. Available from http://maps.google.com/.

[4] TelEduc Project. Home | teleduc.org.br. Available from http://www.teleduc.org.br/.

[5] L. Nigay and J. Coutaz, "A Generic Platform for Addressing the Multimodal Challenge," The SIGCHI Proceedings of the 13th Conference On Human Factors in Computing Systems (CHI 1995), ACM Press, May. 1995, pp. 98-105.

[6] B. Dumas, D. Lalanne, and S. Oviatt, "Multimodal Interfaces: A Survey of Principles, Models and Frameworks," Human-Machine Interaction, D. Lalanne and J. Kohlas, Eds. Berlin: Springer Berlin / Heidelberg, pp. 3-26, 2009.

[7] N. O. Bernsen, "Multimodality Theory," Multimodal user Interfaces: From signal to interaction, D. Tzovaras, Ed. Berlin: Springer, pp. 5-28, 2008.

[8] A. C. da Silva, A. L. C. Viana, and D. Marques, "Identifying Interaction Problems on Internet Navigation Caused by Change of Input Mode: a study about motino sensor controller, Google Maps and Google Street View," Proc. of 8th International Conference on Interfaces and Human Computer Interaction 2014, Game and Entertainment Technologies 2014 and Computer Graphics, Visualization, Computer Vision and Image Processing 2014, IADIS, Jul. 2014, pp. 271-275.