IntechOpen

Human Computer Interaction

Edited by Ioannis Pavlidis

Human-Computer Interaction

Edited by
Ioannis Pavlidis

Human Computer Interaction
http://dx.doi.org/10.5772/83
Edited by Ioannis Pavlidis

First published in Croatia, 2008 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019.
IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Human Computer Interaction
Edited by Ioannis Pavlidis
p. cm.
ISBN 978-953-7619-19-0
eBook (PDF) ISBN 978-953-51-5797-7

Preface

This book is an assembly of essays on state of the art concepts, methodologies, and systems in the fast moving field of human-computer interaction. Twenty-one chapters describe either progress in main themes or introduce totally new themes, never conceived before.

As the use of mobile devices proliferates, a major challenge is how to present effectively visual information on a limited real-estate. K.B. Lee introduces Zoomable User Interfaces (ZUI) to address this problem.

R. Michalski examines the effect of graphical panel's characteristics on usability. He further connects this to gender differences.

N. Juristo et al. argue that particular usability issues have a significant impact on software design and thus, need to be considered early in the development cycle, stating with the requirements formulation.

F. Naya et al. introduce Computer Aided Sketching (CASk) as a means to bridge the disconnect between CAD and the initial design stages of a product.

S. Nilsson provides a detailed account of Augmented Reality (AR) as the technology that aims to merge the real and virtual world. AR is seen not only as a productivity enhancing tool but also as an entertaining spice to the daily routine.

R. de Oliveira and H.V. da Rocha, propose multi-device design via the maintenance of a Consistency Priorities hierarchy defined in three levels. The first two levels give support to the user's expectation, while the third level provides task personalization.

M. Ferre et al. introduce MasterFinger-2, a novel two-finger haptic interface that improves haptic interaction. The interface is based on an open architecture, which allows the control of each finger independently via Ethernet.

J. Park and S.H. Park address the issue of effective visual interfaces in vehicle information systems and aircraft cockpits. The aim is to reduce cognitive workload, which is paramount to safety in vehicular and avionic applications.

D. Shastri et al. describe a new contact-free methodology to measure the cognitive load arising when the vehicle driver speaks at the same time over the cell phone. This technology opens the way for objective usability measurements in vehicular applications and more.

S. Rangarajan et al. describe a novel multimodal system combining a pressure sensing floor and a visual motion capture system.

X. Ren et al. introduces the Adaptive Hybrid Cursor technique that takes advantage of pressure-sensitive input devices. In fact, pressure is used to control the zoom ratio of interface contents.

U. Seifert and J.H. Kim delve into cognitive musicology as a novel approach to human-robot interaction in artistic contexts.

B. Sener and O. Pedgley address the inadequacies of 3D CAD systems in the early (and most creative) stages of industrial design.

Y. Lu and S. Smith present a new type of e-commerce system, AR e-commerce, which visually brings virtual products into real physical environments for user interaction.

A. Song, on one chapter, focuses on the design of multi-dimensional force sensors for haptic human- computer interaction. He pays particular attention on the design principles of a novel 4 DOF force/torque sensor. On another chapter, he focuses on the design of soft haptic display devices for human-computer interaction.

F. Steinicke et al. introduce new collaborative 3D user interface concepts for everyday working environments. They describe a system that allows displaying and interacting with both mono- as well as stereoscopic content in parallel.

Y. Suh et al. present a Context-Aware Augmented Reality (CAMAR) system that supports two main functionalities. One is the intuitive and personalized control of smart appliances. The other is enabling media contents to be shared selectively an interactively among a group of people.

J. Takatalo et al. concentrate on the psychological analysis of the user experience in digital games. They present three different psychological frameworks that have been used in the study of complex environments.

E.V. Mora et al. introduce a new model to improve the quality of life of people who live with chronic diseases following Ambient Intelligence principles. The model is validated in a solution to assess heart failure patients remotely.

T. Yamaguchi et al. present "SharedWell," a collaborative public and private, interactive display, suited for strategic cooperative tasks. The display enables users to dynamically choose negotiation partners, create cooperative relationships and strategically control the information they share and conceal.

Editor
Ioannis Pavlidis
Eckhard Pfeiffer Professor
Director, Computational Physiology Lab
Department of Computer Science
Email: ipavlidis@uh.edu
Houston, TX
USA

Contents

1

Using Zooming Applications for Mobile Devices

Kwang B. Lee
Department of Computer Science
University of Northern Virginia
U.S.A.

1. Introduction

Although mobile devices, cellular phones, Personal Digital Assistants (PDAs) and other handheld devices have hardware constraints such as small screen, small storage, low CPU speed, and low resolution, their popularity has been enormously increasing. This is enhancing the opportunities for researchers to overcome those open problems caused of small size hardware constraints. However, still the biggest difficulty is that these devices are too difficult to load today's visual information because most current information is based on a large visual documentation. How do we present information effectively on mobile devices? This is a main challenge for mobile interface developers since viewing is becoming more and more crucial point in our daily lives.

One of the methods is to build a Zoomable User Interfaces (ZUIs) by using several zooming tools so that the amount of information, which needs to be handled by users, can be shown on a small screen unlimitedly. Smooth zooming technology helps users to interact with their sense of the focus-and-context by shifting the cognitive load to the human visual system, and it can provide a possible solution that satisfies the above demands by means of increasing the effectiveness for using the small screen. Thus supporting zooming tools on mobile devices is a necessary item with enabling users to effectively control the zooming methods.

The goal of this paper is to increase the performance of user interfaces by developing zooming tools on mobile devices. Three zooming approaches will be introduced in this paper. First, focus zooming tools, which consists of the magnifying glass that was introduced from a "Bifocal Display" (Apperley, Tzavaras, and Spence, 1982), the gray scaling and blurring lens that was introduced from a "Focus+Context Visualization" (Giller, Tscheligim, Schrammel, Fröhlich, and Rabl, 2001), will be proposed. Second, file zooming tools including zoom-in and zoom-out functions to enlarge or reduce data and images based on the geometric zooming technology will be proposed. Finally, search zooming tools, which have two functions support a popup zooming and a shadow zooming functions to assist user easy to control for seeing many files on the device, will be introduced. Furthermore, the paper addresses a new usability testing method which combines heuristic, scenarios, and questionnaire approaches in order to effectively take experimental results from users. Its testing methods and procedures will be introduced by conducting usability test with user.

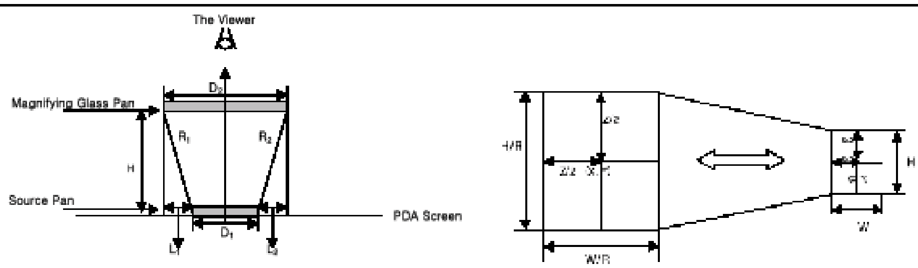
In this paper, we first describe the basic zooming technique and prototype on a PDA in section 2. In section 3, we introduce a new mobile usability testing method, and conduct usability testing and show the results in section 4. Finally, we conclude by describing some of our experiences in building the system and outlining future work.

2. Basic Zooming Structure

In this section, we discuss the theoretical background for basic zooming techniques on mobile devices, and mention the necessary concept needed to support our approaches.

2.1 Magnifying Process

The magnifying processes transfer pixels from a specified source pan to a specified destination pan, which is the magnifying glass, altering the pixels according to the selected raster operation code. To magnify data, a source pan containing source data on the original screen is smaller than the magnifying glass pan, as seen in Figure 1 (a). Thus, the ratio to magnify data will be decided by comparing the size of two pans, a source pan and a magnifying glass pan.



D_1: The size of source pan, ($x_1 \times y_1$)
D_2: The size of the magnifying glass pan, ($x_2 \times y_2$)
R_1, R_2: Magnifying ratio
L_1, L_2: Data loss area
H: Height

(a) Magnifying process (b) Zoom-in and zoom-out process

Fig. 1. Zooming process on a PDA screen

A degree of interest (DOI) function map which was introduced by Furnas (Furnas, 1986) indicates the level of interest in the item and it can be used to control how available space is allocated amongst items. The DOI concept has been changed over time according to the change of user interaction such as searching process and the moving focus. So, instead of using DOI, we suggest another approach to calculate how much data is magnified and how much data is lost.

In the magnifying glass, the technique allows user to determine the glass size and magnifying ratio based on the varying interest levels of parts. To calculate the magnifying ratio, we can use three factors: the source pan (D_1) that is based on the original source data; the data loss (L_1+L_2) that is an obscured region due to the big magnifying glass; and finally, the destination pan (D_2) that is the target window for copied the source data. Thus, the magnifying ratio depends on the size of each pan and data loss

2.2 Zooming Process

Zooming processes are copying bitmaps from the source rectangle and transferring them into a destination rectangle by stretching or compressing the size of bitmaps to fit the dimensions of the destination rectangle, if necessary. By the zooming size, S, and the zooming ratio, $R=\{0<\ldots\leq 1\}$, defined by the user, we are able to reduce the size by transforming a bitmap image into the zooming size. The PDA screen address, $P=(X, Y, Z)$, has both a location and a scale defined by the rectangle size, $Z=(S/R)$, is defined by the linear transformation, T_p: $(X-(Z/2), Y-(Z/2)) \leftrightarrow (X-(S/2), Y-(S/2))$. A zooming region, $A=[P, W, H]$, is a rectangle defined by an address together with a pixel width and height (W, H), as seen in Figure 1 (b).

The other level of zooming applications is to visibly display windows to the users as a popup or shadow zooming style. Every display window in the popup or shadow zooming applications has a region, $[P_i, W_i, H_i]$ where i is window number, which is the portion of the PDA screen, and these windows are located behind the original window by the fingernail-viewing file or the icon-viewing file. In particular, the shadow zooming has another window area, $[P_{i+1}, W_{i+1}, H_{i+1}]$, which is the small magnifying glass to magnify the hidden data instead of showing all data. Here we summarize the properties of zooming methods as follows:

- Visibility window: The visibility range of objects for user.
- Background window: The range of popup or shadow viewing objects which include the source image copied.
- Magnifying window: The glass to magnify data should have a certain range of magnification that allows users to see a small part in which they are interested.

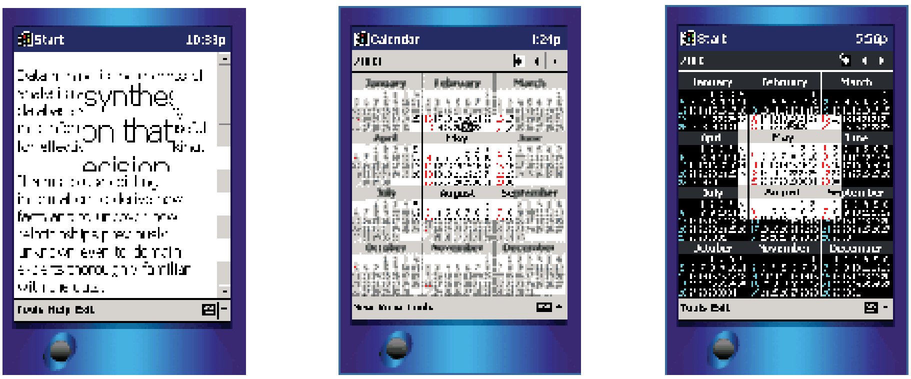
2.3 Basic Structures and Prototypes of Zooming Tools

In this part, we design and implement various zooming tools we mentioned with focusing on their usefulness and extensionality on a PDA, as seen Figures 2 - 5. Those tools were written in embedded Visual C++ supported by Microsoft® and developed on common Pocket PC.

2.3.1 Focus Zooming Tools

We introduce the focus zoom mechanism for increasing users' focusing ability in term of two facts. One is that the human transition is based on focusing on a magnified moving object according to the human perceptual system. The other is that the human eye is used to ignoring blurred objects because the eye also has a limited depth of filed by blurring currently irrelevant objects (Giller, Tscheligim, Schrammel, Fröhlich, and Rabl, 2001). When these tools are moving on the screen, their movements are represented with interactive magnification or the amount of blur, so that the user easily focuses on what is being displayed. These focus zooming tools give users more detail of certain parts of the screen, which is particularly helpful when lots of data is showing on the device.

As seen Figure 2, the focusing glass, (x_n, y_n) rectangle, shows a detailed view as a 2D form, and the other regions remain de-magnified or blurring areas according to the physical position. Cutting, pasting, and blurring various sections of bitmaps in real time will generate this focusing area. Even though the focusing glass does not provide good spatial continuity between regions, it is relatively simple in terms of using and implementation, and also it does not need much of the computational power of system memory. Therefore, focus zooming tools will be a good viewing technique for relatively low memory and mobile devices.



(a) Magnifying glass (b) Blurring lens (c) Gray scaling lens

Fig. 2. Focus zooming tool processes on a PDA

2.3.2 File Zooming Tools

Zooming methods can be categorized by the following presentation techniques: distorted and non-distorted zooming, geometric zooming, and semantic zooming methods (Furnas & Bederson, 1995), (Leung & Apperley, 1999). While the semantic zooming changes the shape or context in which the information is being presented, the geometric zooming has a scale operator to perform a geometric transformation, which can be used to shrink or magnify the size of an image. Each method will be adapted for making an efficient view for a user by considering the characteristics of hardware, software, and necessary environment.

In Figure 3, the file zoom has zoom-in and zoom-out methods in mobile devices based on this geometric method, which allows the user to specify the scale of magnification to increase or decrease the image or display screen. This shrunken file can be saved on the mobile device without changing the original content of the file. To expand its contents, the user touches the icon or the small image. These two viewing methods represent how to save files into a database, and also efficiently retrieve files from the database in mobile devices.

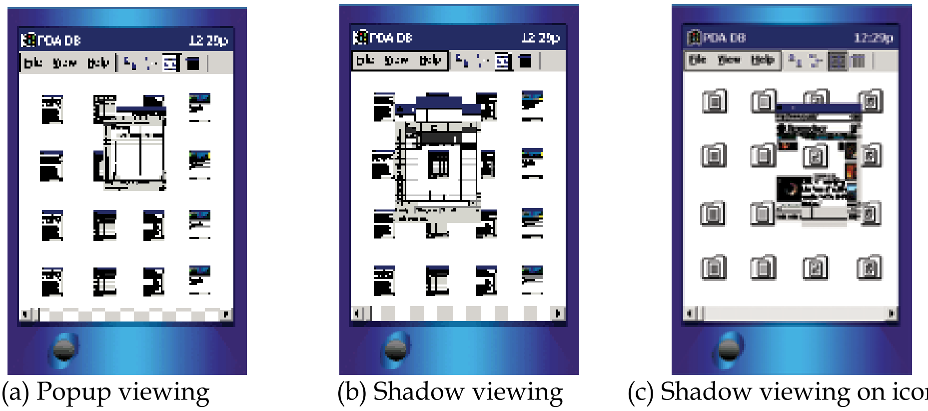
The first method is to save a file as a fingernail view by using a geometric zooming method. The second method, an icon view, uses the semantic method, so a certain icon can cover the small zoom-out file. Both are useful ideas for making user interaction with a database on the mobile device easier by saving screen space, and also providing visual abstractions to the user for what kinds of files are saved on the database. Moreover, the method should be used for memory buttons that are needed in graphic interaction in the small screen interface where the current file and its status can be saved as a small image or an iconic representation (Gheel & Anderson, 1999). Thus, if users desire to access the previous file state or browser, just handling the graphical memory button will bring up the previous file or browser status

(a) Zoom-out process (b) Fingernail viewing files (c) Icon viewing files

Fig. 3. File zooming tool processes on a PDA

2.3.3 Search Zooming Tools

This section introduces the search zoom implementation using a popup and a shadow zooming methods to make another viewing tool of the above zoom-out files. These two viewing tools will be used for searching and retrieving files in a database of mobile devices according to users' preferences.



(a) Popup viewing (b) Shadow viewing (c) Shadow viewing on icons

Fig. 4. Search zooming tool processes on a PDA

As seen in Figure 4 (a), the first method is a popup viewing tool which allows for the touching of the area of the original zoom-out files, and then a bigger zooming window is shown as a popup style. The popup zooming window promptly fades away when the user's attraction is moved to another place. So, the user easily knows what file or data included in the zoom-out file. In Figure 4 (b) and (c), the other method is a shadow viewing tool. When a user has located the point of interest with a small magnifying glass, which is embedded, the magnifying glass reveals the content of the file as a background, like a shadow. If the user touches the area of the file with the magnifying glass, then the embedded background of the

zoom-out file will be shown. The more the magnifying glass moves on the area, the more background image appears. When the user's attention moves to another place, then the zooming promptly fades away.
In this way, the user easily knows what files or data are saved in the original file. Here, we summarized two types of methods as follows:

- Popup Style: This shows the overview file information as a thumbnail image size like a popup menu when a user points to the file location.
- Shadow Style: This shows the overview file information with the magnifying glass, where the magnifying caption views the content of a file.

Both techniques will be potentially powerful tools by saving the searching time to see a full text or image in the mobile device database. We expect that great research efforts should be focused on exploring the application for searching methods and building databases with these applications on mobile devices

3. Key Components for Usability Test on Mobile Devices

In this section, we discuss a usability testing method we designed, and discuss how to build the testing plan for mobile devices.

1. Preparing Guidelines: In doing mobile application evaluation, well-defined guidelines enable the developers to easily assist participants for operating tools when they have problems caused by unstable prototypes and limited domain expertise.

Function Level	Given Tasks
Focus Zoom	• Task 1 – Use magnifying lens to see the interesting content of a file • Task 2 – Use blurring lens to see the interesting content of a file
File Zoom	• Task 3 – Use zooming operation on a file • Task 4 – Save a file as Fingernail view by zooming out function • Task 5 – Save a file as Icon view by zooming out function
Search Zoom	• Task 7 – Search a file by using popup viewing on Icon based PDA database • Task 6 – Search a file by using popup viewing on Fingernail based PDA database • Task 8 – Search a file by using shadow viewing on Fingernail based PDA database • Task 9 – Search a file by using shadow viewing on Icon based on PDA database

Table 1. Given tasks to participants for usability test

2. Developing Prototype: To build final tools much faster and much more cheaply in mobile devices, we use a prototype on a PDA. In many cases, using the prototype reduces the likelihood that people have erroneous communications about the product, and increases better understanding of the user interface designed.

3. Making Scenario-Based Tasks: Scenarios describe tasks in a way that takes some of the artificially out of the test such as explaining situations and environments, and also they can have an encapsulated description of an individual user or group using a specific set of computer facilities to achieve a specific outcome under specified circumstances over a certain time interval. Thus, the scenario-based task is a necessary approach for building mobility tasks to simulate certain mobile environments
4. Preparing Presentation: We prepare enough presentation time for introducing each developing function or interface which includes a step-by-step description of how to manipulate functions and complete given tasks. By using a checklist, we prepare detailed examples and steps for inexperienced participants.
5. Conducting Test: In conducting the testing itself on mobile devices, we have a time to interact with the participants without expressing any personal opinions or indicating whether the user is doing well or poorly. Users are interacting physically and emotionally with the prototype.
6. Debriefing Session and Post Testing: After all evaluations have been completed, we should prepare a meeting in which all participants come together to discuss their impressions from the test. This session is conducted primarily in a brainstorming mode and focused on discussions of the major usability problems and general problematic aspects of the design on mobile devices.

4. Experimental Results

4.1 Conduct Usability Test

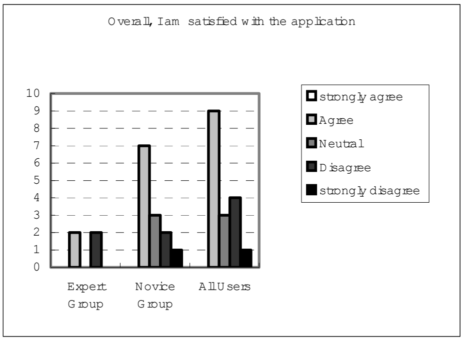
To achieve usability test under mobile environments, we design a combination method prescribed in the previous section, and then conduct usability test for zooming tools using this method (Lee & Grice, 2004), (Johnson, 1998), (Lindroth & Nilsson, 2001). This combined testing method include heuristic evaluation, questionnaires and scenario-based tasks, and it consists of six attributes such as overall impression, success and timing, satisfaction and preferences, feature use and understanding, compatibility and interoperability and functionality.

To conduct the test, we recruited 17 students who were students, and then classified them into two groups based on pre-screening testing results:

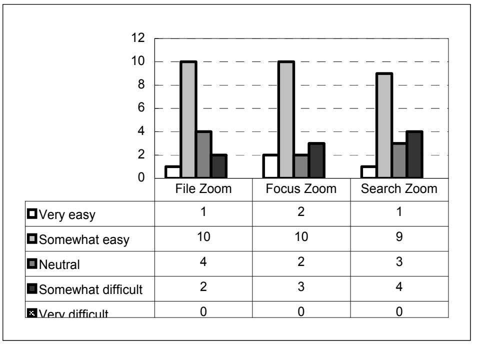
- Expert Group: The group members have substantive knowledge regarding mobile devices, and they are familiar with both HCI and usability. Also, they have fundamental knowledge of mobile devices and personal computers. We choose four students from all participants who have a history of operating mobile devices.
- Novice Group: The group members lack substantive knowledge regarding mobile devices, but they are reasonably familiar with HCI and usability. We recruited 13 students who do not have any experience with mobile devices; however they can manipulate personal computers

All testing was conducted in the class lab, and individual users were asked to inspect the prototype alone. After each user tested the prototype and completed given tasks at least two times, the user started to answer questions, which were based on dialogue elements based on heuristic categories. There were nine tasks that included a proper scenario to help users to evaluate the prototype well, as seen Table 1.

We evaluated the users' satisfaction and preference for how the prototype solves physical limitations, provides an easy way to use zooming tools, and supports feedback for increasing users' interaction. Usually we focus on job completion time: rapid, accurate and completeness of task processing for each main task because it is important that the application is working rapidly and accurately as the user request. Generally, all users agree that the application allows for rapid, accurate and complete task processing.



(a) Overall satisfaction



(b) Post-test evaluation

Fig. 5. Usability testing result

As seen Figure 5 (a), in testing users' overall satisfaction with the zooming tools, we find that the expert group has split opinions, both agreeing and disagreeing. One reason is based on the expert group's experience, because they may have expected a desktop quality application on a mobile device. However, the actual zooming functions on mobile devices are of much lower quality. A second reason is the choice of sample files, graphics and figures in the test may have highlighted the low quality more than the text files, so we chose non-suitable files for the test. The novice group however is satisfied with this application.

In Figure 5 (b), even though some users disagree with each approach, typically more than 65% of the users are satisfied with each function, the focus zoom, the file zoom, and the search zoom. The preference rates of the focus zoom and the file zoom is bigger than the search zoom method. However, many users answered "neutral" and "somewhat difficult" because of unclear terminology and non-suitable tasks. Therefore, to increase usability, we have to find proper terminology and modify the application according to the results when we redesign the product.

4.2 References and Recommendations

In order to ameliorate the most pressing usability problems with the zooming tool by considering the global problems, we describes each group's preference and recommend future changes for developers as follows:

- Expert User Preferences: Usually, expert users want fast and accurate functions to complete tasks, and they need good feedback from the tool. Also, they want to modify the tools to be more compatible with other tools, and feel the difficulty of handling several functions. Finally, expert users need clearly defined instruction to properly use them.

- Novice User Preferences: Many novice users are satisfied with the tool more than expert users. They think this approach is a very useful tool for solving small devices' problems, however they have great difficulty with the tools' feedback, and they want to easily access and exit each function. They think the tool needs more compatibility working with other programs and exchanging information

Here, we summarize user recommendations for redesigning the product.

1. Reducing many clicking steps: The tools ask users to click the pen several times to operate the menu. This procedure might be awkward for users.
2. Making well-organized menu interfaces and more functions: Preparing useful functions and constructing well-organized menus are critical points to increase usability.
3. Trying to develop other uses of the zooming function: Developers should try to develop other uses of the zooming function and to find kinds of tasks/areas to which it would be useful.
4. Drawing borderline when zooming functions are working on the screen: All zooming windows should have a borderline because users cannot easily recognize which parts are zoomed.

5. Conclusion

This paper has described specialized zooming tools on mobile devices with designing and developing basic geometric and semantic zooming methods in order to increase the usability of the device. Based on three zooming methods, we created new zooming tools for the device by describing the detail prototype and implementation of these applications were introduced in this paper.

However, we have problems because the tools were designed for the PDA simulation program in a desktop computer. So, we do not know how many different results exist between the application program and the real physical device. Especially, in terms of hardware, the current PDA does not have enough pixels, so users could potentially encounter broken characters when the magnifying glass is used.

Although these zooming tools are not substantially implemented in commercial PDAs, it will be used for new interfaces in mobile devices by supporting various zooming functions. We look forward to continuing the research and development of this tool according to the future development of the PDA's hardware performance and elements. Therefore, the biggest contribution of this paper is the creation of zooming tools on PDAs by encouraging the development of practical zooming methods over theoretical methods.

6. References

Apperley, M. D., Tzavaras, I. and Spence, R. (1982). A Bifocal Display Technique for Data Presentation. *In Proceedings of Eurographics '82*, pp. 27-43

Bartram, L., Ho, A., Dill, J. and Henigman, F. (1995). The Continuous Zoom: A Constrained Fisheye Technique for Viewing and Navigating Large Information Spaces. *Proceedings of the 8th annual ACM symposium on User Interface and Software Technology*, pp. 207-215, Pittsburgh PA

Furnas, G. W. (1986). Generalized Fisheye Views. *In Proceedings of ACM SIGCHI '86*, pp. 12-16, ACM Press

Furnas, G. W. and Bederson, B. B. (1995). Space-Scale Diagrams: Understanding Multiscale Interfaces. *In Proceedings of the SIGCHI conference on Human Factors in Computing Systems CHI '95*, pp. 231-241, ACM Press

Gheel, J. and Anderson, T. (1999). Data and Metadata for Finding and Reminding. Proceeding of the 1999 *International Conference on Information Visualization 4th*, pp. 446-451, Washington DC

Giller, V., Tscheligim, M., Schrammel, J., Fröhlich, P. and Rabl, B. (2001). *Experimental Evaluation of Semantic Depth of Field, a Preattentive Method for Focus+Context Visualization.* Technical Paper TR-2001-3, Vienna University of Technology

Johnson, P. (1998). Usability and Mobility; Interactions on the move. *In Proceedings of the First Workshop on Human Computer Interaction with Mobile Devices*

Khella. A. and Bederson, B.B. (2004). *Pocket PhotoMesa: A Zoomable Image Browser for PDAs.* MUM 2004, pp. 19-24, College Park MD

Lee, K. B. and Grice, R. (2004). Developing a New Usability Testing Method for Mobile Devices. IPCC 2004, *IEEE Professional Communication Society*, pp. 115~ 127, Minneapolis MN

Lee, K. B. and Grice, R. (2003). The Embedded Zooming Application for Personal Digital Assistant. IPCC 2003, *IEEE Professional Communication Society*, pp 109-116, Orland FL

Leung, Y. K. and Apperley, M. D. (1999). Readings in Information Visualization Using Vision to Think. *A Review and Taxonomy of Distortion-Oriented Presentation Techniques,* pp. 350-367, Morgan Kaufman Publishers, Inc

Lindroth, T. and Nilsson, S. (2001). Context Usability, *Rigour meets relevance when usability goes mobile*, pp. 24-26, ECIS Doctoral Consortium

2

The effects of panel location, target size, and gender on efficiency in simple direct manipulation tasks

Rafal Michalski
Wroclaw University of Technology
Poland

1. Introduction

In recent years there was a big increase in the number of personal computers used worldwide. At the same time a lot of research have been conducted on usability of more and more sophisticated interactive systems. On the other hand some (e.g. Whittaker et al., 2000) have argued that too much focus was given to modern styles of human-computer interaction, since the usefulness of these proposals is very limited (Hartson, 1998). Simultaneously, the in-depth exploration of standard means of communication between human beings and computer programs are very often neglected.
The presented research involves the 'search and click' technique, which is a core component of a direct manipulation style of human-computer interaction (Shneiderman, 1982, 1983). Although currently there exist many other methods, the direct manipulation is still one of the most popular, especially among graphical interfaces. The study described in this publication may be situated in the trend of research related both to visual search and visually controlled motor activity (compare Grobelny et al., 2005; Michalski et al., 2006; Michalski & Grobelny, 2008). This area is a combination of the traditional Fitts' approach (Fitts, 1954; Fitts & Peterson, 1964), in which only the movement time related to graphical object selection is taken into account, and the situation where the time of visual search for a particular target among the group of distractors is of a main concern. The rationale of including these two activities simultaneously in the experimental setup arise from the observations presented in the work of Hoffmann & Lim (1997). The researchers argue that concurrent decision and movement tasks are complex, and they should not be analysed separately. Their suggestions were backed up by experimental results. Additionally, there are some evidence at the neural level (Wurtz et al., 1982; Kustov & Robinson, 1996; Colby & Goldberg, 1999) suggesting that a manual response to a stimulus may influence the cognitive processes.
This study is mainly focused on the problem of graphical panel position on the screen and its impact on the accomplishment of simple 'search and click' tasks. Despite some previous works dealing with this subject, there are still several issues that need to be addressed. The earlier research results are not always consistent. Let us take for instance locations of web site menus. McCarthy et al. (2003) demonstrated that the left menu location is faster

searched, but if the user performed another visual search task in the same web page, this advantage was not observed. In the study of Kalbach & Bosenick (2004), the menu location factor did not significantly influence the mean acquisition times either. The inconsistencies also exist, when the visual search of simple graphical objects is concerned (Michalski et al., 2006).

The prior studies in the HCI field discussed mostly left and right upper corner locations (McCarthy et al., 2003; Kalbach & Bosenick, 2004; Michalski et al., 2006) and other positions were rarely examined. Among the works related to other than left and right screen locations of searched targets there are investigations of Campbell & Maglio (1999), Schaik & Ling (2001), and Pearson & Schaik (2003). The study of Campbell & Maglio (1999) demonstrated that the shortest mean response times were observed for the stimuli placed in the upper left corner of the screen, and the longest for targets in the lower right corner. Schaik & Ling (2001) in their investigation showed that menus having the same contrast were operated the slowest in the bottom position, and that the reaction times for right located targets were significantly slower than in the case of left and top positions. Later in a quite similar paper, Pearson & Schaik (2003) obtained similar selection times both for left and right menus as well as for top and bottom ones. The further analysis showed also, that there was meaningful difference between grouped results for left and right locations and grouped top and bottom. The side positioned menus occurred to be worse in terms of the selection speed than both top and bottom layouts.

The other area of interest discussed in the current research concerns possible differences between male and female computer users in executing simple direct manipulation tasks that require some cognitive effort. Gender differences in performing various types of cognitive task have been a topic of multiple studies in the psychology and neuropsychology fields (e.g. Harasty et al., 1997; Adam et al., 1999; Gur et al., 1999; Weiss et al., 2003; Blatter et al., 2006; Reimers & Maylor, 2006; Roalf et al., 2006; Walhovd & Fjell, 2007). It is generally accepted that men do better in spatial and mathematical tasks, whereas women have better verbal ability (MacCoby & Jacklin, 1974). However, the latest research and meta analyses of previous papers suggest these differences to be less salient than in the past (Hyde & McKinley, 1997; Jorm et al., 2004).

When the discrepancies in accomplishing simple pointing activities are concerned, it is assumed that they are a result of different strategies used by both sexes. According to this approach, women perform better when the accuracy is analysed, while men are superior in tasks, where completion time is of a great concern (Ives et al., 1993; Peters & Campagnaro, 1996; Warshawsky-Livne & Shinar, 2002; Barral & Debû 2004; Rohr, 2006a, 2006b). As it was outlined above, there has been a significant amount of research regarding gender differences in performing cognitive and motor tasks separately, however the studies treating these two conditions simultaneously are hardly to find.

The following sections describe a laboratory experiment that was designed and conducted to cast more light on the aforementioned matters. More specifically, this paper in an attempt to explain how square panel locations along with two panel item sizes affect the speed of executing simple search and click tasks. In addition, differences in task performance between sexes are examined. The obtained results are analysed and compared with the outcomes of previous studies. Limitations of this research as well as possible future works are also outlined.

2. Method

2.1 Participants

Overall, forty Wroclaw University of Technology students volunteered in the study. There was an equal number of male and female participants. The students were within the age range of 21–25 years, and they worked with computer programs on a daily basis. They reported having a normal or corrected to normal visual acuity.

2.2 Apparatus

A computer program written in a MS Visual Basic™ 6.0 environment was used to conduct the experiments. The research took place in teaching laboratories on uniform personal computers equipped with the same computer mice and 17″ monitors of the CRT type. The resolution was set at 1024 by 768 pixels and a typical (default) computer mouse parameters were used.

2.3 Experimental design

The graphical panels being investigated comprised of 36 buttons arranged in a square with 26 Latin alphabet characters and ten Arabic numbers placed on these buttons. Two independent variables were manipulated: the graphical object size and panel location on the screen. Two different, commonly occurring in up-to-date computer programs, panel item sizes were used in the experiments. The side square button sizes equalled to 22 (small) and 38 pixels (large). Bolded Times New Roman font types in sizes of 12, and 24 pt were employed. The distance between the user and the screen was set approximately at 50 cm, so the visual angles of these objects amounted to 0°41′, and 0°69′ respectively. The second factor was examined on four levels corresponding to the four corners of the computer screen. The panels were moved away from the screen edges by 18 pixels to minimize the effect of faster selection of items located at the screen borders (Farris et al., 2002, 2006; Jones et al., 2005).

The independent variables resulted in eight different experimental conditions: (two object sizes) × (four panel locations). A mixed model design was applied. The object size factor was treated within subjects whereas the other effect was examined between subjects. Each of the four groups of participants testing the four panel locations consisted of an equal number of males and females. The dependent variables being measured were the 'search and click' task completion time and the number of errors committed. The time was computed from when the START button was pressed, to when the object was clicked. The error occurred if a subject selected different than required graphical object.

2.4 Procedure

Before the examination the subjects were informed about a purpose and course of the experiment. The study started by filling out a general questionnaire concerned with personal data and computer literacy. Next, participants were asked to perform five attemptive trials. After the warm-up, the proper experiment took place. First, instruction dialogue window, presenting a START button and the target to be looked for, appeared. The searched layout was invisible at this instant. After the START button was clicked, the window disappeared and one of the examined panels was shown. The user was instructed to find as fast as possible the required object in the presented structure, and click it using a computer mouse. The instruction window was shown for each trial. The panels were displayed in a random order, different for every subject. Every student performed 10 trials for each of the examined configurations. Every 10 trials, an informative window including mean acquisition times and incorrect attempts was shown, and after clicking the OK button the examination was continued.

3. Results

3.1 Selection times

The subjects performed 800 trials altogether. The proper target item was localized and clicked in 781 cases. Excluding the error searches, the mean value amounted to 2248 ms with the standard deviation 1576 ms and mean standard error 56 ms. The median was equal to 1793 ms. The shortest selection time was 672 ms, whereas the longest – 14 591 ms. Both the skewness and the kurtosis were decidedly different than the values of these parameters characteristic of the normal distribution and amounted to 2.7 and 12 respectively. The basic descriptive statistics for all the examined conditioned (without the mistakes) are presented in table 1.

No.	Target size	Panel location	Gender	N	Median (ms)	Mean (ms)	SE (ms)	SD (ms)
1.	Small	Left-Bottom	Female	50	1547	1938	153	1079
2.	Small	Left-Bottom	Male	49	2193	2880	319	2235
3.	Small	Left-Top	Female	49	1938	2463	211	1478
4.	Small	Left-Top	Male	48	1933	2465	302	2092
5.	Small	Right-Bottom	Female	50	1838	2252	189	1340
6.	Small	Right-Bottom	Male	49	1843	2314	233	1633
7.	Small	Right-Top	Female	49	1893	2426	238	1667
8.	Small	Right-Top	Male	48	1793	1992	150	1039
9.	Large	Left-Bottom	Female	50	1406	1853	189	1337
10.	Large	Left-Bottom	Male	48	1787	2334	252	1748
11.	Large	Left-Top	Female	49	2294	2832	322	2256
12.	Large	Left-Top	Male	48	1728	2247	210	1454
13.	Large	Right-Bottom	Female	48	1577	2154	215	1486
14.	Large	Right-Bottom	Male	50	1507	1826	161	1139
15.	Large	Right-Top	Female	49	1412	1935	175	1226
16.	Large	Right-Top	Male	47	1913	2078	142	971

Table 1. Basic descriptive statistics for all examined conditions

The results regarding selection times were next analysed by means of the Generalized Linear Models (GZLM; Nelder & Wedderburn, 1972) under the assumption that the dependent variable has the inverse Gaussian (IG) distribution. These assumptions are reasonable in light of the results presented by Michalski (2005) and taking into account the dependent variable descriptive data calculated for the present study. A three way ANOVA based on the GZLM was used for examining the factors of the user gender, panel location, and target sizes.

Effect	df	Wald statistics (W)	p
Panel location (PLO)	3	9.6	*0.022
Item size (ISE)	1	3.9	*0.047
Gender (GEN)	1	0.15	0.70
PLO × ISE	3	2.2	0.54
PLO × GEN	3	16.5	*0.00089
ISE × GEN	1	0.899	0.34
PLO × ISE × GEN	3	5.9	0.12

* The results significant at a level 0.05

Table 2. GZLM analysis of variance results

The results of the analysis are presented in table 2 and showed that the panel location along with the item size factor are significant at the level of $a = 0.05$. The effect of gender alone occurred not to be meaningful, however there was a significant interaction between gender and panel location factors. All other interactions were irrelevant.

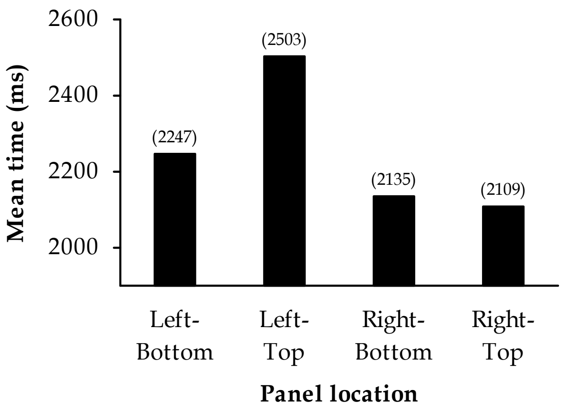


Fig. 1. Mean selection times depending on panel location (df = 3, W = 9.6, p = 0.022)

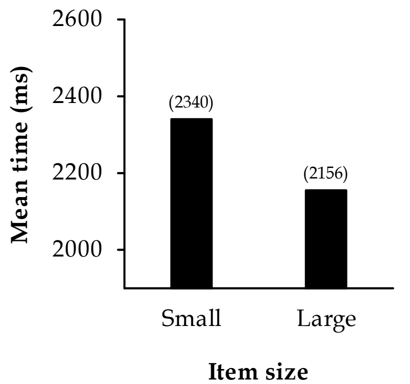


Fig. 2. Mean selection times depending on item size (df = 1, W = 3.9, p = 0.047)

The mean acquisition times along with other basic statistics related to the panel location are presented table 3 and illustrated in fig. 1. The layouts positioned on the right hand side of the computer screen, both top and bottom were operated the fastest, and the difference between their mean selection times were insignificant (df = 1, W = 0.049, p = 0.83). Among the structures located on the left, the bottom layouts were decidedly better ($\alpha = 0.1$) than the top ones (df = 1, W = 2.79, p = 0.095). The left top panel placement was the worst in terms of the selection speed, and the difference in average times between the best and the worst positions amounted to approximately 19% (394ms).

No.	Panel location	N	Median (ms)	Mean (ms)	SE (ms)	SD (ms)	Min (ms)	Max (ms)
1.	Left-Bottom	197	1734	2247	120	1691	672	12 148
2.	Left-Top	194	1930	2503	133	1853	688	14 591
3.	Right-Bottom	197	1673	2135	101	1411	701	8001
4.	Right-Top	193	1772	2109	91	1264	701	8062

Table 3. Results for the panel location factor (df = 3, W = 9.6, p = 0.022)

The graphical illustration of mean acquisition times computed for the target size effect is presented in fig. 2, and the descriptive statistics are put together in table 4. Mean times registered for panels consisting of large objects were substantially shorter than for their small counterparts. The discrepancy was equal 184 ms (8.5%).

No.	Item size	N	Median (ms)	Mean (ms)	SE (ms)	SD (ms)	Min (ms)	Max (ms)
1.	Small	392	1903	2340	82	1629	721	14 591
2.	Large	389	1656	2156	77	1519	672	12 578

Table 4. Results for the item size factor (df = 1, W = 3.9, p = 0.047)

The GLZM analysis of variance revealed that there is an interaction between gender and panel location effects, so in fig. 3 and table 5 there are results presented separately for men and women taking part in the examination.

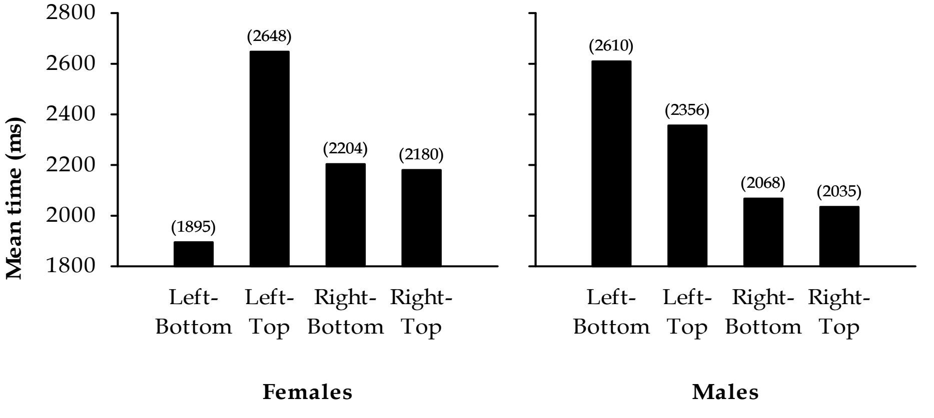


Fig. 3. Mean selection times depending on gender and panel location (df = 1, W = 16.5, p = 0.00089)

The mean operation times for panels on right side of the screen were similar both for women and men as well as for top and bottom positions of these graphical structures. For layouts located on the left hand side of the monitor, females generally outperformed males (df = 1, W = 2.6, p = 0.1097) and left bottom panels were operated faster than left top configurations (df = 1, W = 3.3, p = 0.068).

No.	Gender	Panel location	N	Median (ms)	Mean (ms)	SE (ms)	SD (ms)	Min (ms)	Max (ms)
1.	Female	Left-Bottom	100	1477	1895	121	1209	672	8406
2.	Female	Left-Top	98	2048	2648	193	1906	711	12 578
3.	Female	Right-Bottom	98	1678	2204	142	1407	701	7210
4.	Female	Right-Top	98	1593	2180	149	1476	701	8062
5.	Male	Left-Bottom	97	2062	2610	205	2017	681	12 148
6.	Male	Left-Top	96	1797	2356	183	1795	688	14 591
7.	Male	Right-Bottom	99	1622	2068	143	1419	741	8001
8.	Male	Right-Top	95	1843	2035	103	1002	731	6630

Table 5. Results for the interaction between gender and panel location (df = 1, W = 16.5, p = 0.00089)

However, women had shorter mean selection times for left bottom structures than for left top ones, whereas men did better with left top panels than with left bottom configurations. This interaction between gender and left panel locations was also statistically significant (df = 1, W = 11.8, p = 0.000589).

3.2 Errors

A total of 19 errors were made by participants, which accounts for 2.4% of all performed trials. The percentages of mistakes registered during the examination are put together in table 6. They are broken down by the examined factors.

Factor	Errors (%)
Panel location	
Left-Bottom	1.5
Left-Top	1.5
Right-Bottom	3.5
Right-Top	3.0
Item size	
Small	2.0
Large	2.8
Gender	
Female	1.5
Male	3.3

Table 6. Percentages of wrong selections

Factor	df	χ^2	p
Panel location	3	2.75	0.43
Item size	1	0.49	0.49
Gender	1	2.64	0.104

Table 7. Analysis of differences in the number of errors for examined factors

A nonparametric, Chi-square test was employed to verify the significance of differences in the number of wrong selections for the examined factors. The results of these analyses are presented in table 7. The only meaningful difference in the number of wrong selections was observed for the gender factor. The significance level $\alpha = 0.10$ was slightly exceeded in this case. Women committed decidedly less errors (1.5%) than men did (3.3%). The other two effects were irrelevant.

4. Discussion

Generally, the obtained results showed that the panel location and target item size factors considerably influenced the mean acquisition times. The gender effect was not meaningful, but the interaction between the gender and panel location was statistically significant.

4.1 Panel location

The results showed that the panel location factor considerably influenced the acquisition times. This outcome is generally consistent with the works of Campbell & Maglio (1999), Schaik & Ling (2001), McCarthy et al. (2003), Pearson & Schaik (2003), and Michalski et al. (2006), where the stimuli position, one way or another, significantly influenced the response time. However, this result contradicts with the investigation presented by Kalbach & Bosenick (2004), in which they did not observe the significant influence of the location factor. From among the aforementioned studies, the target locations used by Campbell & Maglio (1999) were most similar to those employed in the described in this chapter experiment. Although, the location factor was significant in their experiment, the detailed results was contradictory with our findings. They explained the results by the nature of the stimulus, which was treated by participants as text to be read. In this paper experiments, it is hardly to associate the outcome to the reading habits, so maybe some other factors come into play. Possibly the obtained results were to some extent influenced by different ways of searching the target by men and women that manifested itself as the statistically significant interaction between location and gender factors. Of course, the discrepancies could have been caused also by a number of other issues including the different type target, screen resolutions, size of the screen, stimuli sizes, as well as the number of distractors and their arrangement.

4.2 Target size

The target object size effect was statistically meaningful. In the case of simple selection tasks where the target is constantly visible to the subject, the Fitts' law (Fitts, 1954; Fitts & Peterson, 1964) applies. According to this well known formula, the movement time is affected by the object size along with the movement amplitude. However, the presented study involves additionally the search process, which may last decidedly longer than the time needed for reaching and clicking the target. In such a case, the Fitts' law may not be relevant. Nevertheless, some recent findings proved that bigger target objects shortened acquisition times (Michalski et al., 2006), and this finding was supported in the present investigation.

4.3 Gender differences

Though the effect of gender alone was not significant, the interaction between gender and panel location effects occurred to be meaningful. This relation was particularly visible for the panels positioned on the left hand side of the screen. Thus in general, the results support the hypothesis that there exists a significant difference between women and men in performing simple search and click tasks (at least for some locations). However, the obtained results seem a bit awkward and it is hardly to draw some reasonable conclusions. For instance, the better results for panels located in the left bottom in comparison with the left top corner obtained by females, could have been attributed to possible inappropriate chair seat height settings. But, if this was the case, so why were the women operation times for panel situated on the right hand side comparable? What is more, for the right placed panels men outperformed women in both bottom and top panel locations, so the interaction did not exist. Taking into consideration only the right locations, the present research outcomes to some extent support the assumption that men do better where the task accomplishment time is evaluated (Ives et al., 1993; Peters & Campagnaro, 1996; Warshawsky-Livne & Shinar, 2002; Barral & Debû 2004; Rohr, 2006a, 2006b). But the differences are not statistically significant ($df = 1$, $W = 1.34$, $p = 0.247$). In light of such inconsistent results, these issues require undoubtedly further more detailed research.

4.4 Incorrect selections

The error analysis proved that males were more prone to make mistakes than females ($\alpha = 0.10$), while the factors Panel location and Item size were irrelevant. The registered data confirm the suggestion that women put more attention to accuracy than male participants. The recorded mean error rate in this research (2.4%) was generally comparable to the values obtained in other research. For example, in the research of Schaik & Ling (2001), Pearson & Schaik (2003), Grobelny et al. (2005), Michalski et al. (2006), Michalski & Grobelny (2008), the mistakes occurred in less than 3% of all trials.

4.5 Limitations and possible future works

There is naturally a number of limitations related with this study. One of the most obvious weaknesses is the difficulty in interpreting especially those data, which are connected with the interaction between gender and the panel location. In light of these inconclusive results, additional studies seem to be necessary. Possibly, increasing the number of subjects or applying some eye tracking techniques would allow for more consistent conclusions.

It should also be stressed that almost all the participants were young and familiar with various computer programs, and were using computers on a daily basis, so their performance may substantially differ from the novice or elderly users. Additionally, the present investigation involved only one and very simple interaction technique, while the real interaction may require a combination of other ways of communicating with a computer. Also the choice of target icons may have an impact on the obtained results. Further research may include other graphical objects (e.g. icons from popular programs), different pointing devices, or subjective assessment of user preferences.

5. Conclusions

According to the obtained results during making the decisions about the design issues both the target size and location of a graphical panel should be considered. The obtained results also showed generally that in simple 'search and point' tasks, the gender factor should rather not be neglected. Although the influence seems not to be clear, the presented findings support the assumption of different ways of performing these kinds of tasks by men and women. As it was mentioned in the introductory section of this chapter, the obtained in this research differences may constitute a juxtaposition of the differences in performing the visually controlled motor tasks as well as discrepancies in executing cognitive tasks.

The presented research results enrich our knowledge in the area of simple pointing tasks combined with a visual search, and show the need for further studies concerned with the subject. However, because of some inconsistencies in the present and past research, one should be cautious in recommending any given design solution. In practice, decisions regarding the graphical features of toolbars should, obviously, take into account limitations of scientific investigations. Possibly, some additional research may be necessary to test the ecological validity of a particular proposal.

6. References

Adam, J.J., Paas, F.G.W.C., Buekers, M.J., Wuyts, I.J., Spijkers, W.A.C., & Wallmeyer, P. (1999). Gender differences in choice reaction time: evidence for differential strategies. *Ergonomics*, **42**(2), 327–335.

Barral, J., & Debû, B. (2004). Aiming in adults: sex and laterality effects. *Laterality*, **9**(3), 299–312.

Blatter, K., Graw, P., Munch, M., Knoblauch, V., Wirz-Justice, A., & Cajochen, C. (2006). Gender and age differences in psychomotor vigilance performance under differential sleep pressure conditions. *Behavioural Brain Research*, **168**(2), 312–317.

Campbell, C.S., & Maglio, P.P. (1999). Facilitating navigation in information spaces: Road-signs on the World Wide Web. *International Journal of Human-Computer Studies*, **50**, 309-327.

Colby, C.L., & Goldberg, M.E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, **22**, 319–349.

Farris, J.S., Johnson, B.R., & Jones, K.S. (2006). Width guidelines for rectangular objects with penetrable and impenetrable borders. *Behaviour & Information Technology*, **25**(1), 83–90.

Farris, J.S., Jones, K.S., & Anders, B.A. (2002). Factors affecting the usefulness of impenetrable interface element borders. *Human Factors*, **44**(4), 578–591.

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, **49**, 389–391.

Fitts, P.M., & Peterson, J.R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology*, **67**, 103–112.

Grobelny, J., Karwowski, W., & Drury, C. (2005). Usability of Graphical icons in the design of human-computer interfaces. *International Journal of Human-Computer Interaction*, **18**, 167–182.

Gur, R.C., Turetsky, B.I., Matsui, M., Yan, M., Bilker, W., Hughett, P., & Gur, R.E. (1999). Sex differences in brain gray and white matter in healthy young adults: correlations with cognitive performance. *Journal of Neuroscience*, **19**(10), 4065–4072.

Harasty, J., Double, K.L., Halliday, G.M., Kril, J.J., & McRitchie, D.A. (1997). Language-associated cortical regions are proportionally larger in the female brain. *Archives of Neurology*, **54**(2), 171–176.

Hartson, H.R. (1998). Human-computer interaction: Interdisciplinary roots and trends. *The Journal of Systems and Software*, **43**, 103–118.

Hoffmann, E.R., & Lim, J.T.A. (1997). Concurrent manual-decision tasks. *Ergonomics*, **40**, 293–318.

Hyde, J.S., & McKinley, N.M. (1997). Gender differences in cognition. results from meta-analyses. In P.J. Caplan, M. Crawford, J.S. Hyde, & J.T.E. Richardson (Eds.), Gender Differences in Human Cognition (pp. 30–51). New York: Oxford University Press.

Ives, J.C., Kroll, W.P., & Bultman, L.L. (1993). Rapid movement kinematic and electromyographic control characteristics in males and females. *Research Quarterly for Exercise and Sport*, **64**(3), 274–283.

Jones, K.S., Johnson, B.R., & Farris, J.S. (2005). GUI objects with impenetrable borders: Instruction (not practice) makes perfect. *International Journal of Human-Computer Studies*, **62**(6), 687–712.

Jorm, A.F., Anstey, K.J., Christensen, H., & Rodgers, B. (2004). Gender differences in cognitive abilities: The mediating role of health state and health habits. *Intelligence*, **32**(1), 7–23.

Kalbach, J., & Bosenick, T. (2003). Web page layout: A comparison between left- and right-justified site navigation menus. *Journal of Digital Information*, **4**, Article No. 153, 2003-04-28. Retrieved from http://jodi.tamu.edu/Articles/v04/i01/Kalbach/ on November 20, 2007.

Kustov, A.A., & Robinson, D.L. (1996). Shared neural control of attentional shifts and eye movements. *Nature*, **384**, 74–77.

MacCoby, E., & Jacklin, C. (1974). *The psychology of sex differences*, Palo Alto: Stanford University Press.

McCarthy, J.D., Sasse, M.A., & Riegelsberger, J. (2003). Could I have the menu please? An eye tracking study of design conventions. In Proceedings of HCI2003, Bath, UK, pp. 401–414.

Michalski, R. (2005). Komputerowe wspomaganie badań jakości ergonomicznej oprogramowania (Computer-aided research of software ergonomic quality). PhD thesis, Wroclaw University of Technology.

Michalski, R., Grobelny, & J., Karwowski, W. (2006). The effects of graphical interface design characteristics on human-computer interaction task efficiency. *International Journal of Industrial Ergonomics*, **36**, 959–977.

Michalski, R. & Grobelny, J. (2008), The role of colour preattentive processing in human-computer interaction task efficiency: a preliminary study, *International Journal of Industrial Ergonomics*, **38**(3/4), 321–332.

Nelder, J.A., & Wedderburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society A*, **135**, 370–384.

Pearson, R., & Schaik, P. (2003). The effect of spatial layout of and link color in web pages on performance in a visual search task and an interactive search task. *International Journal of Human-Computer Studies*, **59**, 327–353.

Peters, M., & Campagnaro, P. (1996). Do women really excel over men in manual dexterity? *Journal of Experimental Psychology/Human Perception & Performance*, **22**(5), 1107–1112.

Reimers, S., & Maylor, E. (2006). Gender Effects on Reaction Time Variability and Trial-to-Trial Performance: Reply to Deary and Der (2005). *Aging, Neuropsychology & Cognition*, **13**(3/4), 479–489.

Roalf, D., Lowery, N., & Turetsky, B.I. (2006). Behavioral and physiological findings of gender differences in global-local visual processing. *Brain and Cognition*, **60**(1), 32–42.

Rohr, L.E. (2006a). Upper and lower limb reciprocal tapping: evidence for gender biases. *Journal of Motor Behavior*, **38**(1), 15–17.

Rohr, L.E. (2006b). Gender-specific movement strategies using a computer-pointing task. *Journal of Motor Behavior*, **38**(6), 431–437.

Schaik, P., & Ling, J. (2001). The effects of frame layout and differential background contrast on visual search performance in web pages. *Interacting with Computers*, **13**, 513–525.

Shneiderman, B. (1982). The future of interactive systems and the emergence of direct manipulation. *Behaviour and Information Technology*, **1**, 237–256.

Shneiderman, B. (1983). Direct manipulation. A step beyond programming languages. *IEE Computer*, **16**, 57–69.

Walhovd, K.B., & Fjell, A.M. (2007). White matter volume predicts reaction time instability. *Neuropsychologia*, **45**(10), 2277–2284.

Warshawsky-Livne, L., & Shinar, D. (2002). Effects of uncertainty, transmission type, driver age and gender on brake reaction and movement time, *Journal of Safety Research*, **33**(1), 117–128.

Weiss, E., Siedentopf, C. M., Hofer, A., Deisenhammer, E. A., Hoptman, M. J., Kremser, C., Golaszewski, S., Felber, S., Fleischhacker, W.W., & Delazer, M. (2003). Sex differences in brain activation pattern during a visuospatial cognitive task: a functional magnetic resonance imaging study in healthy volunteers. *Neuroscience Letters*, **344**(3), 169–172.

Whittaker, S., Terveen, L., & Nardi, B.A. (2000). Let's stop pushing the envelope and start addressing it: A reference task agenda for HCI. *Human-Computer Interaction*, **15**, 75–106.

Wurtz, R.H., Goldberg, M.E., & Robinson, D.L. (1982). Brain mechanisms of visual attention. *Scientific American*, **246**, 124–135.

3

Moving usability forward to the beginning of the software development process

Natalia Juristo*, Ana Moreno*, Maria-Isabel Sanchez-Segura¥
*School of Computing - Universidad Politécnica de Madrid
Spain
¥Department of Computing - Universidad Carlos III de Madrid
Spain

1. Introduction

Software usability is a quality attribute found in a number of classifications (IEEE, 1998), (ISO9126, 1991), (Boehm, 1978). Nielsen gave one of the most well-known descriptions related to software system learnability and memorability, efficiency of use, ability to avoid and manage user errors, and user satisfaction (Nielsen, 1993). In spite of the relevance of usability in software development it is still insufficient in most software systems (Seffah, 2004) (Bias, 2005).

For the past two decades, software usability has been perceived, from a software development perspective, as related to the presentation of information to the user (Seffah & Metzker, 2004) (Folmer et al., 2004). Software engineers have treated usability primarily by separating the presentation portion from the system functionality, as recommended by generally accepted design strategies (e.g. MVC or PAC (Buschmann et al.)). This separation would make it easier to modify the user interface to improve usability without affecting the rest of the application. Accordingly, there is a belief that usability can be considered late in the development process (generally after testing) as it should not take too much rework to improve this quality attribute.

Recently, however, usability's implications in the application core have been highlighted. Some authors have already illustrated, albeit informally, a possible relationship between usability and architectural design (Bass, 2003) (Folmer, 2004). If this relationship is confirmed, the cost of rework to achieve an acceptable level of usability would be much higher than expected according to the hypothesis of separation. If this is the case, usability should be dealt with earlier in the development process in order to define and evaluate its impact on design as soon as possible. Notice that this approach is consistent with the tendency in SE to carefully consider quality attributes early on in the development process (Barbacci, 2003). This strategy has already been applied to other quality attributes like performance, modifiability, reliability, availability and maintainability, where a number of authors have proposed techniques to deal with these attributes, for example, at architectural design time (Klein, 1999) (Bass 1999) (Eskenazi, 2002) (Bosch, 2003).

In this context, the objective of this paper is twofold. On the one hand, we will offer some evidence to demonstrate the relationship between usability and software design. To do this, we will first analyze what different sorts of impact usability heuristics and guidelines discussed in the human computer interaction (HCI) literature are likely to have on a software system. We have identified usability features with a potential impact on the user interface, on the whole development process, and on the design models. To confirm this potential impact, we are particularly interested in design models. Therefore, the next step in this book chapter is to examine what sort of effect such usability features have on software design. With this aim in mind, we have surveyed several real systems that have the usability features in question built in. By way of an illustration, we will detail our study in terms of new classes, methods and relationships derived from adding one particular usability feature. As a result of this analysis we are able to demonstrate that usability really does effect the system's core functionality. With the goal of quantifying as far as possible the effect of usability features with an impact on design, this book chapter goes on to discuss the data gathered from applying a number of usability features to develop several real systems. We used these data to demonstrate the relationship between usability and software design and to get an informal estimation of the implications of building these features into a system.

Consequently, we can demonstrate that particular usability issues have a real impact on software design. Such issues have big functional implications and, therefore, need to be considered from as of the early development phases to avoid design rework, like any other functional requirement.

Accordingly, the second objective of this paper is to discuss how to deal with such usability issues at requirements time. In particular, we present some completeness problems caused by incorporating functional usability features as requirements, and discuss how the traditional solutions for dealing with incompleteness are hard to apply in this case. Then we present the approach we followed to avoid such problems, using a pattern-oriented approach to capture the knowledge to be managed to elicit and specify usability requirements. Finally we show some results related to pattern use.

To achieve the above objectives, this book chapter has been structured as follows. Section 2 discusses the different usability recommendations that can be found in the HCI literature. Section 3 provides some evidence about the relationship between particular usability recommendations and software design. This evidence shows that incorporating such recommendations involves modifying the core of the software design, as required to build in any other functionality. Section 4 shows the problems of dealing with usability during the requirements phase, and section 5 discusses a pattern-based representation of usability recommendations that avoid such limitations. Finally, section 6 presents some data gathered from the evaluation of such patterns.

2. Usability Recommendations in the HCI Literature

The usability literature has provided an extensive set of guidelines to help developers to build usable software. Each author has named these guidelines differently: design heuristics (Nielsen, 1993), principles of usability (Constantine, 1998) (Shneiderman, 1999), usability guidelines (Hix, 1993), etc. Although all these recommendations share the same goal of improving software system usability, they are very different from each other. For example, there are very abstract guidelines like "prevent errors" (Nielsen, 1993) or "support internal locus of control" (Shneiderman, 1999), and others that provide more definite usability solutions like "make the user actions easily reversible" (Hix, 1993) or "provide clearly marked exits" (Nielsen, 1993). It is not our aim to provide a detailed classification of these usability features, as this is outside the scope of software engineering. What we can do, though, is structure these features depending on their potential impact on software development. Accordingly, such features can be divided into three groups:

1) Usability recommendations with a potential impact on the UI. Examples of such recommendations refer to presentation issues like buttons, pull-down menus, colors, fonts, etc. Building these recommendations into a system involves slight modifications to the detailed UI design.
2) Usability recommendations with a potential impact on the development process, which can only be taken into account by modifying the development process itself, e.g. recommendations referring to reducing the user cognitive load, involving the user in software construction, etc.
3) Usability recommendations with a potential impact on the design. They involve building certain functionalities into the software to improve user-system interaction. We have termed these set of usability recommendations Functional Usability Features (FUFs). Examples of FUFs are providing cancel, undo, feedback, etc. Let's suppose that we want to build the cancel functionality for specific commands into an application. To satisfy the requirements for this functionality the software system must at least: gather information (data modifications, resource usage, etc.) that allow the system to recover the status prior to a command execution; stop command execution; estimate the time to cancel and inform the user of progress in cancellation; restore the system to the status before the cancelled command; etc. This means that, apart from the changes that have to be made to the UI to add the cancel button, specific components should be built into the software design to deal with these responsibilities. Table 1 shows the most representative FUFs that can be foreseen to have a crucial effect on system design. This table also includes the HCI authors that suggest each recommendation.

In this chapter, we are interested in usability recommendations with a potential impact on design and try to provide real evidence about their impact on design.

3. Analyzing the Effect of Usability on Software Design

To study the relationship between FUFs and software design we have worked on a number of real development projects carried out by UPM Master in Software Engineering students as part of their MSc dissertations from 2004 to 2005. Students originally developed the respective systems without any FUFs. These designs were then modified to include the FUFs listed in Table1.

Functional Usability Features	Goal
FEEDBACK (Tidwell, 1999) (Brigthon, 1998) (Coram, 1996) (Welie, 2003) (Tidwell, 2005) (Nielsen, 1993) (Constantine, 1999) (Shneiderman, 1998) (Hix, 1993) (Rubinstenin, 1994) (Heckel, 1991)	To inform users about what is happening in the system
UNDO (Tidwell, 2005) (Welie, 2003) (Brigthon, 1998)	To undo system actions at several levels
CANCEL (Tidwell, 2005) (Brigthon, 1998) (Nielsen, 1993)	To cancel the execution of a command or an application
USER INPUT ERRORS PREVENTION/CORRECTION (Tidwell, 2005) (Brigthon, 1998) (Shneiderman, 1998) (Hix, 1993) (Rubinstein, 1984) (Constantine, 1999)	To improve data input for users and software correction as soon as possible
WIZARD (Welie, 2003) (Tidwell, 2005) (Constantine, 1998)	To help to do tasks that require different steps involving user input
USER PROFILE (Tidwell, 1999) (Welie, 2003) (Hix, 1993) (Rubinstenin, 1994) (Heckel, 1991)	To adapt system functionality to users' profile
HELP (Tidwell, 2005)(Welie, 2003)(Nielsen, 1993)	To provide different help levels for different users
COMMAND AGGREGATION (Nielsen, 1993) (Constantine, 1999) (Hix, 1993)	To help users to create commands to execute more than one task at a time
SHORTCUTS (Nielsen, 1993) (Constantine, 1999) (Hix, 1993) (Shneiderman, 1998)	To allow users to activate a task with one quick gesture.
REUSE INFORMATION (Constantine, 1999)	To allow users to easily move data from one part of a system to another

Table 1. Preliminary list of usability features with impact on software design

The projects used were interactive systems (an on-line table booking for a restaurant chain, an outdoor advertising management system, a car sales system, an adaptable surface transport network system, a computer assembly warehouse management system; an on-line theatre network ticket sales and booking system; and an employee profile and job offer processing and matching software system). We deliberately chose interactive systems because usability is more relevant in these cases, and FUFs can be expected to have a bigger impact.

For each of the systems to which the FUFs listed in Table 1 were added, we quantified a number of criteria:

- *FUF impact on system functionality* (FUF-Functionalities). This parameter mirrors the number of functionalities (in terms of expanded use cases) affected by the FUF in question. To assess this criterion we calculated the percentage of expanded use cases affected by each FUF, which was rated as low, medium or high depending on the interval to which the percentage belongs (under 33%, from 33% to 66%, over 66%).
- *FUF-derived classes* (FUF-Classes). This criterion refers to the number of classes that appear in the design as a result of adding a FUF. This has been assessed by calculating the percentage of new classes derived from the feature, which was rated as low, medium or high depending on the interval to which the percentage belongs (under 33%, from 33% to 66%, over 66%).
- *FUF-derived methods complexity* (FUF-Methods Complexity). The criterion refers to how complex the methods that need to be created as a result of incorporating a given FUF into the system are. It is not easy to provide a measure of the complexity of a method at design time. For the purposes of our study, however, we have classified the possible class methods based on their functionality as follows:
    - Methods related to displaying information, running checks, etc., have been rated as low
    - Methods related to filters, error corrections, etc., have been rated as medium.
    - Methods related to returning to the earlier state of an operation, saving the state, etc., have been rated as high.
- *Interaction with other system components* (FUF-Interaction). This parameter represents how the classes involved in FUF design couple with the other system classes. To assess this parameter, we measured the percentage of interactions between the FUF-derived classes or between these and other system classes that can be observed in the interaction diagrams. The value of this criterion will be low, medium and high depending on what third this percentage belongs to (under 33%, from 33% to 66%, over 66%).

The need to build different FUFs into a particular project will depend on the project features. For example, *shortcuts* will have a low value for impact on system functionality (FUF-Functionality) if we are dealing with a software system that will only be executed from time to time, whereas it will have a high value if the application runs continuously and performs the same tasks again and again. Similarly, the other FUFs could be designed to affect more or fewer parts of the software system. In our study, all usability features addressed were specified as being included in the whole system and related to the maximum number of functionalities to which they applied. For example, when *feedback* was added, it was considered that the whole breadth of this feature was needed, including progress bars, clocks, etc., for all the tasks that have need of this functionality.

Additionally, the FUF-Classes, FUF-Methods and FUF-Interactions criteria will very much depend on the type of design. The values output for our systems should not be construed as absolute data. On the contrary, they are intended to illustrate to some extent what effect adding the respective FUFs could have on design.
Readers are referred to (Juristo et al. 2007) for details of this study for one of the above applications. Table 2 summarizes the mean values of the metrics derived from incorporating the FUFs in the above systems. It is clear from this table that the *cancel* and *undo* FUFs have the biggest impact on design. Not many more classes (FUF-Classes) are added (as in the chosen design a single class is responsible for saving the last state for whatever operations are performed, although another equally valid design could have envisaged a separate class to save the state for each operation). However, the complexity of the methods (FUF-Methods-Complexity) that need to be implemented is high, as is the number of interactions between the different classes (FUF-Interactions). In the *cancel* case especially, this feature is closely related to all system functionalities (FUF-Functionalities), because the HCI literature recommends that easy exit or cancellation should be provided for each and every one of the tasks that the user uses the system to do (Tidwell, 99).

Another FUF with a big impact on all system functionality is *feedback*. Apart from the *system status feedback* discussed in the last section, the HCI literature also recommends that the user should receive feedback reporting the *progress* of the operations when the user is doing long tasks (Tidwell, 2005) (Brigthon, 1998) (Coram, 1996) (Welie, 2003), when the *tasks are irreversible* (Brigthon, 1998) (Welie, 2003) and, additionally, every time the user *interacts* with the system (Brighton, 2003). It is this last recommendation especially that leads to the high FUF-functionality for this feature, as it means that feedback affects all a software system's non-batch functionalities.

On the other hand, we find that the impact of adding other FUFs, like for example *user profile*, are less costly because they can be easily built into a software system and do not interact very much with the other components. A similar thing applies to *help*. In this case, though, despite its low impact on functionality (because this functionality was designed as a separate use case, yielding a 5% and therefore low FUF-functionality value), its interaction is high, as it can be called from almost any part of the system.

It is noteworthy that no big differences where found among the applications because they were similar, i.e. they were all management systems. Note that these same FUFs may have a slightly different impact on other software system types, for example, control systems (in which FUFs like *user input errors prevention/correction* or *commands aggregation* may have a bigger impact than shown in Table 2 due to the criticality of the tasks performed) or less interactive systems (in which *feedback* or *cancel* will have less impact).

In sum, the data in Table 2 confirm that some usability recommendations, in particular the ones we have named FUFs, affect the core functionality of a software system. As with any other functionality, specific design components will have to be created to build such FUFs into a software application. The approach we take is to consider such FUFs as functional requirements and deal with them during the requirements process as any other functionality. The rest of the chapter focuses on how to address this proposal.

Summary	FUF-Functionality	FUF-Classes	FUF-Methods Complexity	FUF-Interactions
Feedback	HIGH 90%	LOW 27%	MEDIUM	MEDIUM/HIGH 66%
Undo	MEDIUM 40%	LOW 10%	HIGH	MEDIUM/HIGH 66%
Cancel	MEDIUM 95%	LOW 8%	HIGH	MEDIUM/HIGH 66%
User Input Errors Prevention/Correction	MEDIUM 36%	LOW 11%	MEDIUM	LOW 6%
Wizard	LOW 7%	LOW 10%	LOW	HIGH 70%
User Profile	LOW 8%	MEDIUM 37%	MEDIUM	LOW 10%
Help	LOW 7%	LOW 6%	LOW	HIGH 68%
Commands aggregation	LOW (10%)	LOW (5.8%)	MEDIUM	LOW 15%

Table 2. Mean values for design impact of FUF

4. Limitations of Usability Requirements

The idea of dealing with usability at the requirements phase is not new. Both HCI (Jokela, 2005) and SE (Swebok, 2004) have considered usability as a non-functional requirement. In this context, usability requirements specify user effectiveness, efficiency or satisfaction levels that the system should achieve. These specifications are then used as a yardstick at the evaluation stage: “A novice user should learn to use the system in less than 10 hours”, or “End user satisfaction with the application should be higher than Z on a 1-to-5 scale”. Dealing with usability in the shape of non-functional requirements does not provide developers with enough information about what kind of artifacts to use to satisfy such requirements.

Recent studies have targeted the relationship between usability and functional requirements. Cysneiros et al. suggest identifying functional requirements that improve particular usability attributes (Cysneiros, 2005). We propose a complementary approach in which usability features with major implications for software functionality, FUFs, are incorporated as functional requirements.

Usability functionalities could be specified by just stating the respective usability features. For example, “the system should provide users with the ability to cancel actions” or “the system should provide feedback to the user”. This is actually the level of advice that most

HCI heuristics provide. However, descriptions like these provide nowhere near enough information to satisfactorily specify the feedback functionality, let alone design and implement it correctly.
To illustrate what information is missing let us look at the complexity and diversity of the feedback feature. As we will see later, the HCI literature ((Tidwell, 1996)(Welie, 2003)(Laasko, 2003)(Brighton, 1998)(Coram, 1996)(Benson, 2002)) identifies four types of Feedback: Interaction Feedback to inform users that the system has heard their request; Progress Feedback for tasks that take some time to finish; System Status Display to inform users about any change in the system status, and Warnings to inform users about irreversible actions. Additionally, each feedback type has its own peculiarities. For example, many details have to be taken into account for a system to provide a satisfactory System Status Feedback: what states to report, what information to display for each state; how prominent the information should be in each case (e.g., should the application keep control of the system while reporting, or should the system let the user work on other tasks during status reporting), etc. Therefore, a lot more information than just a description of the usability feature must be specified to properly build the whole feedback feature into a software system. Developers need to discuss this information with and elicit it from the different stakeholders.

Note that the problem of increasing functional requirements completeness is generally solved by adding more information to the requirements (Kovitz, 2002)(Benson, 2002). However, in this case, neither users nor developers are good sources of the information needed to completely specify a usability feature. Users know that they want feedback; what they do not know is what kind of feedback can be provided, what is best for each situation, and less still what issues need to be detailed to properly describe each feedback type. Neither do software engineers have the necessary HCI knowledge to completely specify such functional usability requirements since they are not usually trained in HCI skills (Kazman et al, 2003).

The HCI literature suggests that HCI experts should join software development teams to provide this missing expertise (ISO, 1999)(Mayhew, 1999). However, this solution has several drawbacks. The first is that communication difficulties arise between the software developer team and HCI experts, as HCI and SE are separate disciplines (Seffah & Metzker, 2004). They use different vocabulary, notations, software development strategies, techniques, etc. Misunderstandings on these points can turn out to be a huge obstacle to software development. Another impediment is the cost. Large organizations can afford to pay for HCI experts, but many small-to-medium software companies cannot.

4. Generating Usability Elicitation Patterns

Our approach consists of packaging guidelines that empower developers to capture functional usability requirements using the information provided by the HCI literature as input. We have analyzed this information from a software development point of view and have elaborated elicitation and specification guidelines that have been packaged in a pattern format.

The first task was to analyze the different varieties of usability features identified by HCI authors. We denoted these subtypes as usability mechanisms, and gave them a name that is indicative of their functionality (see Table 3)
Then we defined the elicitation and specification guides for the usability mechanisms, focusing on the information provided by HCI authors. We analyzed and combined all the recommendations on the same mechanism, and then removed redundancies. The resultant HCI recommendations cannot be used directly to capture software requirements, but they can be studied from a development point of view to generate issues to be discussed with the stakeholders to properly specify such usability features.

The outcome of the previous tasks is packaged in what we call a usability elicitation pattern. Other authors have already used patterns to reuse requirements knowledge. Patterns that capture general expertise to be reused during different requirements activities (elicitation, negotiation, documentation, etc.) are to be found in (Hagge, 2005)(Repare, 2005), for example. In (Whitenak, 1995), the author proposes twenty patterns to guide the analyst through the application of the best techniques and methods for the elicitation process.

Our usability elicitation patterns capitalize upon elicitation know-how so that requirements engineers can reuse key usability issues intervening recurrently in different projects. These patterns help developers to extract the necessary information to completely specify a functional usability feature.

We have developed one usability elicitation pattern for each usability mechanism in **Pogreška! Izvor reference nije pronađen.**3 (second column). They are available at http://is.ls.fi.upm.es/research/usability/usability-elicitation-patterns. Table 4 shows an example of the elicitation pattern for the System Status Feedback mechanism.

The developer can use the *identification* part of the pattern to find out the basics of the usability mechanism to be addressed. The discussion with the stakeholders starts by examining the pattern *context* section that describes the situations for which this mechanism is useful. If the mechanism is not relevant for the application, it will not be used. Otherwise, the respective usability functionality will be elicited and specified using the *solution* part of the pattern.

The solution part of the pattern contains two elements: *the usability mechanism elicitation guide* and *the usability mechanism specification guide.* The *usability mechanism elicitation guide* provides knowledge for eliciting information about the usability mechanism. It lists the issues that stakeholders should discuss to properly define how the usability mechanism should be considered, alongside the respective HCI rationale (i.e. the HCI recommendation used to derive the respective issues). Developers should read and understand the HCI rationales in the guide. This will help them to understand why those issues need to be discussed with stakeholders.

The elicited usability information can be specified following the pattern *specification guide*. This guide is a prompt for the developer to modify each requirement affected by the incorporation of each mechanism. An example of the application of this usability elicitation pattern is given in (Juristo, et al, 2007a) .

Usability Feature	Usability Mechanism	HCI Authors' Label	Goal
Feedback	System Status	Modeless Feedback Area (Coram, 1996) Status Display (Tidwell, 1996)	To inform users about the internal status of the system
	Interaction	Interaction Feedback (Brighton, 1998) Modeless Feedback Area (Coram, 1996) Let Users Know What is Going On (Benson, 2002)	To inform users that the system has registered a user interaction, i.e. that the system has heard users
	Warning	Think Twice (Brighton, 1998) Warning (Welie, 2003)	To inform users of any action with important consequences
	Long Action Feedback	Progress Indicator (Tidwell, 1996) (Tidwell, 2005) Show Computer is Thinking (Brighton, 1998) Time to Do Something Else (Brighton, 1998) Progress (Welie, 2003) Modeless Feedback Area (Coram, 1996) Let Users Know What is Going On (Benson, 2002)	To inform users that the system is processing an action that will take some time to complete
Undo Cancel	Global Undo	Multi-Level Undo (Tidwell, 1996) (Tidwell, 2005) Undo(Welie, 2003) Global Undo (Laasko, 2003) Allow Undo (Brighton, 1998) Go Back One Step (Tidwell, 1996)	To undo system actions at several levels
	Object-Specific Undo	Object-Specific Undo (Laasko, 2003)	To undo several actions on an object
	Abort Operation	Go Back One Step (Tidwell, 1996) Emergency Exit (Brighton, 1998) Cancellability (Tidwell, 2005)	To cancel the execution of an action or the whole application
	Go Back	Go Back to a Safe Place (Tidwell, 1996) Go Back One Step (Tidwell, 1996)	To go back to a particular state in a command execution sequence
User Input Error Prevention/ Correction	Structured Text Entry	Forms, Structured Text Entry (Tidwell, 1996) Structured Format (Tidwell, 2005) Structured Text Entry (Brighton, 1998)	To help prevent the user from making data input errors
Wizard	Step-by-Step Execution	Step-by-Step (Tidwell, 1996) Wizard (Welie, 2003) (Tidwell, 2005)	To help users to do tasks that require different steps with user input and correct such input
User Profile	Preferences	User Preferences (Tidwell, 1996) Preferences (Welie, 2003)	To record each user's options for using system functions
	Personal Object Space	Personal Object Space (Tidwell, 1996)	To record each user's options for using the system interface.
	Favorites	Favorites (Welie, 2003) Bookmarks (Tidwell, 1996)	To record certain places of interest for the user
Help	Multilevel Help	Multilevel Help (Tidwell, 2005)	To provide different help levels for different users
Command Aggregation	Command Aggregation	Composed Command (Tidwell, 1996) Macros (Tidwell, 2005)	To express possible actions to be taken with the software through commands that can be built from smaller parts.

Table 3. Usability mechanisms for which usability elicitation and specification guides have been developed

IDENTIFICATION	
Name: System Status Feedback	
Family: Feedback	
Alias: Status Display Modeling Feedback Area (Coram, 1996)	
PROBLEM	
Which information needs to be elicited and specified for the application to provide users with status information.	
CONTEXT	
When changes that are important to the user occur or when failures that are important to the user occur, for example: during application execution; because there are not enough system resources; because external resources are not working properly. Examples of status feedback can be found on status bars in windows applications; train, bus or airline schedule systems; VCR displays; etc.	
SOLUTION	
Usability Mechanism Elicitation Guide:	
HCI Rationale	**Issue to discuss with stakeholders**
1. HCI experts argue that the user wants to be notified when a change of status occurs (Tidwell, 1996)	*Changes in the system status can be triggered by user-requested or other actions or when there is a problem with an external resource or another system resource.* *1.1 Does the user need the system to provide notification of* ***system statuses?*** *If so, which ones?* *1.2 Does the user need the system to provide notification of system* ***failures*** *(they represent any operation that the system is unable to complete, but they are not failures caused by incorrect entries by the user)? If so, which ones?* *1.3 Does the user want the system to provide notification if there are not* ***enough resources*** *to execute the ongoing commands? If so, which resources?* *1.4 Does the user want the system to provide notification if there is a problem with an* ***external resource*** *or device with which the system interacts? If so, which ones?*
2. Well-designed displays of information to be shown should be chosen. They need to be unobtrusive if the information is not critically important, but obtrusive if something critical happens. Displays should be arranged to emphasize the important things, de-emphasize the trivial, not hide or obscure anything, and prevent one piece of information from	*2.1. Which information will be shown to the user?* *2.2. Which of this information will have to be displayed obtrusively because it is related to a critical situation? Represented by an indicator in the main display area that prevents the user from continuing until the obtrusive information is closed.* *2.3. Which of this information will have to be*
being confused with another. They should never be re-arranged, unless users do so themselves. Attention should be drawn to important information with bright colors, blinking or motion, sound or all three – but a technique appropriate to the actual importance of the situation to the user should be used (Tidwell, 1996).	*highlighted because it is related to an important but non-critical situation? Using different colors and sound or motion, sizes, etc.* *2.4. Which of this information will be simply displayed in the status area? For example, providing some indicator.* Notice that for each piece of status information to be displayed according to its importance, the range will be from obtrusive indicators (e.g., a window in the main display area which prevents the user from continuing until it has been closed), through highlighting (with

HCI Rationale	Issue to discuss with stakeholders
being confused with another. They should never be re-arranged, unless users do so themselves. Attention should be drawn to important information with bright colors, blinking or motion, sound or all three – but a technique appropriate to the actual importance of the situation to the user should be used (Tidwell, 1996).	*highlighted because it is related to an important but non-critical situation? Using different colors and sound or motion, sizes, etc.* *2.4. Which of this information will be simply displayed in the status area? For example, providing some indicator.* Notice that for each piece of status information to be displayed according to its importance, the range will be from obtrusive indicators (e.g., a window in the main display area which prevents the user from continuing until it has been closed), through highlighting (with different colors, sounds, motions or sizes) to the least striking indicators (like a status-identifying icon placed in the system status area). Note that during the requirements elicitation process, the discussion of the exact response can be left until interface design time, but the importance of the different situations about which status information is to be provided and, therefore, which type of indicator (obtrusive, highlighted or standard) is to be provided does need to be discussed at this stage.

Table 4. (a) System status feedback usability elicitation pattern

SOLUTION (Cont.)	
Usability Mechanism Elicitation Guide (Cont.):	
HCI Rationale (Cont.)	**Issue to discuss with stakeholders (Cont.)**
3. As regards the location of the feedback indicator, HCI literature mentions that users want one place where they know they can easily find this status information (Coram, 1996). On the other hand, aside from the spot on the screen where users work, users are most likely to see feedback in the centre or at the top of the screen, and are least likely to notice it at the bottom edge. The standard practice of putting information about changes in state on a status line at the bottom of a window is particularly unfortunate, especially if the style guide calls for lightweight type on a grey background (Constantine, 1998). The positioning of an item within the status display should be used to good effect. Remember that people born into a European or American culture tend to read left-to-right, top-to-bottom, and that something in the upper left corner will be looked at most often (Tidwell, 1996).	*3.1. Do people from different cultures use the system? If so, the system needs to present the system status information in the proper way (according to the user's culture). So, ask about the user's reading culture and customs.* *3.2. Which is the best place to locate the feedback information for each situation?*

Usability Mechanism Specification Guide:
The following information will need to be instantiated in the requirements document. - The system statuses that shall be reported are X, XI, XII. The information to be shown in the status area is..... The highlighted information is … The obtrusive information is…. - The software system will need to provide feedback about failures I, II, III occurring in tasks A, B, C, respectively. The information related to failures I, II, etc…. must be shown in status area…. The information related to failures III, IV, etc , must be shown in highlighted format. The information related to failures V, VI, etc , must be shown in obtrusive format. - The software system provides feedback about resources D, E, F when failures IV, I and VI, respectively, occur. The information to be presented about those resources is O, P, Q. The information related to failures I, II, etc….must be shown in the status area..... The information related to failures III, IV, etc , must be shown in highlighted format. The information related to failures V, VI, etc , must be shown in obtrusive format. - The software system will need to provide feedback about the external resources G, J, K, when failures VII, VIII and IX, respectively, occur. The information to be presented about those resources is R, S, T. The information related to failures I, II, etc….must be shown in the status area..... The information related to failures III, IV, etc., must be shown in highlighted format. The information related to failures V, VI, etc., must be shown in obtrusive format.
RELATED PATTERNS[1]:

Table 4. (b) System status feedback usability elicitation pattern (cont.)

5. Preliminary Evaluation of Usability Elicitation Patterns

The potential benefits of the usability elicitation patterns have been evaluated at different levels.

We studied how useful the patterns were for building the usability mechanisms into a software system. We expected pattern use to lead to an improvement on the original situation where developers did not have any compiled or systematic usability information. We worked with SE Master students. In particular, we worked with five groups of three students. Each group was given a different software requirement specification document (for a theatre tickets sale system, for a PC storage and assembly system, for a temping agency job offers management system, for a car dealer vehicle reservation and sale system, and for a travel agency bookings and sale system). All the systems were real applications,

[1] *Related patterns* refer to other usability elicitation patterns whose contexts are related to the one under study and could also be considered in the same application. In this case, no related patterns have being identified. However, readers are referred to other patterns, like Long Action Feedback or Abort Operation, at the above-mentioned web site.

and each one was randomly allocated to a group. Each of the three students in the group was asked to add the functionality derived from the functional usability features listed in section 4 to the original SRS independently and to build the respective software system. The procedure was as follows:

- We gave one of the students the usability elicitation patterns discussed in this paper. This student used the pattern content to elicit the corresponding usability functionality.

- Another student was given reduced patterns. See Appendix, including the reduced pattern for System Status Feedback, to get a taste of the difference between the reduced and full patterns. This short pattern is just a compilation of information from the HCI literature about the usability mechanisms. We have not elaborated this information from a development perspective, i.e. the reduced patterns do not include the "Issues to be discussed with stakeholders" column in Table 3.
  The idea behind using the reduced patterns was to confirm whether our processing of the HCI information resulting in the formulation of specific questions was useful for eliciting the functionality related to the mechanisms or whether developers are able to extract such details just from the HCI literature.
- Finally, the third student was given just the definitions of the usability features according to the usability heuristics found in the HCI literature and was encouraged to take information from other sources to expand this description.

Students of each group were randomly allocated the usability information they were to use (completed patterns, reduced patterns, no patterns) to prevent student characteristics from possibly biasing the final result.

Final system usability was analyzed differently to determine how useful the elicitation patterns were for building more usable software. We ran what the HCI literature defines as usability evaluations carried out by users and heuristic evaluations done by usability experts (Constantine, 1998) (Shneiderman, 1999)(Nielsen, 1993).

6.1. Users' usability evaluation

The usability evaluations conducted by users are based on usability tests in which the users state their opinion about the system. We used an adaptation of the QUIS usability test (QUIS, 2007). Each test question is scored on a scale of 1 (lowest usability) to 5 (highest usability). The final usability score is the mean of the responses to each question. We worked with three representative users for each system. Each user evaluated the three versions of each application (the one developed with the full patterns, with the reduced patterns and with no patterns) in different order.

The mean usability values for the five applications are 4.4, 3.2 and 2.5, with standard deviations of 0.3, 0.2, and 0.4, respectively. The Kruskal-Wallis test confirmed that there was a statistically significant difference among these usability means ($p\text{-value}<0.01$; chi-square = 36.625). The Tamhane test (for unequal variances) showed that the usability value for the systems developed using the full patterns was statistically greater than the score achieved using the reduced patterns, and both were greater than the usability value attained without any pattern (in all cases p-value<0.01). Therefore, we were able to confirm that the users perceived the usability of the systems developed with the full usability elicitation patterns to be higher.

With the aim of identifying the reasons that led users to assess the usability of the different types of applications differently, we had an expert in HCI run a heuristic evaluation.

6.2. Usability Expert Evaluation

A paid independent HCI expert ran the usability evaluation of the applications developed by our MSc students. The expert analyzed the applications focusing on how these systems provided the usability features listed in Table 2.

Table 5 shows the results of the heuristic evaluation. It indicates the extent to which the evaluated software incorporates the functionality related to each usability mechanism. In the case of feedback, for example, the developers that used the respective elicitation patterns included, on average, 94% of the functionalities associated with this mechanism. Developers that used the reduced patterns incorporated 47% of the respective functionalities. Finally, developers that used no pattern included only 25%.

Applying the Kruskal-Wallis test to the expert results for each usability feature we found that there were statistically significant differences among the three groups of data (see last column of Table 5 with p-value<0.01 in all cases). Again the Tamhane test showed that all the usability features were built into the systems developed using the full patterns better than they were into systems developed using the reduced patterns, and both provided more usability details than systems developed without patterns (with feature definitions only). This explains why users perceived differences in the usability of the systems.

	Full usability elicitation patterns	**Reduced patterns**	**No pattern**	**Kruskal-Wallis(chi-square; p-value)**
Feedback	94%	47%	25%	12,658; 0,002*
Undo/Cancel	90%	66%	43%	12,774; 0,002*
User Profile	95%	80%	65%	12,597; 0,002*
Users Input Errors Prevention/ Correction	97%	85%	72%	12,727; 0,002*

	Full usability elicitation patterns	Reduced patterns	No pattern	Kruskal-Wallis(chi-square; p-value)
Wizard	100%	89%	71%	13,109: 0,001*
Help	100%	81%	74%	13,109; 0,001*

* Statistically significant at 99% of confidence

Table 5. Mean percentage of functonality added for each usability mechanism by each information type

Note that the functionality added using the full elicitation patterns is less than 100% for the most complex patterns like Feedback and Undo. These differences are due to the fact that the complexity of these features calls for a very thorough analysis of the specifications to properly identify what parts of the system are affected. The final result then depends on how detailed and thorough the analyst is.

Although bringing an HCI expert into systems development could possibly have led to 100% of all the usability details being identified, elicitation pattern use is an efficient alternative because of its cost. Also, developers should become more acquainted with the patterns as they apply them, and efficiency in use should gradually improve.

Although these are interim data and further checks need to be run, the usability evaluations performed have revealed trends that need to be formally tested with a larger group of users and applications. The users' evaluation has shown that users perceive usability to be better in the versions of the application developed with the full usability elicitation patterns. On the other hand, the expert evaluation found no significant weaknesses in the usability functionality provided in the applications built using such patterns, whereas it detected sizeable gaps in applications built with reduced patterns or without any pattern at all.
These findings give us some confidence in the soundness of the usability elicitation patterns as a knowledge repository that is useful in the process of asking the right questions and capturing precise usability requirements for developing software without an HCI expert on the development team.

7. Conclusions

The goal of this chapter was first to provide some data about the impact of including particular usability recommendations in a software system. The data gathered show that building certain usability components into a software system really does entail significant changes to the software system design. Therefore, it is important to move usability issues forward to the early development phases, i.e. to requirements time (like any other functionality). However, this is not a straightforward objective primarily due to the fact that development stakeholders are not acquainted with HCI.

We propose a possible solution to overcome these snags. To do this, we have developed specific guidelines that lead software practitioners through the elicitation and specification process. This approach supports face-to-face communication among the different stakeholders during requirements elicitation to cut down ambiguous and implicit usability details as early as possible. These guidelines help developers to determine whether and how a usability feature applies to a particular system, leading to benefits for the usability of the final system.

Evidently, the use of usability patterns and any other artifact for improving software system usability calls for a lot of user involvement throughout the development process. This is a premise in the usability literature that is also necessary in this case. If this condition cannot be satisfied, the final system is unlikely to be usable. In our opinion, then, a balance has to be struck between user availability, time and cost constraints, on the one hand, and usability results, on the other, at the beginning of development.

8. References

Andrés, A., Bosch, J., Charalampos, A., Chatley, R., Ferre, X., Folmer, E., Juristo, N., Magee, J, Menegos, S., Moreno, A. 2002. Usability Attributes Affected by Software Architecture. Deliverable. 2. STATUS project. Available at: http://www.ls.fi.upm.es/status

Barbacci M, Ellison R., Lattanze A., Stafford J.A, Weinstock C.B., Wood W.G. 2003. Quality Attribute Workshop, 3rd ed. CMU/SEI-2003-TR-016, Pittsburgh, Software Engineering Institute, Carnegie Mellon University, USA.

Bass, L. and John, B.E. 2003. Linking usability to software architecture patterns through general scenarios. Journal of Systems and Software, Volume 66, Issue 3, pp. 187-197

Bass, L., et al. 1999. Architecture-Based Development. CMU/SEI-1999-TR-007. SEI/CMU.

Bass, L., John, B., Kates, J. 2001. Achieving Usability Through Software Architecture. Technical Report. CMU/SEI-2001-TR-005

Battey, J. 1999. IBM´s redesign results in a kinder, simpler web site. Available at: http://www.infoworld.com/cgi-bin/displayStat.pl?/pageone/opinions/hotsites/hotextr990419.htm .

C. Benson, A. Elman, S. Nickell, C. Robertson. GNOME Human Interface Guidelines. http://developer.gnome.org/projects/gup/hig/1.0/index.html

Bias, R.G., Mayhew D.J. 2005. Cost-Justifying Usability. An Update for the Internet Age. Elsevier.

Black, J. 2002. Usability is next to profitability. BusinessWeek Online. Available at: http://www.businessweek.com/technology/content/dec2002/tc2002124_2181.htm

Boehm B., et al. 1978. Characteristics of Software Quality. North Holland, New York.

Bosch, J., Lundberg, L. 2003. Software architecture – Engineering quality attributes. Journal of Systems and Software, Volume 66, Issue 3, pp. 183-186.
Brighton. 2003. Usability Pattern Collection. http://www.cmis.brighton.ac.uk/research/patterns/
Constantine, L., Lockwood, L. 1999. Software for Use. A Practical Guide to the Models and Methods of Usage-Centered Design. Addison-Wesley.
F. Buschmann, R. Meuneir, H. Rohnert, P. Sommerland, M. Stal. Pattern-Oriented Software Architecture, A System of Patterns. J. Wiley and Sons, 1996
Coram, T., Lee, L. 1996. Experiences: A Pattern Language for User Interface Design. http://www.maplefish.com/todd/papers/experiences/Experiences.html
Chidamber, S., Kemerer, C. 1994. A Metrics Suite for Object Oriented Design. IEEE Transactions on Software Engineering, June, pp. 476-492.
Donahue G. M. 2001. Usability and the Bottom Line. IEEE Software, vol. 16, n. 1. pp 31-37.
Eskenazi, E.M., Fioukov, A.V., Hammer, D.K., Obbink, H., 2002. Performance Prediction for Software Architectures, in Proceedings of PROGRESS 2002 workshop, Netherlands.
Ferre, X. ,Juristo, N., Moreno, A. 2006. Integration of HCI Practices into Software Engineering Development Processes: Pending Issues. Encyclopedia of Human-Computer Interaction. pp. 422-428. C. Ghaoui (ed.). Idea Group Inc.
Fetcke, T., Abran, A., Nguyen, T., 1997. Mapping the OO - Jacobson. Approach into Function Point Analysis. Software. Technology of Object-Oriented Languages and Systems. Proceedings of TOOLS 1997.
Folmer, E., Group, J., Bosch, J. 2004.. Architecting for usability: a survey. Journal of Systems and Software, vol 70. pp. 61-78.
Griffith, J. 2002. Online transactions rise after bank redesigns for usability. The Business Journal. 2002. Available at: http://www.bizjournals.com/twincities/stories/2002/12/09/focus3.html
L. Hagge. K. Lappe. "Sharing Requirements Engineering Experience Using Patterns". IEEE Software. Jan-Feb 2005, pp. 24-31.
Heckel, P. 1991. The elements of friendly software design. (2nd ed.) CA: Sybex Inc.
Hix, D., Hartson, H. R.. 1993. Developing User Interfaces: Ensuring Usability Through Product and Process. John. Wiley & Sons, New York.
IBM, 2005. Cost Justifying Ease of Use. Available at: http://www-3.ibm.com/ibm/easy/eou_ext.nsf/Publish/23 (current 18 May. 2005).
IEEE. 1998. IEEE Std 1061: Standard for a Software Quality Metrics Methodology.
ISO. 1998. ISO 9241-11, 98: Ergonomic Requirements for Office work with Visual Display Terminals. Part 11: Guidance on Usability. ISO.
ISO. 2000. ISO 18529, 00: Human-Centered Lifecyle Process Descriptions. ISO.
ISO/IEC, 1991. ISO 9126: Information Technology - Software quality characteristics and metrics.
ISO/IEC. 1999, ISO14598-1, 99: Software Product Evaluation: General Overview. ISO/IEC.
ISO Std 13407: Human-Centred Design Processes for Interactive Systems. ISO, 1999.1
T. Jokela. "Guiding Designers to the World of Usability: Determining Usability Requirements through Teamwork". In Human-Centered Software Engineering. A. Seffah, J. Gulliksen and M. Desmarais, Kluwer 2005

Juristo, N., Moreno A.M., Sánchez-Segura, M. Guidelines for Eliciting Usability Functionalities (pp. 744-758). IEEE Transactions on Software Engineering. November 2007, vol 33 (11).

Juristo N., Moreno A.M., Sánchez-Segura M. Analysing the Impact of Usability on Software Design (pp:1506 – 1516). Journal of System and Software. Vol. 80(9) September 2007.

R. Kazman, J. Gunaratne, B. Jerome. "Why Can't Software Engineers and HCI Practitioners Work Together?" In Human-Computer Interaction Theory and Practice. C. Stephanidis, L. Erlbaum (Eds.). Elsevier, 2003.

Klein, M., et al. 1999. Attribute-Based Architectural Styles. CMU/SEI-99-TR-022. SEI/CMU.

B. Kovitz. "Ambiguity and What to Do about It". IEEE Joint International Conference on Requirements Engineering 2002 (Key talk).

Laasko, S. A. 2003. User Interface Designing Patterns. http://www.cs.helsinki.fi/u/salaakso/patterns/index_tree.html Visited October 2004.

Mayhew, D.J. 1999. The Usability Engineering Lifecycle. Morgan Kaufmann.

McKay, E.N. 1999. Developing User Interfaces for Microsoft Windows, Microsoft Press

Nielsen, J. 1993. Usability Engineering, AP Professional, Boston, Mass.

Perry, D., Wolf, A. 1992. Foundations for the Study of Software Architecture. ACM Software Engineering Notes, vol 17 (4), pp. 40-52.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T. 1994. Human-Computer Interaction. Addison Wesley.

QUISTM Questionnaire For User Interaction Satisfaction. http://lap.umd.edu/QUIS/

REPARE http://repare.desy.de/Repare/RepareController.

Rubinstein, R, Hersh, H. 1984. The Human Factor. Digital Press, Bedford, MA.

Scapin, D. L., Bastien, J.M. C. 1997. Ergonomic criteria for evaluating the ergonomic quality of interactive systems, Behaviour & Information Technology, vol 16, no 4/5, pp.220-231.

Seffah, A., Metzker E. 2004. The Obstacles and Myths of Usability and Software Engineering. Communications of the ACM, Vol. 47(12), pp. 71-76.

Shneiderman, B. 1998. Designing the User Interface: Strategies for Effective Human-Computer Interaction. (3rd ed. ed.). Menlo Park, CA: Addison Wesley.

SWEBOK. Guide to the Software Engineering Body of Knowledge. 2004 Version. http://www.swebok.org

Thibodeau, P. 2002. Users Begin to Demand Software Usability Tests. ComputerWorld. Available at: http://www.computerworld.com/softwaretopics/software/story/0,10801,76154,00.html

Tidwell, J. 2005. Designing Interfaces. Patterns for Effective Interaction Design. O´Reilliy, USA.

Tidwell, J. 1999. Common Ground: A Pattern Language for Human-Computer Interface Design. http://www.mit.edu/%7Ejtidwell/interaction_patterns.html

Trenner L., et al. 1998. The Politics of Usability. Springer, London, UK.

Welie M. 2003. Amsterdam Collection of Patterns in User Interface Design. http://www.welie.com/.ž

B.G. Whitenak.. "RAPPeL: A Requirements-Analysis Pattern Language for Object Oriented Development". In Pattern Languages of Program Design, J.O. Coplien and D.C. Schmidt (eds.), Addison-Wesley, 1995.

4

Sketch-Based Interfaces for Parametric Modelling

Ferran Naya, Manuel Contero & Nuria Aleixos
Universidad Politécnica de Valencia
Spain
Joaquim A. Jorge
INESC-ID & Universidade Técnica de Lisboa
Portugal
Pedro Company
Universidad Jaume I
Spain

1. Introduction

Sketching is still widely used by designers and engineers as it continues to be a useful and powerful tool that helps designers during the conception of a new product (Tversky, 2002). If engineers and designers generally use sketches the question is, why sketching is not integrated in the digital design process? Available Graphical User Interfaces (GUI) for CAD applications are still by and large constrained by the WIMP (Windows, Icons, Menus and Pointing) paradigm and current commercial CAD systems not support sketch-based design. Therefore, the problem is that the sketches continue to be unplugged to the rest of the design process. In other words, in spite of recent advances in Computer Aided Design, current CAD tools are not well suited to the initial design stages of product development, because many techniques and idioms characteristic of hand-made drawings cannot be used directly in CAD systems. To sum up, there is a disconnection between sketching and CAD tools in the new product development process and true Computer-aided Sketching (CASk) tools are required.

During last decades different research lines have been explored to improve the human-computer interface in CAD systems. In this context, some CASk systems have been developed to support freehand drawings as a way to create and edit three-dimensional geometric models. These advanced CASk systems try to provide more functionality than paper or a whiteboard, giving an added value to sketching on a digital environment. This extra functionality usually has been directed either to improve the graphic quality of the sketch by means of a beautification process or it has been oriented to automatically transform the 2D sketch into a 3D model. Interest in CASk systems has increased in the last years as new hardware devices such as Tablet-PCs and LCD graphics tablets have been launched to the market.

In this chapter we show our main contributions in the field of computer aided sketching. The aim of this work is to explore new interaction paradigms in CASk tools, geared at exploiting sketching skills of designers and engineers. Through this chapter the GEGROSS application developed by our research group (www.regeo.uji.es) will be used to illustrate the important concepts. GEGROSS is a CASk application than performs an online conversion of a raw sketch into a 3D model supporting parametric control of geometry.

2. Sketch Based Interfaces and Modelling (SBIM)

Over the last decades different research lines have been explored to improve the human-computer interface in CAD systems. One of these new approaches is termed as "Sketch-based interfaces and modelling" (SBIM) that is an emerging research field oriented to the creation of new computer tools to promote a shift (Igarashi & Zeleznik, 2007) to a new paradigm where sketches would be used as input to create 3D digital engineering models. Recent advances in SBIM applications promise better integration of sketching and CAD tools, integrating a paradigm shift to change the way geometric modelling applications are built, in order to focus on user-centric systems, rather than systems that are organized around the details of geometry representation. While most of the activity in this area in the past has been focused in off-line algorithms, where an application analyzes a complete sketch and then proposes a plausible 3D model, the growing focus on sketches and modelling has brought forth a new emphasis on approaches geared towards interactive applications. These interactive applications interpret in real time the input generated by a digitizing tablet and a pen, an approach also termed calligraphic interface (see Computers & Graphics vol. 24, special issue "Calligraphic Interfaces: towards a new generation of interactive systems"). This kind of interface relies on the analysis of the pen strokes generated by the user, and exploits the space-time information provided by those to yield richer and more expressive interaction. A common feature of these systems is to use gestures (a special graphic symbol or stroke sequence) as commands (Fonseca and Jorge, 2001). These interfaces are specially suited to applications requiring capturing rough shapes and ideas, usually associated to the conceptual design stages of new product development. In these interfaces the artificial dialogue constraints imposed by the previous generation of WIMP user interfaces are removed and designers can interact with the computer in ways evocative of more traditional media, such as paper and pencil.
To sum up, there is a growing research interest in using freehand interaction and sketches as a way to create more natural interfaces, especially for the creation and edition of three-dimensional geometric models. Digital sketching can offer an added value with respect to paper-and-pencil sketching, exploiting a more "natural" environment that does not disturb the user while he is creating the drawing. The availability of proper hardware as Tablet-PCs, electronic whiteboards and other devices supporting touch or stylus input is other of the reasons that support growing interest in this kind of interfaces.
The main requirement for designing an advanced CASk system should be to provide more functionality than paper or a whiteboard, trying to give an augmented digital paper. This extra capability with respect plain paper in some cases has been oriented to improve the graphic quality of the sketch by means of a beautification process as mentioned previously, or it has been oriented to automatically transform the 2D sketch into a 3D model. Here, it is possible to distinguish two principal approaches to transform the 2D sketch into a 3D

model. One method relies on gesture alphabets as commands for generating objects from sketches (a gesture in this context represents a graphical symbol that is translated into a command). Examples of gestural systems are SKETCH (Zeleznik et al., 1996), Teddy (Igarashi et al., 1999), GIDeS (Pereira et al., 2000) and Blobmaker (De Araujo & Jorge, 2003). Gestural systems provide predefined gesture alphabets that encode some geometric modelling operations; basically these systems substitute the selection of icons and menus by graphic gestures. The second approach, derived from computer vision, uses algorithms to reconstruct geometric objects from sketches that depict their two-dimensional projection. Examples of reconstruction systems are Stilton (Schweikardt & Gross, 2000), Digital Clay (Turner et al., 2000) and CIGRO (Contero et al., 2005).

In summary, two basic alternatives exist to create 3D models from sketches: reconstruction based and gesture based. From these approaches the reconstruction based is the most transparent to users, since they have only to create a sketch which does not require a priori knowledge of a gestural command set. Chronologically, reconstruction systems appeared before gestural ones, because reconstruction systems took advantage of previous work in offline line drawing recognition. On the other hand, gestural systems require more elaborate recognition engines for distinguishing geometry information from gestural codes and most importantly, must provide elaborate user feedback in real time. Partly due to this and because of the restricted computing power available in earlier tablet PCs, some early systems avoided the disambiguation step by using icons and menus to explicitly provide this information to the system.

3. The GEGROSS application

The REGEO research group has developed in recent years the GEGROSS system, which is a CASk interactive application that converts raw sketches to three-dimensional models. The GEGROSS application follows the gestural approach and allows the user to generate three-dimensional models using some gestural commands. In this system, it is possible to draw two-dimensional parametric freehand sections combined with the use of a simple gesture alphabet that encode some geometric modelling operations.

As can be seen in Fig. 1, in this application, the user introduces the freehand sketch directly onto a Tablet PC, using a reduced-instruction set calligraphic interface. The design goal of this interface is to create two-dimensional parametric sections and three-dimensional models in a very simple way, using the conventions of technical drawing to define the shape of the section. The user interface is designed to minimize the interaction with menus or icons in an attempt to emulate the traditional use of pen and paper.

Fig. 1. User stroke input on a Tablet PC (GEGROSS)

In the design of this system some assumptions have been taken to simplify the recognition process to interpret gestures. The main assumption is that the final user of this system will have a technical or engineering background. That means that users know the conventions of technical drawing, and the application is designed to recognize the typical drawing procedures that designers/engineers follow to create a technical sketch.
The gestural modelling process is organized in two stages. In first place a 2D profile is defined. To do this, the user introduces the geometry of the 2D section using a two-dimensional parametric freehand sketch module called ParSketch (Naya et al., 2007). This module offers many of the features provided by current commercial parametric CAD applications as it is built on top of the most common parametric engine of the market (D-Cubed components from the SIEMENS firm, www.plm.automation.siemens.com). Later, in the second stage and also using gestures, it is possible to make an extrusion or a revolution of the parametric section generated in the previous stage to create a 3D solid model. Then this process, the user can continue sketching new 2D sections onto the faces of the generated object and applying the corresponding modelling gestures.
The ParSketch module implements a calligraphic interface to manage the geometric entities and the geometric constraints found in two-dimensional sections. The system distinguishes two modes of operation: one where the strokes made by the user are interpreted as geometric entities and other where the strokes are considered as commands. In table 1 the supported gestural alphabet is presented. Majority of gestures are inspired in the typical symbols used in technical drawing. When the user introduces a new stroke, ParSketch uses the drawing pressure as a mode discriminator (geometry or gesture). Then, the application interprets the type of stroke drawn by the user using a geometry recognizer (RecoGeo) or a gestural recognizer (RecoGes). Next, an automatic beautification stage is executed transforming the strokes in the corresponding geometry entities and constraint symbols.
The geometric recognizer RecoGeo supports complex strokes that after interpretation are split into its constituent primitives, allowing users to build simple sketches composed by line segments and arcs, which are automatically tidied and beautified. The application cleans up input data and adjusts edges to make sure they meet precisely at common endpoints in order to obtain geometrically consistent figures by filtering all defects and errors of the initial sketches that are inherent to their inaccurate and incomplete nature.

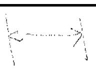
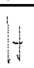
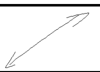
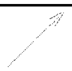
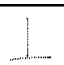
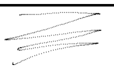
Constraint gestures	Class	Constraint gestures	Class
	Concentric		Vertical
	Linea dimension		Horizontal
	Diametral dimension		Parallel
	Radial dimension		Perpendicular
	Tangent		Cross-out (erase)

Table 1. Gesture alphabet for constraining 2D geometry implemented in ParSketch

Once the designer has introduced the complete outlined sketch, it can be edited, dimensioned and constrained using the gesture recognizer RecoGes. RecoGes has been developed to provide an alphabet of geometric/dimensional constraints to parameterise the sketches. In other words, if user wants to generate design alternatives, or adjusting some sketch to reach some dimensional condition, the system provides parametric capabilities and handwritten dimensional control to the two-dimensional freehand sections. Handwritten number recognition is provided by the Windows XP Tablet PC Edition operative system.

As explained before, the mode detection has been solved using the electronic pen pressure information, since the system is intended to be used by persons with basic engineering drawing skills. It can be said that line width is the mode-change feature when reading an engineering drawing. The usual practice is that thick lines are associated to geometry and thin lines to dimensions and other type of annotations. As line width is related to increasing pressure with the pencil while drawing, this information is used to discriminate among geometry or gesture. In other words, drawing making high pressure on the screen is intended for geometry input, while soft pressure is associated to auxiliary information. The user can configure a pressure level threshold to classify strokes as geometry or gestures.

An example of interaction with ParSketch is presented in Fig. 2. In this example the user draws the whole contour in 2.a. One single stroke is accepted as input, and it is later decomposed by the application into six rectilinear and connected strokes. When the application shows the beautified version (Fig. 2.b), the user adds another complex stroke composed by two segments and one arc. The geometry is then beautified (Fig. 2.d). In Fig. 2.e we can see the use of the scratching gesture to refine the geometry. Drawing this gesture is interpreted by the application as a command to delete those geometric entities intersecting the smallest quadrilateral that encloses the gesture. Then a parallel constraint is applied by simply sketching its associated gesture over the two segments we want to make parallel (see 2.f, 2.g, 2.h). Once the desired shape has been obtained, we can proceed with dimensional control. A first action is to draw a dimension without the dimension text (see Fig. 2.i). This is interpreted by the application as a measure command, and the current value of that dimension is shown, as seen in Fig. 2.j. If the user wants to change the current dimension value, he or she writes the new value next to the current one. Then the system regenerates and displays the new geometry (Fig. 2.k and 2.l). In this way, the system provides a very natural form of imposing the desired dimensions over the sketch.

As can be seen, once the designer has introduced the complete outlined sketch, it can be edited, dimensioned and constrained. In other words, the interface offers some innovative ways of controlling the shape after a beautified constrained model is presented to the user.

The application manages two types of constraints and dimensions: “automatic” and “user defined”. Automatic constraints and dimensions are those provided by the system. The “user defined” ones are sketched by user. As can be seen in Fig. 2, the user can add new constraints drawing their associated gestures (Table 1) near the geometric entities where they must be applied. These gestures can be written by the user to impose some desired constraint. In this context, the scratch gesture can be used to remove undesired constraints. The automatic beautification process (automatic constraints and dimensions) is in charge of adjusting the input sketch in real time and provides an immediate feedback to the user, because it operates as the user draws the sketch.

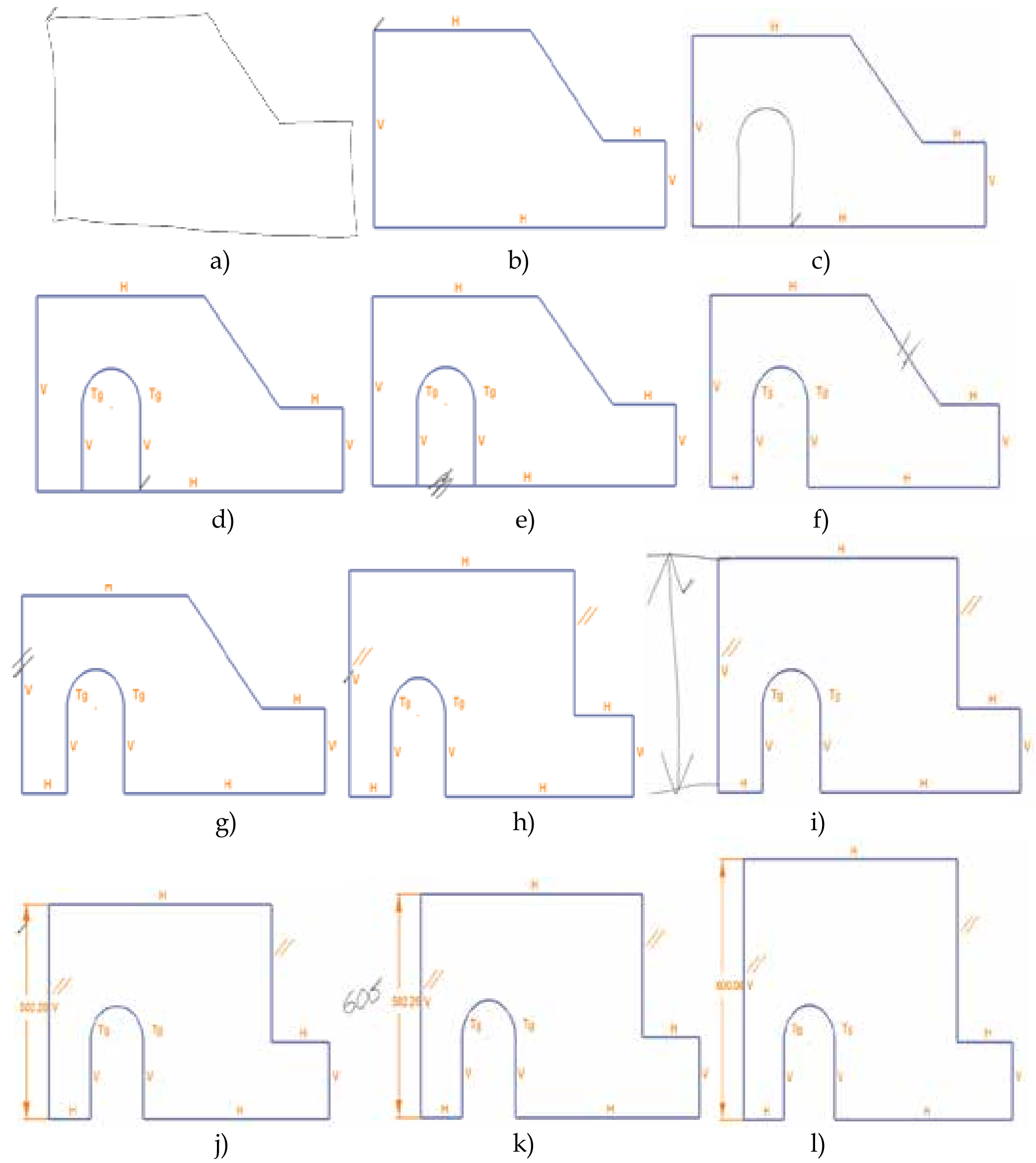


Fig. 2. Sketching sequence in ParSketch

The user can configure how the parametric engine controls the geometry. This control is implemented by a set of threshold values used to decide whether a geometric constraint is verified (see Fig. 3 for details) or not. The user has the possibility of enabling or disabling a specific constraint by an on/off selection box. Also it is possible to establish the order in which the constraints will be applied, using the "sequence" field in the dialog box presented in Fig. 3. These tolerance settings are intended to provide a tool for controlling the beautification action. Some of the supported constraints are: coincident (if a coincident constraint is defined between a point and any geometry then this implies that the point lies on the geometry), concentric, parallel, tangent, equal radius (it implies that the radii of the geometries are the same), perpendicular, equal distance (this constraint is used for search geometries with the same length), distance, angle and radius.

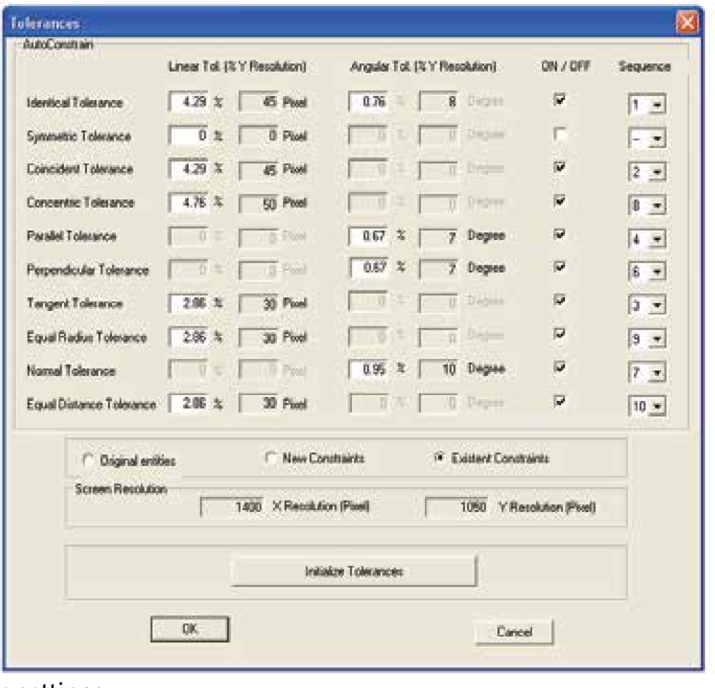


Fig. 3. Tolerance settings

Once a two dimensional section has been defined with the ParSketch module it is possible, using gestures, to make an extrusion or a revolution of the parametric section to create a 3D model. Then this process can continue sketching new 2D sections onto the faces of the generated object and applying the corresponding modelling gestures. In this second stage the command set includes the three gestures listed in Table 2.

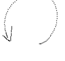
Modeling gestures	Class
	Extrusion
	Revolve-right
	Revolve-left

Table 2. Gesture alphabet for modelling operations implemented in GEGROSS

The application recognizes the type of stroke drawn by the user using the gestural recognizer (RecoGes). Fig. 4, Fig. 5 and Fig. 6 show examples of modelling with the GEGROSS application.
The system uses the geometric kernel ACIS (www.spatial.com) to store the geometric entities. The points of the stroke are taken by means of the Wintab API (www.pointing.com), which is an open interface that directly collects pointing input from a digitizing tablet and passes it to applications. This API allows retrieving additional information as the pressure the user applies at each point of the stroke over the tablet.

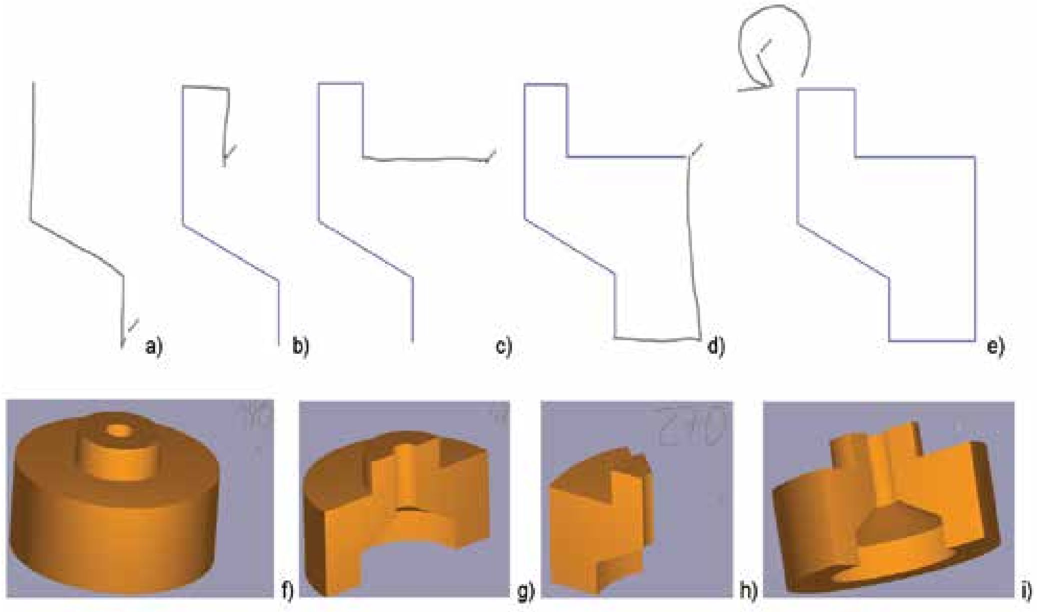


Fig. 4. Modelling sequence in GEGROSS. Example of revolution shape

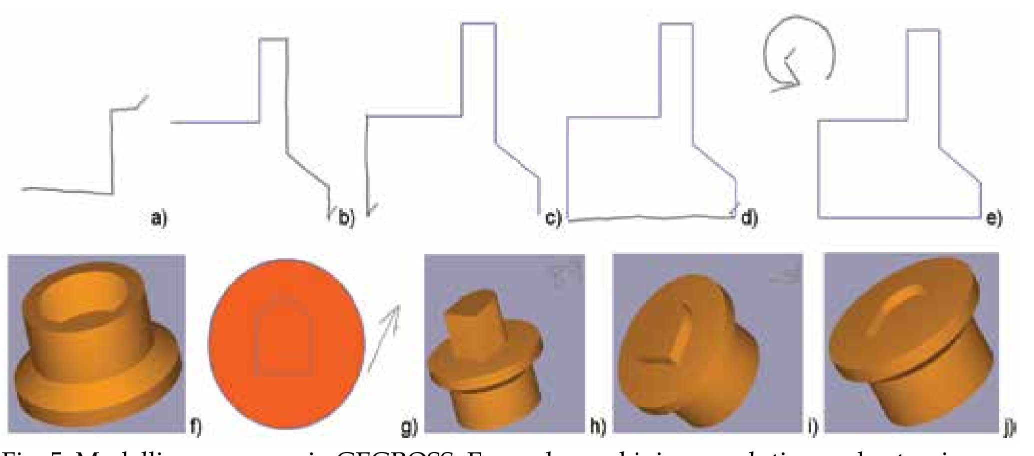


Fig. 5. Modelling sequence in GEGROSS. Example combining revolution and extrusion

4. Sketching-Based vs. WIMP Interfaces for Parametric Drawing

From a theoretical point of view we can show that if the sketching application supports complex strokes, i.e. strokes composed by several basic primitives as line segments and arcs (see Fig. 7 as an example) this means a potential advantage over WIMP interaction. For instance, analyzing sections composed exclusively by arcs and line segments, we can make an approximated calculation of the number of interactions required by a WIMP application to complete the drawing task.

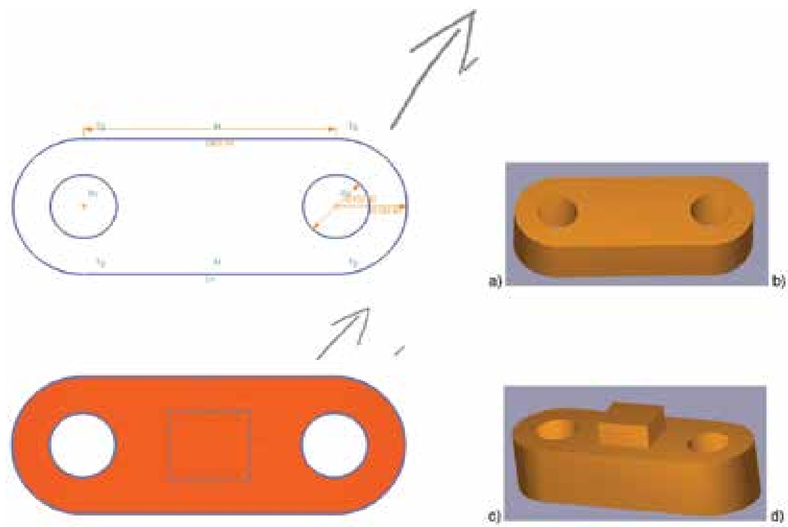


Fig. 6. Modelling sequence in GEGROSS. Example of extrusion-modelled shape

Usually one interaction is required to initiate the drawing process (one mouse click) and another one for finishing (a double click or pressing the enter key, for example). For drawing the line segments and tangent arcs in Fig. 7 two more interactions per elements are required: one is for defining the connecting vertex and the other for the selection of the proper geometric constraint as the horizontal, vertical, perpendicular or tangent conditions in this example. We count for this second interaction although, in modern parametric sketchers, geometric constraints are dynamically added as the user moves the drawing cursor. Only after the user detects the proper constraint is when he/she introduces the next entity vertex. This requires user attention, so we add it to the global number of interactions.

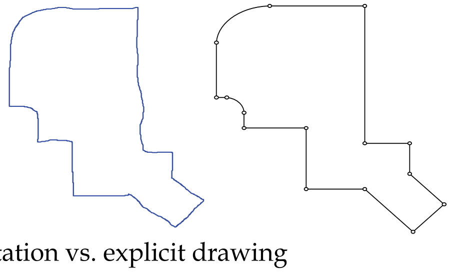


Fig. 7. Automatic segmentation vs. explicit drawing

The last term in the calculation of total number of interactions is related to primitive switching operation. When the user wants to link for example a tangent arc to a previous polyline, he/she must spend one interaction, providing this information to the system (using for example a contextual menu or icon selection), and then spend a second interaction to come back to polyline mode.
In sum, if n_l and n_a represents the number of line segments and arcs respectively, the total number of interactions N spent by the user is:

$$N = 2 + 2(n_l + n_a) + 2n_a \quad (1)$$

Even for not too complex figures (N= 36 in figure 4) the last equation shows that although a user could employ several strokes to complete the shape and require some corrections to overcome recognition errors by the sketching application, there is a wide margin to compete with WIMP-based interaction in terms of efficiency. So it is feasible to implement a robust geometric segmentation and recognition to keep advantage over WIMP interaction. We think that this is one of the keys for success in providing a real alternative or at least a complement to a WIMP interface. But as Igarashi and Zeleznik noted (Igarashi & Zeleznik, 2007), we must adapt the design of our applications to exploit the pen's intrinsic capacity for rapid, direct, modeless and expressive 2D input.
To improve segmentation results, our system can be adapted to each user way of sketching by means of a tolerance control panel previously described (see Fig. 3) that defines some key parameters for improving recognition. As explained before, the mode detection has been solved using the electronic pen pressure information, since GEGROSS is intended to be used by persons with basic engineering drawing skills.
In relation with other typical operations in a parametric 2D application, as imposing geometric constraints or performing dimensional control, the number of interactions required by both systems is similar. So we can conclude that from the efficiency point of view the sketch based approach is a viable option.

4.1 Usability Study

The usability of digital thinking sketches as opposed to traditional paper-and-pencil sketches was measured elsewhere (Company et al., 2006). In this analysis, we have centred our study in the user satisfaction component of usability (Hornbæk, 2006), following the usability definition provided by ISO 9241-11, where it stands for "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use".
As noted previously, the main design goals of the GEGROSS application are:

- Expeditious creation of shapes composed by polylines, arcs, and circles.
- Dimensional and geometric shape-control though the use of technical drawing conventions.

The evaluation involved six CAD instructors and six students with parametric CAD experience. All but one of the CAD instructors were male with an average age of 55. All students were male with an average age of 24.
We allowed 30 minutes for the evaluation, which had four parts: an overview of the system where some short videos showed the system operation, an instruction stage with a modified

version of GEGROSS that explicitly informs the user about the recognized entities or gestures (typically 10 minutes were employed in this training), a drawing task, and a final discussion with participants. After the discussion, users filled a questionnaire to evaluate GEGROSS and express their comments about it.
Each participant used a Toshiba Tecra M4. This Tablet PC has a 14,1" screen, with a resolution of 1400x1050 dots, and employs Ms Window XP Tablet-PC Edition.
We asked users to accomplish three drawing tasks using the ParSketch module. Shapes presented in Fig. 8 where used to propose several drawing exercises. The first exercise was to create a parametric section similar to the left shape of Fig. 8. The other two exercises employed the other shapes, and the users had to create the shape and impose some dimensional and geometric constraints.

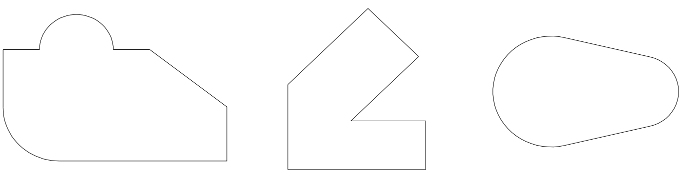


Fig. 8. Shapes for usability study

In relation with the evaluation of effectiveness, we have measured the accuracy and completeness with which users achieved the goals presented previously, using experts' assessment of the produced sketches. This assessment is based on the accuracy of the generated shapes. All the participants completed the requested drawing tasks satisfying all the conditions imposed to the generated shapes.
Efficiency has been measured taking into account the resources expended in relation to the accuracy and completeness with which users completed the drawing tasks. In our study we have used the task completion time and the number of events logged by a modified macro recording application. The most interesting result in this measure was the comparison between the best results obtained with the ParSketch module and the minimum number of interactions required by PTC's Pro/Engineer Wildfire 3 to complete the drawing tasks (this data are presented in Table 3). For all the participants in the study, this was their first contact with a Tablet-PC, and some of them had problems to control the pressure threshold that changes input mode.

Exercise	ParSketch (# of strokes)	Pro/E (# mouse click + # menu selection)
#1 (left)	3	12+4
#2 (middle)	1	8+1
#3 (right)	4	10+4

Table 3. Efficiency comparison

From Table 3 we can extract a first topic of discussion. Is it comparable the mental effort to generate a stroke on the Tablet-PC with the equivalent mouse operations to define the same geometry? We think that for users with previous experience in sketching on plain paper, drawing is practically an automatic task, which requires less concentration and effort than

the mouse operation. Perhaps this justifies that 100% of participants evaluated as easier, the use of the ParSketch module with respect to the CAD tools known by them.
Finally, user satisfaction has been measured using an adapted version of the QUIS Questionnaire (Chin,1998) using a 10 point scale from 0 to 9. A selection of the questions is presented in Table 4. In general, all participants expressed a very positive attitude towards the application, and all of them learnt in a few minutes to use it. Majority of comments about the system came from the pressure-based mode selection and about recognition errors. With respect to the pressure, none of participants had had previous experience with pressure sensible application and this had a distracting effect, requiring some concentration effort to change from the geometry input mode to the gesture one. We think that with more time of use, this mode change would not require so much effort. Also we are thinking about the convenience of providing some kind of online indicator (feedback) of what kind of input is receiving the system. Now, the application uses a paradigm similar to drawing on plain paper. Thickness of the rendered stroke in screen is related to the pressure done by the user while it is drawing. We are thinking on a color-based indication system that will represent geometry strokes in one color, and gesture strokes in another different one. This color assignment should be done dynamically, because in this way, if the user inadvertently begins to draw the stroke in the wrong mode he/she can correct it on the fly.

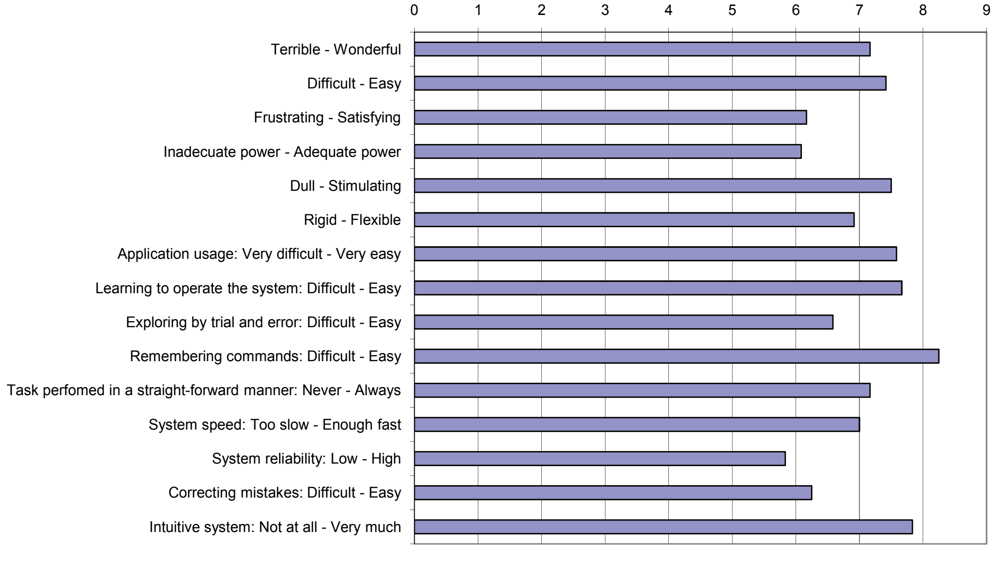


Table 4. User satisfaction measures

The other part of the comments about the system came from mistakes in the recognition process. This creates some kind of frustration in the user, when he/she draws a stroke or a gesture and a wrong interpretation is provided. The recognition rate for gesture recognition was 90 percent. Rates for geometry recognition were very variable, depending on the complexity of the generated stroke and the ability of the user creating the sketch.
In order to improve recognition results, we are studying the creation of a training mode in the application to adapt and tune recognition process to each user procedure of sketching.

5. Conclusion

In this chapter, an approach to create three-dimensional parametric objects using a gesture alphabet has been described. The main objective of this work has been to provide dimensional and geometrical control over the sections in an easy and natural way. The system offers a very simple interface to create parametric sections with an interesting possibility: “dimensional control”. In this way, user can impose some dimensional condition drawing the corresponding dimension and writing its value. This handwritten dimensions offer a natural and simple method to change dimension values that is known by any engineer.

Comparing the operation of the GEGROSS system with a standard WIMP parametric CAD application we can say that the basic functionality is practically equivalent. As can be seen in previous examples, usability is enhanced as well since the interface has been particularly tailored to detect standardized symbols. New symbols are not invented for any existing ones, but they are “borrowed” from the set of meaningful engineering symbols currently defined in the standards (ISO, ASME…). Therefore, the improvement in usability results from the fact that those symbols are commonplace for potential users: no learning is required, and unconscious user actions are readily interpreted by the computer.

In other words, learnability of GEGROSS has proven to be very high. Actually, users only have required ten minutes of introduction and demonstration before using the system. This in part is justified by the engineering background of participants. But GEGROSS has been specifically designed for this kind of users, trying to exploit their knowledge of technical drawing conventions and their sketching abilities. Perhaps this is one of the reasons of this positive reaction. Users feel that this tool adapts to them, not requiring a special effort for learning to use.

Preliminary tests have shown encouraging results and have concluded the GEGROSS application is a feasible alternative to current approach used in commercial CAD applications in order to create shapes of small or medium complexity. In that situation it presents a more effective modelling time, and it has been rated as easier to learn than comparable commercial applications. Users that have an engineering background find very natural the system behaviour, and the learning process to manage the application is very fast. Therefore, user satisfaction has been very high during the usability study. Users enjoy the simplicity of the system and its powerful control of geometry. However, improvements are needed to give a clearer feedback of pressure mode selection.

GEGROSS offers in many cases higher efficiency than a comparable WIMP application. This is much related to supporting complex strokes, i.e. strokes composed by mixed basic primitives as line segments and arcs, for defining the shape’s geometry. However a high efficiency in terms of complex stroke support can have an undesired side effect: worse system effectiveness because of the increasing difficulty of the recognition and segmentation tasks. So we can conclude, than the best alternative for getting the best results is the combination of several medium complex strokes, instead of trying to define the whole geometry in one only stroke. Besides, the user can take advantage of the edition strokes (erase and the like) to follow another good strategy: recursive refinement of a first rough version of the stroke. It has some advantages. First, reduces the fail rate of the system. Second, forces the user to sketch in a more convenient way: concentrating in the major shapes, and letting the details for subsequent refinements.

Acknowledgments. The Spanish Ministry of Science and Education and the European Union (Project DPI2007-01373) supported this work. It was also partially supported by the Portuguese Foundation of Science and Technology, under grant DecorAR - POSC/EIA/59938/2004.

6. References

Chin, J.P.; Diehl, V.A. & Norman, K.L. (1988). Development of an instrument for measuring user satisfaction of the human–computer interface. *Proceedings of ACM Conference on Human Factors in Computing Systems*, pp. 213–218, ISBN 0-201-14237-6, April 1998, ACM Press, New York, NY

Company, P.; Contero, M.; Naya F. & Aleixos, N. (2006). A Study of Usability of Sketching Tools Aimed at Supporting Prescriptive Sketches, *Proceedings of the 4th Eurographics Workshop on Sketch-Based Interfaces and Modeling*, pp. 139–146, ISBN 3-905673-39-8, September 2006, Eurographics, Vienna

Contero, M.; Naya, F.; Company, P.; Saorín, J.L. & Conesa J. (2005). Improving Visualization Skills in Engineering Education. *IEEE Computer Graphics and Applications*, Vol. 25, No. 5, (September/October 2005) 24-31, ISSN 0272-1716

De Araujo, B. & Jorge, J. (2003). BlobMaker: Free-form modelling with variational implicit surfaces, *Proceedings of 12° Encontro Português de Computação Gráfica*, pp. 17-26, October 2003, Porto

Fonseca, M. & Jorge, J. (2001). Experimental Evaluation of an On-Line Scribble Recognizer. *Pattern Recognition Letters*, Vol. 22, No. 12 (October 2001), 1311-1319, ISSN 0167-8655

Igarashi, T.; Matsuoka, S. & Tanaka, H. (1999). Teddy: a sketching interface for 3D freeform design, *Proceedings of ACM SIGGRAPH '99*, pp. 409-416, ISBN 0-20148-560-5, August 1999, publisher, Los Angeles, California

Igarashi, T. & Zeleznik, B. (2007). Sketch based interaction. *IEEE Computer Graphics and Applications*, Vol. 27, No. 1, (January/February 2007) 26–27, ISSN 0272-1716

Hornbæk, K. (2006). Current Practice in Measuring Usability: Challenges to Usability Studies and Research. *International Journal of Human-Computer Studies,* Vol. 64, No. 2, (February 2006) 79–102, ISSN 1071-5819

Naya, F.; Contero, M.; Aleixos, N. & Company, P. (2007). ParSketch: A Sketch-Based Interface for a 2D Parametric Geometry Editor. *Lecture Notes in Computer Science*, Vol. 4551, (August 2007) 115-124, ISSN 0302-9743

Pereira, J.; Jorge, J.; Branco, V. & Nunes, F. (2000). Towards calligraphic interfaces: sketching 3D scenes with gestures and context icons, *Proceedings of WSCG'2000 Conference*, pp. ,ISBN 80-7082-612-6, February 2000, Skala V. Ed., Plzen

Schweikardt, E. & Gross, M.D. (2000). Digital Clay: deriving digital models from freehand sketches. *Autom. in Construction*, Vol. 9, No. 1(Jan. 2000) 107-115, ISSN 0926-5805

Turner, A.; Chapmann, D. & Penn, A. (2000). Sketching space. *Computers & Graphics*, Vol. 24, No. 6, (December 2000) 869-879, ISSN 0097-8493

Tversky, B. (2002). What Do Sketches Say About Thinking?. AAAI Spring Symposium Series - Sketch Understanding, pp. 148-152, ISBN 978-1-57735-153-5

Zeleznik, R.C.; Herndon, K.P. & Hughes, J.F. (1996). SKETCH: an interface for sketching 3D scenes, *Proceedings of SIGGRAPH '96*, pp. 163-170, ISBN 0-89791-746-4, August 1996, ACM, New Orleans

5

Having Fun at Work: Using Augmented Reality in Work Related Tasks

Susanna Nilsson
Linköping University
Sweden

1. Introduction

Imagine driving down a windy road on a dark and rainy night. The rain makes it nearly impossible to see more than a few feet in front of the car. Now image that through this rain you can suddenly see the lines dividing the lanes as clearly as on a nice day. Or imagine opening a box full of parts that supposedly are the parts of furniture that you just bought from one of the bigger furniture houses in northern Europe. You try to figure out which part is which and how to follow the very simple instructions included in the box. You struggle to hold the paper in one hand and identifying where to put the pieces together with your other hand. Now instead imagine that you put on a pair of glasses and when you open the box you se an arrow pointing at the first piece you need to pick up and then you see how another part in the box is highlighted and how an image of that part is moved through the air showing you how, and with what tool, to put that part together with the first part. In this way you are guided with virtual instructions all the way until you finally can start putting books on the finished shelf or things in the finished drawer. This may sound a bit futuristic but actually it is a very real way of presenting information through a technique called Augmented Reality.

Augmented Reality (AR) is a technology that aims at merging the real and the virtual world, and thereby enhancing, or augmenting, the user's perception of the surrounding environment in varying aspects, as the examples above illustrate. Today there is a wide range of industrial and military AR applications, as well as applications for more entertaining or informative purposes. AR can be used in many different ways, not only to increase productivity in an assembly process (although there may be such effects), but also for the fun of it. However there is a lack of user focus in the development and evaluation of AR systems and applications today. This chapter aims at giving a short introduction to the technology of AR in general but foremost to give an example of an end user AR application used in a regular work related task in the natural environment of the user. We also discuss how usability should be addressed when designing a system without a traditional desktop interface. Current methods within the field of human-computer interaction (HCI) are largely based on findings from cognitive and perceptual theories, focusing on performance and quantitative usability measures. An AR-system, or any other novel mode of interaction is likely to receive poor results in an evaluation based on traditional HCI theories and methods, which are more favourable towards interaction methods the user already is familiar with. Therefore, other ways of evaluating and measuring user experience are more relevant for new ways of interacting with technology. For example, user experience in terms of enjoyment is a large part of user acceptance of a product, deciding if it actually is going to be used or not. In this chapter we describe how new technologies such as AR can be evaluated from a holistic perspective, focusing on the subjective user experience of the system.

2. Background: Augmented Reality and its applications

Augmented Reality is part of a field of technologies usually described as Mixed Reality (MR), which Milgram and Kishino (1994) described as a continuum of real and virtual information (see figure 1).

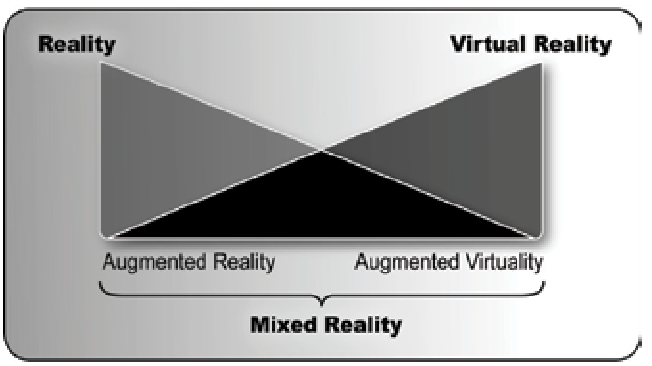


Figure. 1. The Mixed Reality continuum (Milgram & Kishino 1994)

The term 'mixed reality' aims at capturing the conceptual ideas behind the technologies used – the blending, merging or mixing of different realities. Even though it may be an interesting question, this chapter will not discuss the notion of 'reality' or go in to a philosophical debate about what is real and what is not. For the purpose of this text 'reality' is simply what humans perceive as their surrounding in their normal life. Given this definition of reality, 'merging realities' could simply mean merging two images of the world together. However, Azuma (1997) put some constraints on the term and mentions three criteria that have to be fulfilled for a system to be classified as an AR system: they all combine the real and the virtual, they are supposedly interactive in real time (meaning that the user can interact with the system and get response from it without delay), and they are registered and aligned in three dimensions. As an example, motion pictures with advanced 3D effects might have elements of AR, but they are not interactive so they do not qualify in the AR category.

AR applications can be found in diverse domains such as medicine, the military, entertainment and infotainment, technical support and industry applications, distance operation and geographical applications. A common application using AR is to provide instructions on how to operate new or unfamiliar equipment, or how to assemble a more or less complicated object. Tang (2003) describes an experimental evaluation of AR used in object assembly, Zauner et al. (2003) describe how AR can be used as an assembly instructor for furniture applications. The ARVIKA project illustrated several different applications for development, production and service (Friedrich, 2004).

2.1 Augmented Reality technology

The concept of merging realities in this way is not a novel idea of the 21st century. Military history, for instance, illustrates the use of blending real and virtual information by the use of predicted impact points when aiming a weapon, where static markers on a glass lens guides the aim of the shooter. When it comes to allowing a human to perceive immersion into virtual realities, Ivan Sutherland proposed the idea of an "ultimate display" already in 1965. The idea was to let a person wear a display and through that display see, and be in a virtual world. In fact a few years later he proposed and built the first head mounted display prototype (Sutherland, 1968). Technology has come a far way since the late sixties, and nowadays head mounted displays are easy to get a hold of and projecting images into these displays is relatively easy to do. The technology combining or merging real and virtual information does still create some difficulties however. One of the most important issues in AR is tracking or registration - a process necessary in order to align virtual objects with the real world as seen though the display. This chapter will not to great detail deal with technical issues of realising the concept of AR as other sources do a much better job of this (see for instance Kiyokawa, 2007). For an understanding of the technology and its possibilities and limitations some information must however be provided regarding tracking and hardware choices.

How to merge realities

Using a head mounted display there are principally two different solutions for merging reality and virtuality in real time today – video see-through and optic see-through (Azuma, 1997, Kiyokawa, 2007). In optic see-through AR, the user has a head mounted optical see-through display that allows the user to see the real world as if through a glass lens (Kiyokawa, 2007). The virtual information is then overlaid on the see-through display. Although the technique of blending virtual and real information optically is simple and cheap compared to other alternatives, this technique is known to cause some problems. For one, the virtual projection cannot completely obscure the real world image - the see-through display does not have the ability block off incoming light to an extent that would allow for a non-transparent virtual object. This means that real objects will shine through the virtual objects, making them difficult to see clearly. The problem can be solved in theory, but the result is a system with a complex configuration. There are also some issues with placement of the virtual images in relation to the surroundings in optic see through displays. Since the virtual objects presented to the user are semi-transparent they give no depth clues to the user. Instead the virtual objects seem to be aligned along the same focal plane whereas in natural vision, objects are perceived in different focal planes (Gustafsson et al., 2004; Haller et al., 2007).

A way to overcome some of the problems with optic see-through is by using a technique commonly referred to as video see-through AR, where a camera is placed in front of the users' eyes (see figure 2). The captured camera image is then projected to a small display in front of the users' eyes (Azuma, 1997; Kiyokawa, 2007). The virtual images are added to the real image before it is projected which solves the problem with the semitransparent virtual images described above, as well as gives control over where the virtual objects are placed.

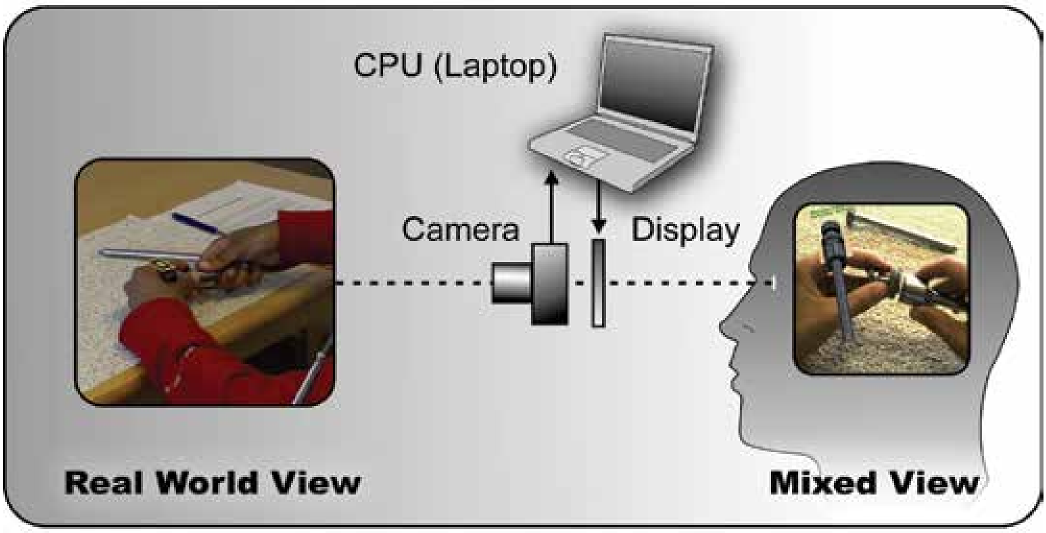


Fig. 2. A schematic view of a video see-through Augmented Reality system.

The video resolution of the camera and display sets the limit for what the user perceivesm The cameras and displays used today offer high-resolution images but unfortunately a field of view that is very limited compared to natural vision. A problem with video based solutions is that there is an eye offset due to the fact that the camera's position can never be exactly where the eyes are located, which gives the user a somewhat distorted experience, since the visual viewpoint is perceived to be where the camera is (Azuma, 1997). The difference between the bodily perceived movement and the visual movement as seen through the display can have effect on the user experience of the system, in some cases even causing motion sickness (Stanney, 1995). Despite these problems there are important vantage points with the video-see through solution. One has already been pointed out – the ability to occlude real objects - and another is that the application designer has complete control over the presented image in real time since it is run through the computer before it is presented to the user. In the optic see through design only the user will see the final augmented image. To conclude; there is a trade-off between the optic see through systems and the camera based systems, and the available resources often determine the choice of solution.

Marker tracking

Regardless of what display solution has been chosen for an AR application the most important issues to solve is how and where to place the virtually generated image. In order to place the virtual information correctly, the AR system needs to know where the user and user view point is. This means that the system has to use some kind of tracking or registration of the surrounding environment. There are different techniques to do this and several of them can be combined to ensure more reliable tracking of the environment (Haller et al., 2007). Tracking is normally done by using different sensors to register the surrounding environment. This sensor information is then used as a basis for placing the virtual information (Azuma et al., 2001). When using video see-through technique the AR system is already equipped with a visual sensor – the camera – which allows for vision based tracking. This tracking technique is one of the most commonly used today and it makes use of visual markers that can be recognized by feature tracking software (Kato & Billinghurst, 1999). Figure 3 below shows an example of a marker that can be used for this purpose.

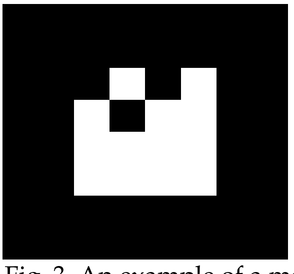


Fig. 3. An example of a marker for vision based tracking.

The marker tracking technique used in the studies presented in this chapter is based on ARToolkit, which is an open software library for building AR applications (Kato & Billinghurst, 1999). By tracking and identifying markers placed in the environment the algorithm calculates the position of the camera relative the marker and hence the virtual information can be placed in the display relative to the marker position. The images seen in figure 2 illustrate how a marker (seen on the user's finger) is used to place the virtual 3D object in the user's field of view.

3. Usability measures and the study of users in Augmented Reality

Although several projects in the AR domain strive to include an end-user perspective there are still few commercially available AR systems, and the research has mainly been focused on technological advances. Very few papers report on results of user studies or HCI evaluations (Bowman et al., 2002; Swan & Gabbard, 2005). Despite the potential of the technology, the research has still primarily focused on prototypes in laboratories, mainly due to the constraints of the hardware currently available to implement the systems (Livingston, 2005). This is also a reason why there are so few end user studies of AR techniques – the hardware constraints also limit the human factor research in the area. Still there have been a few user studies published, and the results point in the same general direction; there are several usability problems that are normally explained by hardware limitations, and despite these problems users respond positively to the use of AR for several different applications (see for example Bach & Scapin, 2004; Haniff & Baber, 2003; Nilsson & Johansson, 2006). There are other issues than hardware that affect the user experience of the

AR system, and these issues may become easier to identify when apparent hardware related issues (such as motion sickness, limited field of view and the lack of depth perception etc) are solved.

The methods used to study AR systems described in the existing literature are mainly based on usability methods used for graphical user interfaces, sometimes in combination with usability for VR applications (Träskbäck et al., 2003; Gustafsson et al., 2005; Dünser et al., 2007). This approach has some complications since it is not based on the experiences from actual AR systems users in actual contexts (Nilsson & Johansson, 2006; 2007b). Usability criteria and heuristics that are considered to be useful for designing new AR systems tend to be general, broad criteria, such as the ones Nielsen presented in his list of usability heuristics in 1993 (Nilsson & Johansson, 2006; Dünser et al., 2007). Usability methods such as cognitive task design (Hollnagel, 2003) where the design approach is based on observations of how a user completes a task in which the system or artefact is involved, also have to deal with the so called 'envisioned world problem' (Woods & Roth, 1998; Hollnagel & Woods, 2005). The 'envisioned world' problem states that even if a good understanding of a task exists, the new design/tool will change the task, rendering the first analysis invalid.

Designing systems based on heuristics developed for computer-based applications may be common practice in the AR field, but there are few examples of studies on how users actually perceive the system in actual use situations. During user studies in a smaller research project users were asked about their experience of the AR system, and none of them even mentioned desktop or laptop computers, or programs when describing what they were interacting through or with (Gustafsson et al., 2005). Instead, words like robot, video game and instructor were used to describe the interaction. The AR system was thus perceived as introducing other properties to interaction than "normal" desktop applications. This could hardly be attributed to the content of the interaction (which mainly was simple instructions of operation), but rather to the fact that the content was presented directly in the context of use. This of course raised questions of how useful it really is to base design of AR systems on desktop computer metaphors and usability criteria.

3.1 User acceptance of technology

When new technologies are introduced into a domain it may affect the user and the task on both a practical and a social level. The process of change requires knowledge, not only about the system introduced but also about the domain. The technical system or interface which is introduced should have as much positive effect on the user and her work as possible, while at the same time minimizing the negative effects of the system both for the user and other individuals.

Fundamental usability awareness implies that the interface or system should not be harmful or confusing to the user, but rather aid the user in her tasks. However, traditional usability guidelines, such as the ones presented by Nielsen (1993) or Shneiderman (1998) often do not include the context of use, the surrounding and the effect the system or interface may have in this respect. Being contextually aware in designing an interface means having a good perception of who the user is and where and how the system can and should affect the user in her tasks.

Davis (1989; 1993) describes two important factors that influence the acceptance of new technology, or rather information systems, in organizations. The perceived usefulness of a system and the perceived ease of use both influence the attitude towards the system, and hence the user behaviour when interacting with the system, as well as the actual use of the system (see figure 4).

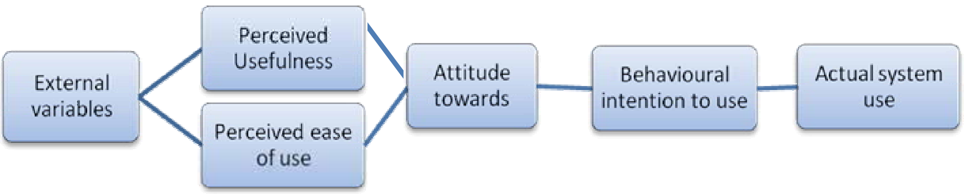


Fig. 4. The Technology Acceptance Model derived from Davis, 1989 (Nilsson & Johansson, 2007b).

If the perceived usefulness of a system is considered high, the users can accept a system that is perceived as less easy to use, than if the system is not perceived as useful. For AR systems this means that even though the system may be awkward or bulky, if the applications are good, i.e. useful, enough, the users will accept it. Equally, if the AR system is not perceived useful, the AR system will not be used, even though it may be easy to use.

3.2 Cognitive Systems Engineering as a basis for analysis[1]

Traditional approaches to usability and human computer interaction assumes a de-composed view with separate systems of humans and artifacts. As noted previously, the idea of the human mind as an information processing unit which receives input and generates output has been very influential in the domain of human computer interaction. A basic assumption in the information processing approach is that cognition is studied as something isolated in the mind. A problem with many of these theories is that they mostly are based on laboratory experiments investigating the internal structures of cognition, and not on actual studies of human cognition in an actual work context (Neisser, 1976; Dekker & Hollnagel, 2004).

A holistic approach to human-machine interaction has been suggested by Hollnagel and Woods called ‘cognitive systems engineering’ (CSE) (Hollnagel & Woods 1983; 2005). The approach is loosely based upon findings and theories from, among others, Miller et al. (1969) and Neisser (1976). The core of this approach is the questioning of the traditional definition of cognition as something purely mental: “Cognition is not defined as a psychological process, unique to humans, but as a characteristic of system performance, namely the ability to maintain control. Any system that can maintain control is therefore potentially cognitive or has cognition” (Hollnagel & Woods, 2005).

In the CSE approach is important to see the system as a whole and not study the parts in isolation from each other. The cognitive system can be comprised of one or more humans interacting with one or more technical devices or other artifacts. In this cognitive system, the human brings in the ‘natural cognition’ to the system and artifacts or technological systems may have ‘artificial cognition’. Hollnagel and Woods (2005) uses the notion ’joint cognitive system‘ (JCS) to describe systems comprised of both human and technological components

[1] This section has in parts previously been presented in Nilsson & Johansson 2007a.

that strive to achieve certain goals or complete certain tasks. The JCS approach thus has a focus on function rather than structure, as in the case of information processing, which is the basis for most traditional HCI. A CSE approach to humans and the tools they use thus focus rather on what such a system does (function) rather than what is (structure). A consequence of that perspective is that users should be studied when they perform meaningful tasks in their natural environments, meaning that the focus of a usability study should be user performance with a system rather than the interaction between the user and a system. A design should thus be evaluated based on how users actually perform with a specific artifact, but the evaluation should also be based on how they experience that they can solve the task with or without the artifact under study.

Using tools or prosthesis

As stated above, the main constituents in a JCS are humans and some type of artifact. Hutchins (1999) defines cognitive artifacts as "physical objects made by humans for the purpose of aiding, enhancing, or improving cognition". Hollnagel and Woods (2005) define an artifact as "something made for a specific purpose" and depending upon this purpose and how the artifact is used, it can be seen as either as a tool or as prosthesis. A tool is something that enhances the users' ability to perform a task or solve problems. Prostheses are artifacts that take over an already existing function. A hearing aid is a prosthesis for someone who has lost her/his hearing while an amplifier can be a tool for hearing things that normally are too quiet to be heard. Another example is the computer which is a very general tool for expanding or enhancing the human capabilities of computation and calculation, or even a tool for memory support and problem solving. But the computer can also be used not only to enhance these human capabilities but also to replace them when needed. A computer used for automating the locks of the university buildings after a certain time at night has replaced the human effort of keeping track of time and at the appropriate time going around locking the doors. The way someone uses an artifact determines if it should be seen as a tool or prosthesis, and this is true also for AR systems. AR systems are often very general and different applications support different types of use. So as with the computer, AR systems can be used either as tools or as prostheses, which can have effect on the perceived usability and hence the appropriate design of the system. It is very rare to evaluate a computer in general – usability evaluations are designed, and intended to be used for specific applications within the platform of the computer. This should also be the case for AR systems – to evaluate and develop usability guidelines for the general AR system platform is both impossible and pointless. Evaluating the AR applications however is necessary to ensure a positive development of future AR systems so that they better support the end user applications.

4. Two examples of end user applications

In this section two end user studies are described as examples of AR applications developed and evaluated in cooperation with the end users. The studies are grounded in the core CSE idea that users should be studied in their natural environment while solving meaningful tasks. The AR applications developed for these user studies were both developed in cooperation and iteration with an experienced operating room nurse and a surgeon. This professional team of two described problematic issues around which we used the AR technology to aid them in performing the task of giving instructions on two common medical tools.

The basic problem for both applications was how to give instructions on equipment both to new users, and to users who only use the equipment at rare occasions. Normally a trained professional nurse would give these instructions but this kind of person-to-person instruction is time consuming and if there is a way to free up the time of these professional instructors (nurses) this would be valuable in the health care organisation. The AR applications were therefore accordingly aimed at simulating human personal instructions. It is also important to note that the aim of these studies was not to compare AR instructions with either personal instructions or paper manuals in any quantitative measures. The focus was not speed of task completion or other quantitative measures, the focus was instead on user experience and whether or not AR applications such as the ones developed for these studies could be part of the every day technology used at work places like the hospital in the studies. The results from the studies have been reported in parts in Nilsson & Johansson 2006, 2007b and 2008.

4.1 The first study

The specific aim of the first study was to investigate user experience and acceptance of an AR system in an instructional application for an electro-surgical generator (ESG). The ESG is a tool that is used for electrocautery during many types of surgical procedures. In general electrocautery is a physical therapy for deep heating of tissues with a high frequency electrical current. The ESG used in this study is used for mono- or bipolar cutting and coagulating during invasive medical procedures (see figure 5). When using this device it is very important to follow the procedure correctly as failing to do so could injure the patient. Part of the task is to set the correct values for the current passing through the device, but most important is the preceding check up of the patient before using the tool – the patient cannot have any piercings or any other metal devices on or in the body, should not be pregnant, and most importantly, the areas around the patient must be dry as water near electrical current can cause burn injuries.

The AR system used in this study included a tablet computer (1GHz Intel®Pentium® M, 0.99GB RAM) and a helmet mounted display with a fire wire camera attached. The camera was also used for the hybrid tracking technology based on visual marker tracking (Gustafsson et al., 2005). A numeric keypad was used for the interaction with the user (see insert, figure 5).

Fig. 5. To the right, the helmet-mounted Augmented Reality System. To the left, the electro-surgical generator prepared with markers for the tracking.

A qualitative user study was conducted onsite at a hospital and eight participants (ages 30 – 60), all employed at the hospital, participated in the study. Four of them had previous experience with the ESG, and four did not. All of the participants had experience with other advanced technology in their daily work. First the participants were interviewed about their experience and attitudes towards new technology and instructions for use. Then they were observed using the AR system, receiving instructions on how to start up the ESG. After the task was completed they filled out a questionnaire about the experience.
The instructions received through the AR system were developed in cooperation and iteration with an experienced operating room nurse at the hospital in a pre study. The instructions were given as statements and questions that had to be confirmed or denied via the input device, in this case a numeric keypad with only three active buttons – 'yes', 'no', and 'go to next step'. An example of the instructions from the participants' field of view can be seen in figure 6.

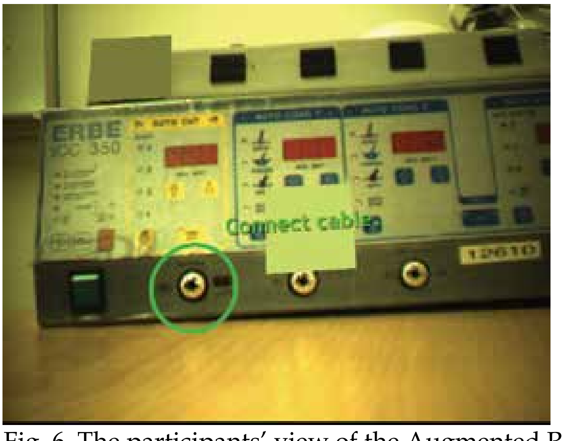


Fig. 6. The participants' view of the Augmented Reality instructions.

Data was collected both through observation and open ended response questionnaires. The questionnaire consisted of questions related to overall impression of the AR system, experienced difficulties, experienced positive aspects, what they would change in the system and whether it is possible to compare receiving AR instructions to receiving instructions from a teacher.

Results of the first study[2]
It was found that all participants but one could solve the task at hand without any other help than by the instructions given in the AR system. In general the interviewed responded that they preferred personal instructions from an experienced user, sometimes in combination with short, written instructions, but also that they appreciated the objective instructions given by the AR system. The problems users reported on related both to the instructions given by the AR system and to the AR technology, such as problems with a

[2] A detailed report on the results of the study is presented in Nilsson & Johansson, 2006.

bulky helmet etc. Despite the reported problems, the users were positive towards AR systems as a technology and as a tool for instructions in this setting.
All of the respondents work with computers on a day to day basis and are accustomed to traditional MS Windows™ based graphical user interfaces but they saw no similarities with the AR system. Instead one respondent even compared the experience to having a personal instructor guiding through the steps: "It would be if as if someone was standing next to me and pointing and then… but it's easier maybe, at the same time it was just one small step at a time. Not that much at once."
Generally, the respondents are satisfied with the instructions they have received on how to use technology in their work. However, one problem with receiving instructions from colleagues and other staff members is that the instructions are not 'objective', but more of "this is what I usually do". The only 'objective' instructions available are the manual or technical documentation and reading this is time consuming and often not a priority. This is something that can be avoided with the use or AR technology – the instructions will be the same every time much like the paper manual, but rather than being simply a paper manual AR is experienced as something more – like a virtual instructor.
The video based observation revealed that the physical appearance of the AR system may have affected the way the participants performed the task. Since the display was mounted on a helmet there were some issues regarding the placement of the display in from of the users' eyes, so they spent some time adjusting it in the beginning of the trial. However since the system was head mounted it left the hands free for interaction with the ESG and the numerical keypad used for answering the questions during the instructions. As a result of the study, the AR system has been redesigned to better fit the ergonomic needs of this user group. Changes have also been implemented in the instructions and the way they are presented which is described in the next study.

4.2 The second study

The second study referred to here is a follow up of the first study. The main differences between the studies are the AR system design and the user task. The AR system was upgraded and redesigned after the first study was completed (see figure 5). It included a head mounted display, an off the shelf headset with earphones and a microphone and a laptop with a 2.00 GHz Intel®Core™ 2 CPU, 2.00 GB RAM and a NVIDIA GeForce 7900 graphics card. Apart from the hardware, the software and tracking technique are basically the same as in the previous study. One significant difference between the redesigned AR system and the AR system used in the first study is the use of voice input instead of key pressing. The voice input is received through the headset microphone and is interpreted by a simple voice recognition application based on Microsoft's Speech API (SAPI). Basic commands are OK, Yes, No, Backward, Forward, and Reset.
The task in this study was also an instructional task. The object the participants were given instructions on how to assemble was a common medical device, a trocar (see fig 7). A trocar is used as a port, or a gateway, into a patient during minimal invasive surgeries. The trocar is relatively small and consists of seven separate parts which have to be correctly assembled for it to function properly as a lock preventing blood and gas from leaking out of the patient's body.

Fig. 7. To the left, the separate parts of a trocar. Tot he right, a fully assembled trocar.

The trocar was too small to have several different markers attached to each part. Markers attached to the object (as the ones in study 1) would also not be realistic considering the type of object and its usage – it needs to be kept sterile and clean of other materials. Instead the marker was mounted on a small ring with adjustable size that the participants wore on their index finger (see figures 8 a and b).

Fig. 8a) The participant's view in the HMD. b) A participant wearing the head mounted display and using the headphones and voice interaction to follow the AR instructions.

Instructions on how to put together a trocar are normally given on the spot by more experienced nurses. To ensure realism in the task, the instructions designed for the AR application in this study was also designed in cooperation with a nurse at a hospital. An example of the instructions and animation can be seen in figure 8a. Before receiving the assembly instructions the participants were given a short introduction to the voice commands they can use during the task; OK to continue to the next step, and back or backwards to repeat previous steps.

Twelve professional nurses and surgeons (ages 35 – 60) at a hospital took part in the study. The participants were first introduced to the AR system. When the head mounted display and headset was appropriately adjusted they were told to follow the instructions given by the system to assemble the device they had in front of them.

As in the previous study, data was collected both through direct observation and through questionnaires. The observations and questionnaire was the basis for a qualitative analysis. The questionnaire consisted of 14 statements to which the users could agree or disagree on a 6 point likert scale, and 10 open questions where the participants could answer freely on their experience of the AR system. The questions related to overall impression of the AR system, experienced difficulties, experienced positive aspects, what they would change in the system and whether it is possible to compare receiving AR instructions to receiving instructions from a teacher.

Results of the second study[3]

All users in this follow-up study were able to complete the task with the aid of AR instructions. The responses in the open questions were diverse in content but a few topics were raised by several respondents and several themes could be identified across the answers of the participants. Issues, problems or comments that were raised by more than one participant were the focus of the analysis.

Concerning the dual modality function in the AR instructions (instructions given both aurally and visually) one respondent commented on this as a positive factor in the system. Another participant had the opposite experience and considered the multimedial presentation as being confusing: "I get a bit confused by the voice and the images. I think it's harder than it maybe is".

A majority among the participants were positive towards the instructions and presentation of instructions. One issue raised by two participants was the possibility to ask questions. The issue of feedback and the possibility to ask questions are also connected to the issue of the system being more or less comparable to human tutoring. It was in relation to this question that most responses concerning the possibility to ask questions, and the lack of feedback were raised.

The question of whether or not it is possible to compare receiving instructions from the AR system with receiving instructions from a human did get an overall positive response. Several of the respondents actually stated that the AR system was better than instructions from a teacher, because the instructions were "objective" in the sense that everyone will get exactly the same information. When asked about their impressions of the AR system, a majority of the participants gave very positive responses and thought that it was "a very interesting concept" and that the instructions were easy to understand and the system as such easy to use. A few of the participants did 7however have some reservations and thought it at times was a bit tricky to use.

The result of this study as well as the previous study, indicate that the acceptance of AR instructions in the studied user group is high. To reconnect with the idea of measuring the usefulness of a system rather than just usability the second study also included questions about the use of AR as a supportive tool for learning how to assemble or use new technology, both in work related tasks and in other situations. The users were in general very positive as these diagrams illustrate:

[3] A detailed report on the results of the study is presented in Nilsson & Johansson, 2007b.

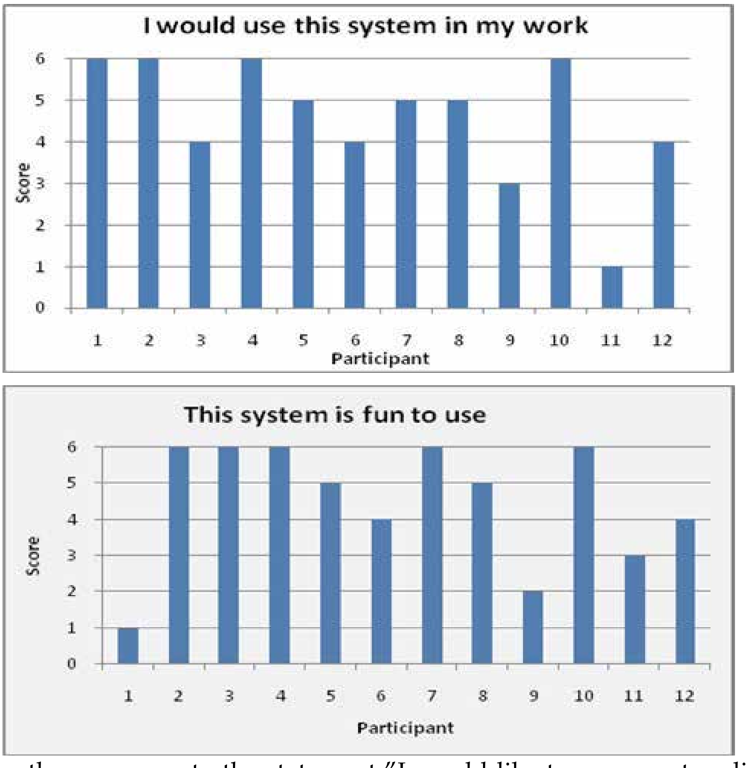


Fig. 9. a) Top, the responses to the statement ”I would like to use a system like this in my work”. b) Bottom, the responses to the statement ”This system is fun to use” (6 is the most positive grade and 1 the most negative for further details see Nilsson & Johansson 2007b).

As can be seen in the graph to the left in figure 9 one of the participants definitely does not want to use this kind of system in their work, while four others definitely do want to use this kind of system in their work. Interestingly enough one participant, who would like to use the AR system at work, does not find it fun to use (see figure 9 above). In general though the participants seem to enjoy using the system and this may be an indicator that they see it as a useful tool in their normal work tasks.

4.3 Lessons learned in the two user studies

The overall results from both studies shows a system that the participants like rather than dislike regardless of whether they received instructions in two modalities or only one. Both studies indicate that the participants would like to use AR instructions in their future professional life. Despite some physical issues with the AR system all users but one did complete the task without any other assistance. However, effects of the physical intrusion of the system upon the users' normal task should not be ignored. Even if the system is lightweight and non-intrusive, it still may change the task and how it is performed. This may not be a problem in the long run – if the system is a positive influence on the task, user and context, it will with time and experience grow to be a part of the task (much like using computers have become part of the task of writing a paper).

Interactivity is an important part of direct manipulating user interfaces and also seems to be of importance in an AR system of the kind investigated in these studies. A couple of the participants who were hesitant to compare AR instructions to human instructions, motivated their response in that you can ask and get a response from a human, but this AR system did not have the ability to answer random questions from the users. Adding this type of dialogue management in the system would very likely increase the usability and usefulness of the system, and also make it more human-like than tool-like. However, this is not a simple task, but these responses from real end users indicate and motivate the need for research in this direction. Utilizing knowledge from other fields, such as natural language processing, has the potential to realize such a vision.

In a sense AR as an instructional tool apparently combines the best from both worlds – it has the capability to give neutral and objective instructions every time and at the same time it is more interactive and human like than paper manuals in the way the instructions are presented continuously during the task. But it still has some of the flaws of the more traditional instructional methods – it lacks the capability of real-time question-answer sessions and it is still a piece of technical equipment that needs updates, upgrades and development.

5. Concluding discussion

AR is a relatively new field in terms of end user applications and as such, the technological constraints and possibilities have been the driving forces influencing the design and development. This techno-centred focus has most likely reduced the amount of reflection that has been done regarding any consequences, other than the technical, of introducing the technology in actual use situations. The impact of the way AR is envisioned (optic see-through and video see-through) has largely taken focus off the use situation and instead lead to a focus on more basal aspects, such as designing to avoid motion sickness and increasing the physical ergonomics of the technology. However, these areas are merely aspects of the platform AR, not of the applications it is supposed to carry and the situations in which they are supposed to be used. Studies of AR systems require a holistic approach where focus is not only on the ergonomics of the system or the effectiveness of the tracking solutions. The user and the task the user has to perform with the system need to be in focus throughout the design and evaluation process. It is also important to remember that it is not always about effectiveness and measures – sometimes user attitudes will determine whether or not a system is used and hence it is always important to look at the actual use situation and the user's attitude towards the system.

The purpose of the system is another important issue when evaluating how useful or user-friendly it is – is it intended for pleasure and fun or is it part of a work setting? If it is somewhat forced on the user by it being part of everyday work and mandatory tasks, the system needs to reach efficiency standards that may not be equally important if it is used as a toy or entertainment equipment. If the system is a voluntary toy the simplicity factor is more important than the efficiency factor. On the other hand, if a system is experienced as entertaining, chances are it may actually also be perceived as being easier to use. It is not a bold statement to claim that a system that is fun and easy to use at work will probably be more appreciated than a system that is boring but still reaches the efficiency goals. However, as the technology acceptance model states – if the efficiency goals are reached (i.e. the users find it useful) the users will most likely put up with some hassle to use it anyway. In the case of the user studies presented in this chapter this means that if the users actually feel that the AR instructions help them perform their task they may put up with some of the system flaws, such as the hassle of wearing a head mounted display or updating software etc, as long as the trade off in terms of usefulness is good enough.

As discussed previously in the chapter there is a chance that the usability focused methodology measures the wrong thing – many interfaces that people use of their own free will (like games etc) may not score high on usability tests, but are still used on a daily basis. It can be argued that other measures need to be developed which are adapted to an interface like AR. Meanwhile, the focus should not be on assessing usability but rather the experienced usefulness of the system. If the user sees what he can gain by using it she will most likely use it despite usability tests indicating the opposite.

The field of AR differs from standard desktop applications in several aspects, of which the perhaps most crucial is that it is intended to be used as a mediator or amplifier of human action, often in physical interaction with the surroundings. In other words, the AR system is not only something the user interacts with through a keyboard or a mouse. The AR system is, in its ideal form, meant to be transparent and more a part of the users perceptive system than a separate entity in itself. The separation between human and system that is common in HCI literature is problematic from this point of view. By wearing an AR system the user should perceive an enhanced or augmented reality and this experience should not be complicated. Although several other forms of systems share this end goal as well, AR is unique in the sense that it actually changes the user's perception of the world in which he acts, and thus fundamentally affects the way the user behaves. Seeing virtual instructions in real time while putting a bookshelf together, or seeing the lines that indicate where the motorway lanes are separated despite darkness and rain will most likely change the way the person assembles the furniture or drives the car. This is also why the need to study contextual effects of introducing AR systems seems even more urgent. When evaluating an AR system, focus has to be on the goal fulfilment of the user-AR system rather than on the interface entities and performance measures gathered from evaluation of desktop applications. This approach is probably valid in the evaluation of any human machine system, but for historical reasons, focus often lays on only one part of the system.

AR as an interaction method for the future is dependent on a new way of addressing usability – if the focus is kept on scoring well in usability tests maybe we should give up novel interfaces straight away. But if the focus is on the user's subjective experience and level of entertainment or acceptance, AR is an interactive user interface approach that surely has a bright future.

References

Azuma, R. (1997) A survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*. 6: 4, pp. 355-385

Azuma, R, Bailot, Y., Behringer, R. Feiner, S., Simon, J. & MacIntyre, B. (2001) Recent

Advances in Augmented Reality. *IEEE Computer Graphics and Applications*. 21: 6, pp. 34-47

Bach, C. & Scapin, D.L. (2004) Obstacles and Perspectives for evaluating Mixed Reality systems usability. *Proceedings of the IUI-CADUI Workshop on Exploring the Design and Engineering of Mixed Reality Systems* (MIXER). Funchal, Island of Madeira, January 13th

Bowman, D., Gabbard, J. & Hix, D. (2002) A survey of usability evaluations in virtual environments: trade classification and comparison of methods. *Presence: Teleoperators and Virtual Environments* 11: 4, pp. 404–424

Davis, F. D. (1989) Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly,* 13: 3 pp. 319-340

Davis, F.D. (1993) User acceptance of information technology: system characteristics, user perceptions and behavioural impacts. *The International Journal of Man-Machine Studies* 38, pp. 475–487

Dekker, S. & Hollnagel, E. (2004) Human factors and folk models. *Cognition Technology and Work* 6: 2, pp. 79-86

Dünser, A. Grasset, R., Seichter, H., & Billinghurst, M. (2007) Applying HCI Principles in AR Systems Design. *2nd International Workshop on Mixed Reality User Interfaces: Specification, Authoring, Adaptation (MRUI 2007)*

Friedrich, W. (2004) *ARVIKA, Augmented Reality für Entwicklung, Produktion und Service.*, Erlangen, Germany: Pubilcis Corporate Publishing

Gustafsson, T., Carleberg, P., Nilsson, S., Svensson, P., Sivertun, Å., & LeDuc, M. (2004) Mixed Reality för tekniskt stöd, FOI-R--1198--SE, (Technical report ISSN 1650-1942)

Gustafsson, T., Carleberg, P., Svensson, P., Nilsson, S., & Sivertun, Å. (2005) Mixed Reality Systems for Technical Maintenance and Gaze-controlled interaction. FOI-R.1695-SE (Technical report ISSN 1650-1942)

Haller, M., Billinghurst, M., & Thomas, B. (eds) (2007) *Emerging Technologies of Augmented Reality: Interfaces and Design.* London, UK: Idea Group Publiching

Haniff, D.J. & Baber C. (2003) User evaluation of augmented reality systems. *Proceedings of the Seventh International Conference on Information Visualization (IV'03)* 16-18 July, pp. 505- 511

Hollnagel, E ed (2003) *Handbook of Cognitive Task Design*. New Jersey: Lawrence Erlbaum Associates, Inc

Hollnagel, E. & Woods, D.D. (1983) Cognitive Systems Engineering: New wine in new bottles. *International Journal of Human-Computer Studies* 51, pp. 339-356

Hollnagel, E. & Woods, D.D. (2005) *Joint Cognitive Systems. Foundations of Cognitive Systems Engineering*. Boca Raton, Forida: CRC Press, Taylor & Francis Group

Hutchins, E. (1999) Cognitive Artefacts. in Wilson, R.A. and Keil, F.C. *The MIT Encyclopedia of the Cognitive Sciences (MITECS)*, Cambridge, Massachusetts: Bradford Books, pp. 126-127

Kato, H. & Billinghurst, M. (1999) Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System. *Proceedings of the 2nd International Workshop on Augmented Reality (IWAR 99).* 20-21 October, San Francisco, USA.

Kiyokawa, K. (2007) An Introduction to Head Mounted Displays for Augmented Reality. In Haller, M., Billinghurst, M., Thomas, B. (eds): *Emerging Technologies of Augmented Reality: Interfaces and Design.* London, UK: Idea Group Publishing, pp. 43-63

Livingston, M. (2005) Evaluating Human Factors in Augmented Reality Systems. *IEEE Computer Graphics and Applications* 25: 6, pp. 6-9

Milgram, P. & Kishino, F. (1994) A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems*, E77-D: 12

Miller, G. A., Galanter, E. & Pribam, K. H. (1969) *Plans and the Structure of Behavior*. Holt, Rinehart & Winston, New York

Nielsen, J. (1993) *Usability Engineering*, San Diego: Academic press

Nilsson, S. & Johansson, B. (2006) A cognitive systems engineering perspective on the Design of Mixed Reality systems. *Proceedings of the 13th European Conference on Cognitive Ergonomics* September 20-22, Zürich 2006, pp. 154-161

Nilsson S. & Johansson B. J.E. (2007a) A Systemic Approach to Usability in Mixed Reality Systems. *Proceedings of the 13th Australia and New Zealand Systems Conference*, 2-5 December, Auckland, New Zealand

Nilsson, S & Johansson, B.J.E. (2007b) Fun and Usable: Augmented Reality Instructions in a Hospital Setting. *Proceedings of the Australasian Computer - Human Interaction Conference*, 28-30 November, Adelaide, Australia

Neisser, U. (1976) *Cognition and Reality*. W. H. Freeman: San Fransisco.

Shneiderman B. (1998) *Designing the user interface. Strategies for effective human-computer interaction*. Reading, Massachusetts: Addison-Wesley

Stanney, K. (1995) Realizing the full potential of virtual reality: human factors issues that could stand in the way, *Proceedings of the Virtual Reality Annual International Symposium (VRAIS'95)* March 11-15, pp.28-34

Sutherland, I. (1968) A head-mounted three-dimensional display. *Proceeding of the Fall Joint Computer Conference. AFIPS Conference Proceedings* 33, Arlington, Virginia: AFIPS, pp. 757-764

Swan, J. E., II, & Gabbard, J. L. (2005) Survey of user-based experimentation in augmented reality. *1st International Conference on Virtual Reality*, July 22-27, Las Vegas, Nevada

Tang, A. (2003) Comparative Effectiveness of Augmented Reality in Object Assembly. *Proceedings of the SIGHI conference on Human Factors in Computing Systems* April 5-10, Ft Lauderdale, Florida, USA

Träskbäck, M., Koskinen, T. & Nieminen, M. (2003) User-Centred Evaluation Criteria for Mixed Reality Authoring Applications. *Proceedings of Human Computer Interaction*, 22-27 June, Crete, Greece, vol. 3, pp. 1263-1267

Woods, D.D. & Roth, E.M (1988) Cognitive engineering: human problem solving with tools. *Human Factors* Volume 30, pp. 415 - 430

Zauner, J., Haller, M. & Brandl, A. (2003) Authoring of a Mixed Reality Assembly Instructor for Hierarchical Structures. *Proceedings of the 2nd IEEE and ACM International conference on Mixed and Augmented Reality (ISMAR*). 7-10 October, Tokyo, Japan, pp. 237-246

6

Multi-Device Design in Contexts of Interchange and Task Migration

Rodrigo de Oliveira, Heloísa Vieira da Rocha
State University of Campinas (UNICAMP)
Brazil

1. Introduction

With the miniaturization of digital components and the vast technological development of the past years, society has remarked the redefinition of "personal computers" by the advent of modern mobile devices. Together with the innovation brought by these handhelds, there was also the challenge to develop multi-device interfaces for today's desktop applications. While some created mobile interfaces from scratch to get the best from the devices, others looked for automatic adaptations to reduce the load imposed to the designer. In both cases, the user wasn't the focus anymore, resulting interfaces so different from each other to the point of compromising usability when performing the same task on many devices.

The proposal described in this chapter claims that there is no multi-device approach capable to provide full usability in every context because the user may choose only one interface to access the application or interchange its use via many devices. In the first case, the user learns to perform tasks with only one device, which makes relevant an approach that takes advantage of its resources and solves its limitations. In the second, the user already knows one of the available interfaces, which generates an expectation for the others. Therefore, it is necessary to combine approaches with different goals and suit the user according to the appropriate context.

In this sense, we propose multi-device design via the maintenance of a Consistency Priorities hierarchy defined in three levels. The first two levels give support to the user's expectation in order to guarantee easiness of learning/remembering and safety in contexts of interchange (prone to task execution with different devices) and task migration (starting tasks with one device and finishing with another). On the other side, the third level provides task personalization according to the user's interest towards higher efficiency and satisfaction of use with a specific device.

We evaluated this proposal by means of a controlled experiment in which an e-learning desktop application was taken as a reference to design three pocket PC interfaces using different approaches: (1) *Direct Migration* to maintain exactly the same layout of the desktop interface; (2) *Linear Transformation* to personalize and adequate the desktop interface to the handheld; (3) and *Overview* applying the first two levels of the Consistency Priorities hierarchy. All participants executed common tasks using each of the three mentioned interfaces.

The subjective evaluation results pointed the Overview approach as the best to maintain the user's mental model by preserving easiness, efficiency and safety of use for inter-device interaction. Additionally, both efficacy (task result accuracy) and efficiency (task average execution time) were the same or even better with this approach. On the other hand, users revealed their preference for the task personalization present in the Linear approach. This result gives support to our proposal, corroborating that the efficacy generated by the first two levels of the Consistency Priorities hierarchy (task perception and execution) should be combined with the third level of personalization. This could be done by letting designers create interface patterns and make them available to users during interaction. Such combination should guarantee usability while constantly accessing one application through the same device or in contexts of interchange and task migration.

This chapter is structured as follows: First, we review some relevant previous work in the area of multi-device design. Then, we describe our proposal and outline supportive theories in multidisciplinary fields. The implementation of these ideas are exemplified and evaluated in the next sections. Finally, we present our conclusions and future directions of research.

2. Related Work

Multi-device design has been addressed in many ways focusing the transition between desktop and mobile interfaces. Generally, this process involves automatic or manual transformations to remove images, reduce sizes, summarize texts, adapt orientation or restructure the whole information to better suit the handheld characteristics. In order to understand the collection of proposals presented recently in this research field, we suggest a division based in four categories: Hypertext Structure, Universal Controller, Adaptive Interface and Layout Consistency.

The *Hypertext Structure* category includes interfaces that outline the structure of related web pages using hypertext. This proposal has been implemented with automatic approaches that create hyperlinks matching the web site structure in a tree-based view. This way, users may first explore the document at high-level and only then visualize details about the information of interest. This visualization technique has proven to be useful in cases of limited bandwidth and processor power. First prototypes were developed for desktop browsing, like WebMap (Dömel, 1995) and WebTOC (Nation et al., 1997), and improved towards the mobile context with projects such as WebTwig (Jones et al., 1999) and Power Browser (Buyukkokten et al., 2000). Other proposals applied these ideas not only to one web site, but also to a set of them belonging to the news context (Banerjee et al., 2003).

The *Universal Controller* category envisions a totally different perspective for multi-device design, adapting handhelds' functionalities to exploit services discovered while entering new environments (*e.g.* controlling of lights, projector, stereo, *etc.*). Examples of this category include the architecture proposed by Hodes et al. (1997) and the ICrafter framework (Ponnekanti et al., 2001), both adequate to rigid ubiquitous environments. On the other hand, the PUC system (Nichols et al., 2002) has a more flexible structure for the mobile context by engaging in a two-way communication with everyday appliances, first downloading a specification of the functions and then translating protocols to automatically create remote control interfaces. Follow-on work had major upgrades in efficacy and efficiency whenever users had to execute tasks using interfaces consistent with their previous experience (Nichols et al., 2007).

The *Adaptive Interface* might be the most predominant category considering its vast number of proposals implemented using model-based design. The methodology builds specifications from an abstract declarative model of how the interface should behave and then automatically generate concrete user interfaces based on such model. Eisenstein et al. (2000) proposed techniques to help designers with the modeling process of platform, presentation and task structure. Lin (2005) also targeted the designers by creating a tool called Damask, which enables the design sketching using patterns optimized for each target device. Many authors implemented prototypes to automatically generate interfaces based in the abstract models (Bergman et al., 2002; Mori et al., 2003; Coninx et al., 2003; Gajos & Weld, 2004). Model extraction from already made applications was also addressed for the web domain (Gaeremynck et al., 2003) and graphical user interface reverse engineering (Santo & Zimeo, 2007). Although the adaptive interface category reduces the heavy load imposed to the developer, the generated interfaces can't guarantee a smooth inter-device transition in contexts of interchange and task migration, which have been considered the primary concern by many authors (Denis & Karsenty, 2004; Florins et al., 2004; Pyla et al. 2006). In fact, the experiments realized with adaptive interfaces tend to focus only on the automatic interface generation efficacy instead of horizontal usability issues.

At last, the *Layout Consistency* category is based on Overview transformations that preserve visual characteristics of the desktop layout. Some of the most used visualization techniques include the fisheye (Baudish et al., 2004), thumbnail (Milic-Frayling & Sommerer, 2002; MacKay et al., 2004; Lam & Baudisch, 2005) and focus + context (Roto et al., 2006). These proposals have revealed better easiness, efficiency, safety and satisfaction of use when compared to other automatic transformations, such as the Direct Migration and the Single Column. However, designers still need a well established theoretical model to guide them towards constructing these interfaces with better usability for multi-device contexts. The following sections describe our user-centered approach that addresses this issue for contexts of interchange and task migration.

3. User-Centered Multi-Device Design

3.1 Mental Model Update Cycle

Norman (1988) proposed a seven stage action model of how people execute tasks. Although it can't be considered a complete psychological theory (stages are not discrete, nor necessarily sequential and most behavior does not go through all stages), the main human cognitive processes involved are well highlighted, like *attention* to world objects, *decision making* to execute actions, *perception* of produced effects, *memory* analysis to interpret the world state and *learning* of final results. Fig. 1 adapts this model to a simplified version that focus on the user's mental model update stage.

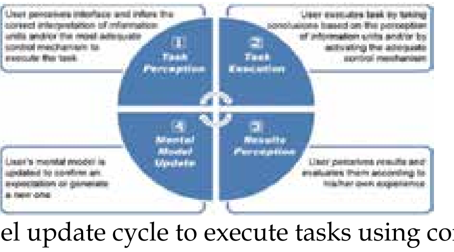


Fig. 1. – User's mental model update cycle to execute tasks using computer interfaces.

3.2 Consistency Priorities

According to the mental model update cycle presented in Fig. 1, the user's first step is to infer what should be the appropriate action towards the goal and, only then, actually execute it. This inductive inference based reasoning process usually contrasts the interface perception (*e.g.* visual, auditory, tactile, *etc.*) with the user's previous experience (mental model). As a result, a particular decision making is drawn according to the user's judgment, increasing the possibility to achieve the desired goal. However, when the system presents similar situations leading by inference to different conclusions, it is likely that the user will make mistakes and store ambiguous information in his/her mental model (forth stage). In order to avoid this, we propose that the application's interfaces should preserve the same perceptual characteristics (which constitute the inference process input) and have a consistent behavior in which one task can be executed following the same actions' flow on different devices maintaining the richness of the distinct interaction types involved. This proposal can be structured in the following Consistency Priorities:

1. *Task Perception:* Inter-device perceptual constancy[1] preserving size, shape and color of every control mechanism and information unit relevant to the task. Also, their relative localization within the interface should be maintained. If relevant differences are found between devices considering their:
   - *sizes*, interface should be adapted maintaining visibility;
   - *shapes*, interface should be adapted maintaining visibility and mapping;
   - *colors*, interface should be adapted maintaining visibility, mapping and feedback.

   Additionally, if the interaction types are incompatible (*e.g.* speech and graphical pointing interfaces), each control mechanism perception and its relative localization should be mapped to demand attention of the correspondent human sense;
2. *Task Execution:* Inter-device consistency of the actions' flow required to execute each user's task. If the control mechanisms had to be adapted by the task perception priority, the actions' flow should be preserved in a logic perspective to maintain the task model structure under a different implementation of the modeled interactions.

By adapting to the user's previous experience, the Consistency Priorities hierarchy shall contribute to multi-device design guaranteeing easiness of learning, remembering and safety of use in contexts of interchange and task migration. However, some users could choose only one device to access the application, thus reducing the concern with his/her experience. Additionally, the varied nature of these devices may restrict the application's executable tasks set, thus compromising efficiency and satisfaction of use with the first two consistency levels. We suggest a third consistency priority to balance the usability attributes:

3. *Task Personalization:* Ability to change both levels of task perception and execution according to the user's preferences and context of use. The goal is to achieve the best design which is the configuration that the user expects. In this sense, we encourage the development of interface patterns at the users' convenience. This priority is related to the personally consistent design concept (Nichols, 2006, p.86), but with an active position for the user. As a result, efficiency and satisfaction are guaranteed to both experts and novices, avoiding the downsides of consistent design (Grudin, 1989).

---

[1] Denotes the tendency of animals and humans to see familiar objects as having standard shape, size and colour regardless of changes in angle of perspective, distance, or lighting. Impression tends to conform to the object as it is or is assumed to be, rather than to the actual stimulus.

4. Applying the Consistency Priorities Approach

The implementation of this approach must be understood in the same context as the original application's design process. In this sense, Fig. 2 highlights the steps required in the lifecycle model towards applying the Consistency Priorities.

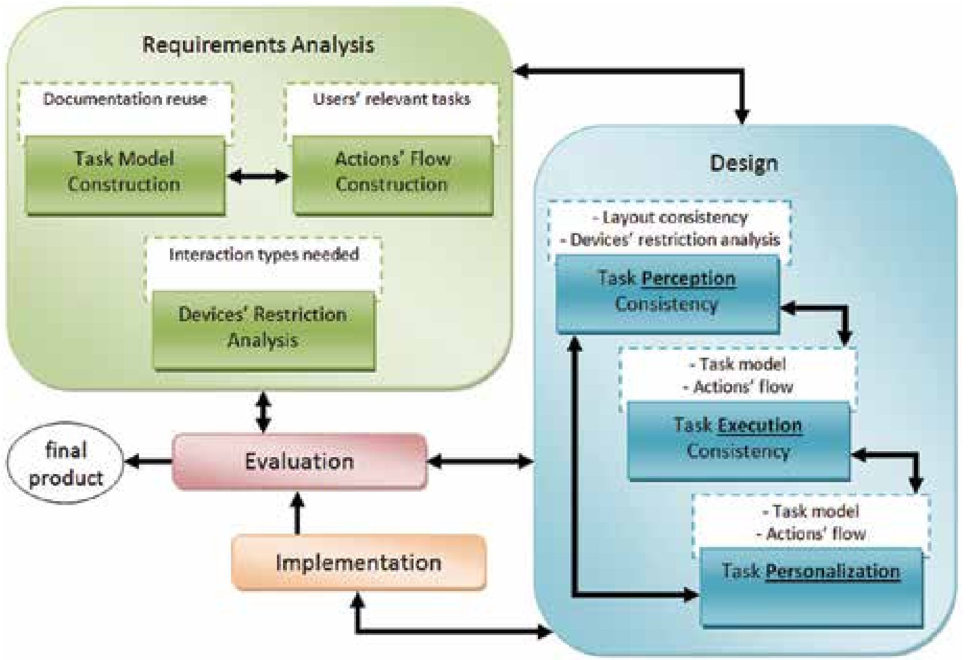


Fig. 2. – Interaction design lifecycle model adapted from Sharp et al. (2007) to focus on the Consistency Priorities implementation steps.

According to Fig. 2, the Interaction Design process can be divided in four main stages: Requirements Analysis, Design, Implementation and Evaluation. Moreover, the Consistency Priorities approach can be embedded in the model, reinforcing the importance of iteration. Pragmatically, we suggest applying this methodology by taking the following steps:

1. *Task Model Construction:* Representation of the user's tasks defined in high level, interaction tasks required to execute such user's tasks, their sequential steps and information units present on the interface;
2. *Actions' Flow Construction:* Description of the user's cognitive effort on relevant tasks concerning perception, execution and memory storage/retrieval activities;
3. *Devices' Restriction Analysis:* Comparison of the application access devices to identify relevant restrictions. This procedure is important to reveal the main design principles to be accounted on the next phase;
4. *Consistency Priorities Implementation:* Design of alternative interfaces following the three priorities of the consistencies hierarchy (perception, execution and personalization).

In order to ease the transition between theory and practice, this section presents an example applying the Consistency Priorities to design a pocket PC interface for a desktop application. Following, we present the chosen application, the task model elaborated for one of its tools, the actions' flow identification process, the restriction analysis for the target devices, and the implementation of two mobile interfaces adequate to different contexts (task migration and sole device access).

4.1 Application Domain

We chose Distance Learning to be the application's domain for this example due to its potential for dissemination and the availability of human resources to conduct experiments. Moreover, the application chosen was the TelEduc[2], an open-source e-learning environment used by more than 3000 institutions around the world, among schools, faculties, universities and companies. Fig. 3 shows a screen from this system designed for desktop.

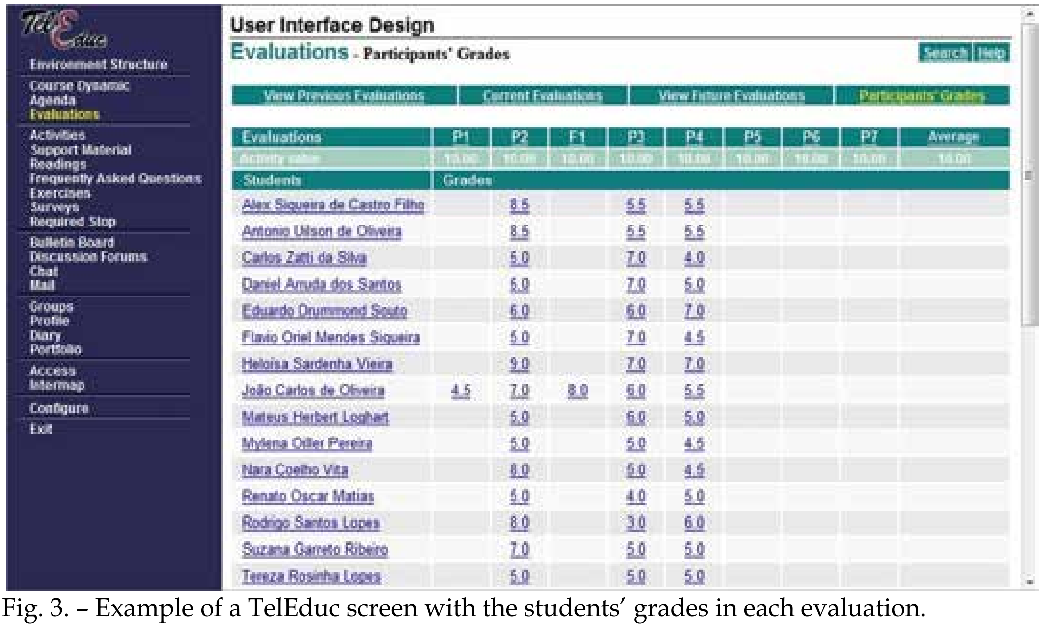


Fig. 3. – Example of a TelEduc screen with the students' grades in each evaluation.

4.2 Task Model

Building the task model is the first step of this methodology and its relevance is due to the fact that it describes interactive systems in terms of tasks needed to be executed towards the users' goals. Hence, the multi-device design process gains support to generate consistent interfaces directing the designers' focus to the system's requirements and behavior instead of implementation details for each platform. No specific notation is required, as long as the chosen language is able to model:

- User's tasks defined in high level;
- Interaction tasks required to execute the user's tasks;
- Sequential steps for the interaction tasks;
- Interface elements or information units present in the interaction.

[2] http://teleduc.nied.unicamp.br/teleduc

One way to start building this model is to considering each user's task and investigate every interaction needed to conclude them. Another way follows the reverse flow, describing all the interaction possibilities in each screen and growing the task model tree level by level until all functionalities were explored. Fig. 4 shows an example of the model built for the TelEduc's Evaluation tool using this latter approach under the ConcurTaskTree notation (Paternò et al., 1997). The result is a task model tree describing all the interaction tasks available on the screens related to this tool and is of great importance for this methodology's further steps, particularly for the Consistency Priorities implementation phase.

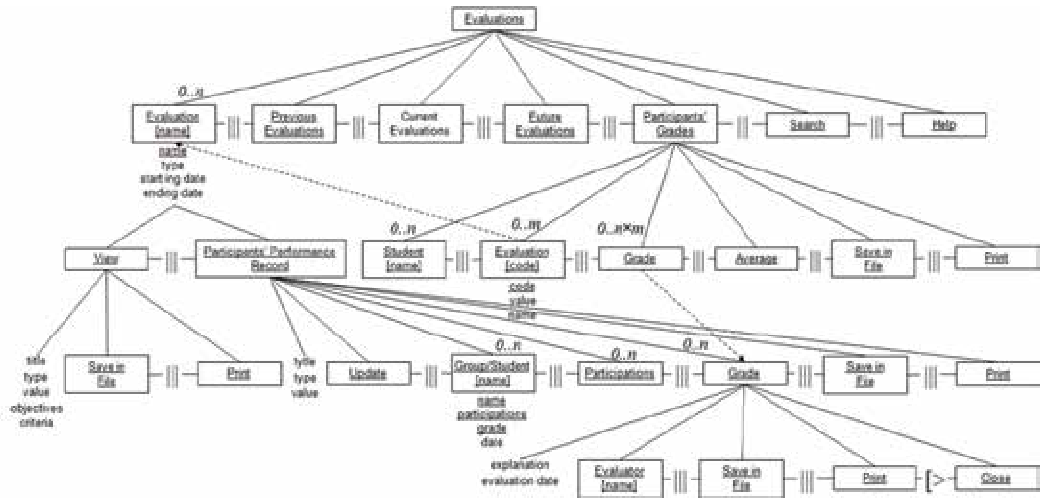


Fig. 4. – Task model of the TelEduc's Evaluation tool using the ConcurTaskTree notation with two extensions: (1) multiplicity in associations to avoid redundancy and (2) explicit declaration of interface elements (*e.g.* attributes like *name*, *type*, *starting date* and *ending date* present on the evaluations' screen with multiple interaction tasks called *Evaluation[name]*).

4.3 Actions' Flow

In this step, the actions' flow for each relevant task to the user should be specified. Again, it isn't necessary to use any specific notation, but it is of great interest to consider the activities listed in the user's mental model update cycle (see section 3.1). We suggest using the following terms:

- **perceive:** effort applied during the interval between searching the object of interest (control mechanism or information unit) and finding it. Every human sense might be involved in this search. Perception must be stored in memory in case the individual needs to use it after the interruption of its finding (see "**store**" below);
- **execute:** effort applied during the interval between decision making and activation of the perceived control mechanism;
- **store:** effort applied for temporary storage in short-term memory.

The actions' flow specification considering these activities assists in the process of task personalization (third level of the Consistency Priorities hierarchy) in which the designer will be concerned with choosing the user's most relevant tasks for their simplification towards better execution efficiency and satisfaction of use. Table 1 presents some examples of actions' flow specified for the TelEduc's Evaluation tool.

User's Task	Actions' Flow
Check evaluation's criteria (evaluation=*y*)	1. **execute perceive** Evaluations 2. **execute perceive** Evaluation[name=*y*] 3. **execute perceive** View 4. **return perceive** criteria
Check student's grade in evaluation (student=*x*; evaluation=*y*)	1. **execute perceive** Evaluations 2. **execute perceive** Participants' Grades 3. *c* **store perceive** Evaluation[name=*y*] 4. *aval* **store perceive** Evaluation[code=*c*] 5. *stud* **store perceive** Student[name=*x*] 6. **return perceive** Grade(*stud*, *aval*)
Check highest grade from the *n* students in evaluation *y*.	1. **execute perceive** Evaluations 2. **execute perceive** Participants' Grades 3. *c* **store perceive** Evaluation[name=*y*] 4. *aval* **store perceive** Evaluation[code=*c*] 5. for each Grade in *aval*'s column 5.1. *temp* **store perceive** Grade(*aval*) 5.2. if *temp* > *highest*, then *highest* **store** *temp* 6. **return** *highest*

Table 1. Example of user's tasks and corresponding actions' flow for the Evaluation tool.

4.4 Devices' Restriction Analysis

This step identifies main differences among target devices considering three attributes pointed by the perceptual constancy principle as the most relevant, *i.e.* size, shape and color. Although other attributes could also lead to ambiguous or erroneous perceptions when drastically changed (*e.g.* light, distance, weight, size, fluidity, flexibility, opacity, *etc.*), we expect that these three characteristics can model most devices in order to guide implementation of the task perception consistency priority.

In this sense, Table 2 presents concise comparative descriptions for a desktop, pocket PC and smartphone input/output (I/O) devices.

Device	Attribute	Desktop	Pocket PC	Smartphone	Relevance
Display	Color	24-bit	16/24-bit	16/24-bit	low
	Size	15", 800x600 pixels	3.5", 240x320 pixels	2", 240x320 pixels	high
	Shape	4:3	3:4 / 4:3	3:4 / 4:3	high/none
Keyboard	Color	variable	variable	variable	none
	Size	40cm/10cm	4.8cm/4cm	4cm/5cm	high
	Shape	QWERTY (hand adapted)	virtual QWERTY (pointing device adapted)	numeric/ QWERTY (thumb adapted)	medium

Table 2. Comparison among I/O devices of a standard desktop, pocket PC and smartphone.

According to Table 2, both input and output devices of a standard desktop have relevant differences in size compared to those available for the pocket PC and smartphone. Also, the pocket PC rotation on the palm of the hand to adjust its orientation regarding the desktop display has undesirable ergonomic implications, pointing the shape as another perceptual attribute to be considered for the interface adaptation. Thus, we expect to focus on visibility and mapping during the Consistency Priorities implementation for these target devices.

4.5 Implementation

Task Perception and Task Execution
The task perception consistency aims to preserve size, shape, color and relative localization of control mechanisms and information units available on interfaces. On the other side, the task execution consistency demands the same actions' flow to execute the user's tasks. A useful baseline to start implementing these consistency priorities is the Direct Migration approach, which consists of the desktop interface presented on the handheld device without any adaptation. According to the results obtained with the devices' restriction analysis in section 4.4, *size* and *shape* were the attributes with the most relevant differences between target devices, indicating that *visibility* and *mapping* shall be the design principles to focus on this consistency level (see section 3.2). The violation of these principles in the Direct Migration can be perceived by the intense interaction required with both vertical and horizontal scrolling to access information throughout the interface. If tasks are not visible, many usability attributes can be compromised, like utility, efficiency and safety of use.

A common solution to adapt desktop interfaces to the pocket PC screen is the Single Column feature, which is able to analyze and partition the web page structure presenting its content without the horizontal scrolling. However, this proposal can violate many task perception consistency requirements by changing relative localization of side menus and content area, losing visibility of the user's tasks and generating ambiguities on semantic mapping by reorganizing information units. These side effects are due to the fact that Single Column considers only the *shape* as an attribute with relevant difference between target devices. Therefore, we must also consider adaptations on *size*.

Among the information visualization techniques focusing on this attribute, we highlight the focus+context and the thumbnail (reduced replica of the desktop interface). Belonging to the latter, Smartview (Milic-Frayling & Sommerer, 2002) and Gateway (MacKay et al., 2004) are examples of proposals that let users first scan the thumbnail and then explore regions of interest. The main advantage is that visual mapping remains consistent with the user's previous experience, but the zoom-out rate makes content unreadable, as it can be noticed comparing Fig. 3 and Fig. 5a. In order to support visibility, the Gateway prototype presents readable texts of the thumbnail regions touched by the user, overlapping the readable region on the thumbnail (see Fig. 5b). However, data comparison tasks on the same interface might demand excessive memorization, besides the additional touch interaction for multiple regions of the same thumbnail. Also, mapping of table structures can be compromised as readable columns will be shown one at a time, losing the correlation between lines and columns.

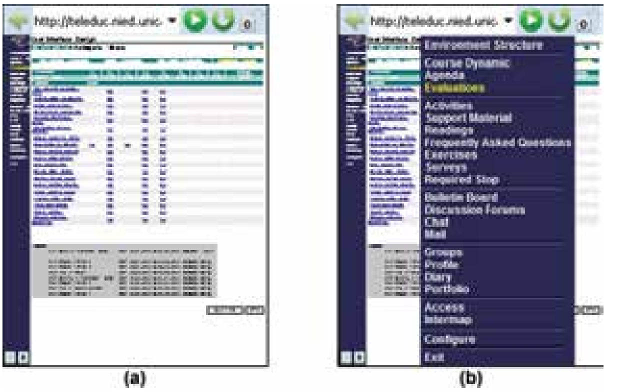


Fig. 5. – Example of the Gateway proposal applied for the TelEduc's participants' grades screen (see Fig. 3). In (a), a reduced non-functional replica of the desktop page; in (b), the TelEduc's side menu overlapped on the thumbnail after touching its region in (a).

We argue that these problems arise because only the attribute size was considered for adaptation. According to the devices' restriction analysis, both size and shape shall be adapted, thus requiring special focus on the visibility and mapping design principles. In this sense, the Summary Thumbnail (Lam & Baudisch, 2005) proposes a more adequate solution by improving the latter prototypes with text font increase and content summarization for the thumbnail. Fig. 6 shows an example of this adaptation for the TelEduc.

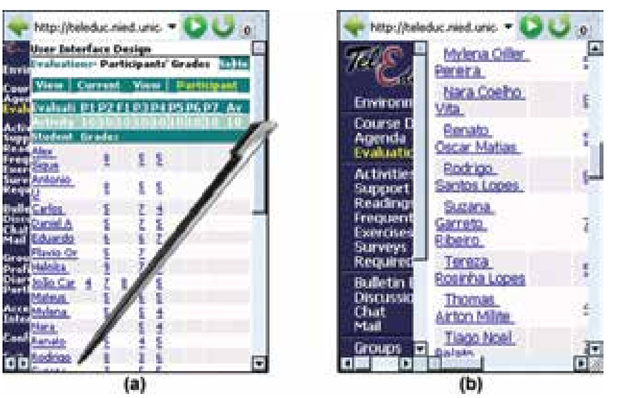


Fig. 6. – Example of the Summary Thumbnail proposal applied for the TelEduc's participants' grades screen (see Fig. 3). In (a), a reduced functional replica of the desktop page with readable and summarized texts; in (b), the detailed view of the region touched by the user's pen (full text and real size images as in the Direct Migration approach).

The interface presented in Fig. 6a reveals good similarity to its desktop equivalent (see Fig. 3) and also enhances the thumbnail legibility obtained with the Gateway (see Fig. 5a). Still, Summary Thumbnail fails to preserve mapping and consistency on task execution.
The mapping failure can be verified in the summarization of links "View Previous Evaluations" and "View Future Evaluations", resulting in two labels called "View" (menu

on top of Fig. 6a). This was due to the automatic summarization process, which crops characters from right to left until text fits within the available space. This methodology might lead to other unexpected results, especially for languages in which adjectives come after nouns (*e.g.* Portuguese, Spanish). This problem could be fixed using a conversion table built statically by the designers and containing every control mechanism label related to its most appropriate summarized form (non-hyperlink texts could still be summarized using the right to left cropping approach). A more dynamic solution could be based on a domain oriented summarization approach that would rip terms not relevant to the given page (low frequency on the page and high frequency in the database) and preserve only the important ones (high frequency on the page and low frequency in the database).

The task execution inconsistency comes from the ability to activate any navigation structure by pointing over it directly on the thumbnail. As the user needs to perform the same action to reveal its full text on the detailed view before deciding if that's the appropriate hyperlink to activate, there will always be an interaction ambiguity. This problem can be fixed substituting the Direct Migration detailed view by an overlapped window (thus avoiding many inconsistencies of such method as presented before) and eliminating the navigation ambiguity with a non-functional thumbnail. In this case, navigation could be provided by activating the full text hyperlink in the detailed view. This way, task execution remains consistent[3] by always revealing the detailed view whenever the user touches the thumbnail. Fig. 7 shows an example of how these adaptations can be implemented to preserve the first two levels of the Consistency Priorities hierarchy.

Fig. 7. – Example of the Consistency Priorities (first two levels) applied to the TelEduc's participants' grades screen (see Fig. 3). In (a), the Summary Thumbnail approach adapted to preserve visibility and mapping principles; in (b), the detailed view overlapped on the thumbnail avoiding context loss while switching views; and in (c), an example of navigation by activating the hyperlink inside the detailed view.

[3] Although task execution claims the interface to be consistent with both task model and actions' flow developed in sections 4.2 and 4.3, we argue that these changes to the Summary Thumbnail (including an extra interaction with the hyperlink inside the detailed view) don't break consistency at this level because it constitutes a new approach for handheld navigation. The concept implies that every object activated by the user reveals a detailed view with full text for summarized texts or real-size images for reduced images; but if the object is a hyperlink, an additional activation is required to actually follow the hyperlink (see Fig. 7c). Therefore, we state that the task model wasn't changed, but the navigation concept implicit on its tree nodes.

Task Personalization

This third level of the Consistency Priorities hierarchy focuses on users that won't access the application in contexts of interchange and task migration. In other words, they plan to access the application using only one device. Therefore, the concern with consistency in the first two levels of the hierarchy loses its relevance for this group of users. In this case, we suggest personalizing the interface with an active position for the user during interaction, which can be implemented in two ways:

- *Customization:* Ability to change perceptual aspects of control mechanisms' (*e.g.* enlarge fonts, shrink side menu, change structures' attributes like shape, color, *etc.*) and to reorganize information (*e.g.* hide images, menu items, table columns; add shortcuts; reveal full texts and/or descriptions; *etc.*);

- *Pre-built Interface Patterns:* Design of alternative interfaces with improved efficiency for tasks considered more relevant to a group of users. The original task model must be adapted towards reducing the actions' flow for such tasks, which could be done by removing leaf nodes, sub trees, or via a hierarchical rearrangement of child nodes as a result of their parent removal. The users' choice over pre-built interface patterns can be implemented by checking their profile on first interactions with the application.

The customization approach demands higher motivation to be accessed during interaction, reason why we encourage designers to build interface patterns that will require less effort by the users and still delegate them an active role in design. This personalization can be exemplified for the TelEduc's Evaluation tool considering the task of checking the students' grade for a given evaluation *y* (see Table 1):

1. **execute perceive** <u>Evaluations</u>
2. **execute perceive** <u>Participants' Grades</u>
3. *c* **store perceive** <u>Evaluation</u>[name=*y*]
4. *aval* **store perceive** <u>Evaluation</u>[code=*c*]
5. *stud* **store perceive** <u>Student</u>[name=*x*]
6. **return perceive** <u>Grade</u>(*stud*, *aval*)

If this task was considered the most relevant to the mobile user, than it should be personalized to reduce complexity and improve efficiency. The first step is to identify changes imposed by the context of use that could simplify the way tasks are currently executed. In the given example, we could assume that the mobile user is not interested in comparing grades, but rather prefer having a faster way to access his/her personal information. This assumption reduces the actions' flow to four simpler steps:

1. **execute perceive** <u>Evaluations</u>
2. **execute perceive** <u>Participants' Grades</u>
3. *aval* **store perceive** <u>Evaluation</u>[name=*y*]
4. **return perceive** <u>Grade</u>(*aval*)

The newer actions' flow removes search and memorization tasks from two information units: the evaluation (which used to associate a code to the evaluation's name) and the student (which required the identification of the adequate row on the students grades table). Fig. 8 shows how these changes reflect on the original task model (see Fig. 4) and Fig. 9a shows the interface obtained with this personalization.

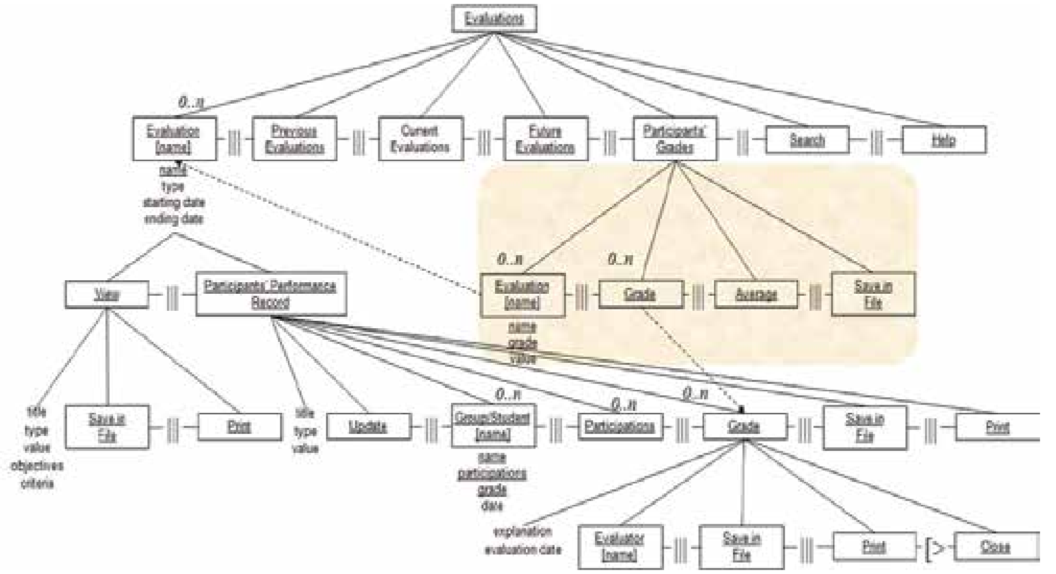


Fig. 8. – Personalization of the original task model (Fig. 4) for the TelEduc's Evaluation tool. The main focus was efficiency on the task of checking personal grades.

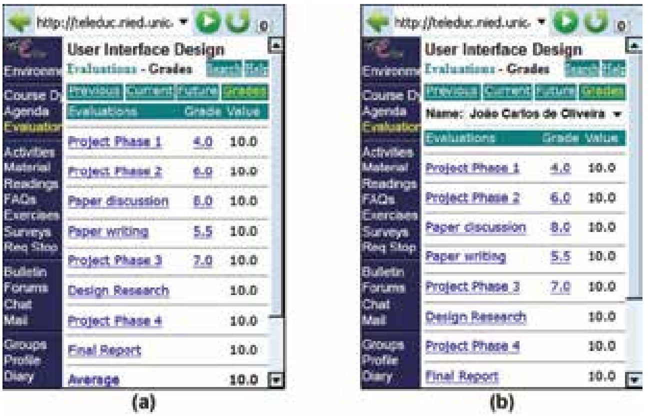


Fig. 9. – Interfaces obtained with the task personalization consistency priority for the TelEduc's Evaluation tool. While both screens consider the task of checking personal grades as the most relevant, in (a) the focus is restricted to the user, and in (b) grade comparison is enabled to balance interaction continuity and efficiency gain for the task.

As expected, the proposed interface for the mobile context described previously focus only on the user's personal information, thus preventing access to other students' grades (see Fig. 9a). However, many authors reinforced the idea that interaction continuity is a key element for multi-device design (Denis et al., 2004; Florins et al., 2004; Pyla et al., 2006; Hajdukiewicz, 2006). Hence, balancing interaction continuity and efficiency gain plays an important role in task personalization. Considering the TelEduc example, this could be done by preserving

the sub tree *Student[name]* (present in the original task model in Fig. 4) and implementing it as a choice structure, like a combo box (see Fig. 9b).
In this third consistency priority, almost every design decision conflicts with those taken in the previous levels. Basically, this happens due to the fact that task personalization means no compromise with the user's previous experience accessing the application via other devices. Additionally, if there isn't enough information to build personalized interface patterns, customization should be another important resource to enhance user experience.

5. Experiment

5.1 Domain, Application and Tasks

The motivation to choose Distance Learning as the experiment domain and TelEduc as the e-learning application was already presented in section 4.1. In order to choose one of its 21 tools as the experiment focus (*i.e.* agenda, evaluations, portfolio, *etc.*), we established the following requirements: the tool should have (1) higher access frequency in the computer science course taken by the population sample and (2) more relevant visualization information challenges for the desktop-handheld adaptation. While the first criterion was applied through an investigation of TelEduc's records, the second was based in the system's analysis, considering as challenges: the variable number of columns and rows inside tables, the need to show popup windows, the deeper menu hierarchy inside each tool and complex visual representations (*e.g.* graphs). Finally, the score combining both criteria for each tool revealed the Evaluations and the Portfolio as the most relevant ones. Therefore, the Evaluations tool was chosen to be the experiment focus because one of its challenges is very appropriate to highlight limitations in pocket PC's shape and size (*i.e.* the extensive matrix containing every students' grades on each test).

5.2 Participants

The experiment had 18 male computer science undergraduate students, ranging in age from 19 to 29 ($\bar{x} = 22$). They all had relevant experience with computers and the TelEduc e-learning system (used before in other seven courses in average). None had used it via a handheld, being their experience restricted to desktop/laptop/tablet PCs. When questioned about the devices they would like to use with TelEduc, only six showed interest in using more than one, which reveals an apparent indifference for task migration activities. On the other side, six subjects chose to access it solely by a desktop/laptop/tablet PC, pointing mobile interfaces as of the majority's interest to access the system (12 subjects). Participants were also questioned about their most frequent task with TelEduc's Evaluation tool. From a total of 15 answers, 12 indicated the checking of grades (two mentioning explicitly the comparison of grades) while three pointed the search for the evaluations' details.

5.3 Material

The experiment was conducted in a computer lab with wireless Internet connection and 18 tablet PCs available in individual desks. During evaluation, all the tablets remained laid or inclined on the desks and the pocket PC pen-based interaction was simulated by the tablet pen. Also, the pocket browser was reproduced in the tablet and equipments were connected to power outlets, which prevented interruptions by battery discharge.

5.4 Treatments and Procedures

The following treatments were applied to the experiment participants while executing a set of tasks to evaluate their effects and contribute to the investigation of the most appropriate multi-device design approach in a context of task migration:

- *Direct Migration:* Applied as a baseline, this treatment proposes a TelEduc pocket PC interface that is exactly like the desktop interface. Although consistency in task perception and execution is preserved, it can't guarantee usability principles like visibility, mapping and/or feedback (see Fig. 10a).
- *Linear Transformation:* A TelEduc pocket PC interface that adapts the desktop version to the handheld constraints, aiming efficiency in tasks of major interest to the user besides the preservation of task model characteristics. This approach loses consistency in task perception and keeps it partially for task execution, like many current approaches (*e.g.* most adaptive interfaces in the related work). Undesirable automatic transformation residues were avoided by manually designing the screens (see example in Fig. 10b).
- *Overview Transformation applying Consistency Priorities:* Adaptation of the TelEduc desktop interface for the pocket PC preserving the first two levels of the Consistency Priorities to focus in contexts of interchange and task migration (see Fig. 10c).

Fig. 10. – Experiment interfaces to visualize the TelEduc participants' grades using the three available treatments: (a) Direct Migration, (b) Linear Transformation and (c) Overview transformation applying Consistency Priorities.

Sample Partition

The 18 participants were fully distributed in six groups of three participants and each group followed a different treatments' application sequence, covering every possible combination. Thus, the residual effects of every treatment application over the other were balanced.

Studied Variables

In order to contrast pros and cons of each treatment applied in each task execution, we studied quantitative (execution time and task accuracy) and qualitative variables (easiness, efficiency and safety subjectively compared to the desktop interface). After finishing each task with the three approaches, users identified the best and the worst interfaces for that task. At last, after the execution of all tasks, they also filled a satisfaction questionnaire identifying their favorite interface and the reasons for that.

Tasks

Subjects had to execute a set of three tasks using each of the available interfaces (Direct Migration, Linear and Overview) in a context of task migration. Initially, the task description was presented and its execution process started on the TelEduc desktop interface to be further completed on the pocket PC simulator (the quantitative variables mentioned before were observed only during the pocket PC simulation). Tasks were chosen based on the subjects indication of the most frequently executed, in which 80% (12 subjects) said it was the checking of grades and 20% (three subjects) pointed the search for evaluations' details (*e.g.* date, criteria, *etc.*). Therefore, we implied that the checking of grades was the most relevant task, which led us to improve its efficiency in the Linear transformation approach (the grades matrix was replaced by a simpler table with only the user's grades, as shown in Fig. 10b). Table 3 compares actions' flow between treatments.

Steps	Treatments	
	Direct Migration/Overview	Linear
1	*c* **store perceive** <u>Evaluation</u>[name=*y*]	*aval* **store perceive** <u>Evaluation</u>[name=*y*]
2	*aval* **store perceive** <u>Evaluation</u>[code=*c*]	**return perceive** <u>Grade</u>(*aval*)
3	*stud* **store perceive** <u>Student</u>[name=*x*]	
4	**return perceive** <u>Grade</u>(*stud, aval*)	

Table 3. Task 1 actions' flow (*x*: student's name; *y*: evaluation's name).

We also wanted to investigate the implications of such improvement in related tasks, like the comparison of grades (two out of the 12 subjects explicitly mentioned it as the most frequently executed task). Thus, we created a second task in which subjects had to count the number of colleagues with a higher grade than his/her in a certain evaluation. While Task 1 should point the Linear interface as the most efficient due to its actions' flow simplification, Task 2 could help us investigate the implications of this consistency loss for a related task.

Finally, the third task was to go after the details of a certain evaluation (elected by three subjects as the most frequent task). As the second most executed task, a common scenario would be the user checking his/her grade in a certain evaluation and only then searching for details of the next evaluations to perform. In this sense, we implemented Task 3 with the same interface presented by the end of Task 1. This way, we provided both the adequate scenario according to the subjects' preferences (Task 3 stimulated by Task 1 or 2) and the means to investigate implications of a mental trace loss (incapacity to suppose the actions taken with device *x* to reach its current state due to a task migration started with device *y*).

Precautions with Tasks' Initial State

The following decisions were taken to make the task's initial state as real as possible and avoid particular cases that could benefit any of the evaluated treatments: (1) a standard user name was chosen to guarantee homogeneity for the subjects' search effort and also consider an average case for the user's name position inside the grades' matrix (Direct Migration and Overview transformation) and the combo box (Linear transformation); (2) evaluations and students' numbers, as well as the evaluations' names, were taken from a previous course; and (3) the students' grades in each evaluation were different to avoid users memorizing solutions with one treatment and reusing for the following.

5.5 Statistical Analysis

In order to adjust the residual effects in the task execution time continuous variable (due to the application of one treatment after the other), we opted for a parametric analysis of variance using latin square balanced for immediate residual effect (Cochran & Cox, 1992). The comparison between paired treatments was performed by the Tukey *post-hoc* test. As for the non-normal Likert scale discrete variables (easiness, efficiency and safety subjectively compared to the desktop TelEduc interface), the Friedman test was chosen according to its suitability for nonparametric analysis with three or more treatments and paired samples. Also, each pair of treatments was compared by the Wilcoxon signed rank test due to its adequacy on checking differences between medians of two groups with paired samples.

5.6 Results and Discussion

Task 1: User's Most Relevant Task

The checking of grades was considered the most executed task by 12 subjects out of 15 (not all the 18 subjects answered this question). Table 4 summarizes the observed data.

Observed Variables	Treatments		
	Direct Migration	Linear	Overview
Efficacy (task response accuracy)	18 (100%)	18 (100%)	18 (100%)
Efficiency (average execution time)*	50.93a	33.92b	45.32ab
Easiness compared to the TelEduc desktop*	3c	6a	4b
Efficiency compared to the TelEduc desktop*	2c	5a	4b
Safety compared to the TelEduc desktop**	3b	5a	4b
Best treatment's choice	0 (0%)	16 (89%)	2 (11%)
Worst treatment's choice	14 (78%)	1 (5%)	3 (17%)

* Treatments with different letters in the same line diverge significantly for $p < 0.05$
** Treatments with different letters in the same line diverge significantly for $p < 0.005$
Table 4. Task 1 results.

According to Table 4, all treatments had a perfect score for the task response accuracy, meaning that the consistency breaking, present in the Linear approach, didn't lead to errors. On the other side, although this approach has reduced considerably the required actions and their complex to perform the task, no significant difference was identified between its average execution time and the one obtained with the Overview treatment ($p < 0.05$). This result wasn't expected since the Linear transformation's major advantage is the efficiency gain by means of device oriented adaptations. Thus, *we conclude that executing a reduced number of simple non-expected actions can take as much time as a greater number of complex* *expected actions due to the user's mental model influence.* Although such conclusion is in agreement with the measured data, it can't be confirmed by the perceived data, which pointed the Linear approach as the most efficient one. This divergence can be explained by the fact that the efficiency subjective evaluation was realized after the task execution, when users eventually understood how to accomplish it. As the demanded cognitive adaptation wasn't relevant, users had the impression that an approach requiring a reduced number of simpler actions to perform the same task would take less time. We don't argue against this assumption (which is the reason why our proposal gives support to the third level of consistency on task personalization), but a context of task migration may prove the contrary, as verified by the measured efficiency.

Another important observation concerns the seven point Likert scale, in which the number four means no difference between handheld and desktop interfaces for the evaluated attribute (*i.e.* easiness, efficiency or safety). In this sense, the Overview transformation was the only approach able to maintain median four for every attribute, besides the significant differences to the Linear transformation's results ($p < 0.05$). Thus, *we conclude that the interface proposed using the first two levels of Consistency Priorities preserved the user's mental model by attending to his/her expectations.* We are confident that such goal has more important implications to multi-device design in order to smooth the transition between devices in contexts of interchange and task migration.

Finally, the user's choice for the best interface confirmed that this task's optimization in the Linear transformation was the key to get the users satisfaction.

Task 2: A Variation of the User's Most Relevant Task

The comparison of grades was explicitly mentioned by two subjects out of the 12 voters of Task 1 as the most relevant task. In this sense, Task 2 demanded the count of students with a higher grade than the user's in a certain evaluation. Table 5 summarizes the observed data.

Observed Variables	Treatments		
	Direct Migration	Linear	Overview
Efficacy (task accuracy)	17 (94%)	3 (33%)	17 (94%)
Efficiency (average execution time)*	69.48a	75.28a	30.15b
Easiness compared to the TelEduc desktop*	2b	2b	4.5a
Efficiency compared to the TelEduc desktop**	2b	2b	4a
Safety compared to the TelEduc desktop**	2b	3ab	4a
Best treatment's choice	2 (11%)	2 (11%)	14 (78%)
Worst treatment's choice	6 (33%)	12 (68%)	0 (0%)

* Treatments with different letters in the same line diverge significantly for $p < 0.05$

** Treatments with different letters in the same line diverge significantly for $p < 0.007$

Table 5. Task 2 results.

According to Table 5, while 94% of the subjects realized Task 2 correctly with both Direct Migration and Overview approaches, only 33% did it using the Linear interface. This result exemplifies how an interface adaptation privileging a certain task and breaking consistency can lead to bad effects on related tasks. Also, task efficiency[4] revealed the same aggravating effect, pointing the Overview transformation as the fastest interface to accomplish Task 2. Perceived data confirmed these results, indicating that the Overview approach was the easiest and most efficient interface. Additionally, the Overview transformation was able to preserve the user's mental model, given that its medians were closer to four than the other treatment's medians. At last, the subjects' preference for the Overview approach and aversion for the Linear confirmed the importance of the task perception consistency priority, which concerns not only with the Perceptual Constancy attributes and relative localization, but also with the design principles compromised by the devices' relevant restrictions. Once again, efficiency was indicated as the major reason for this choice, followed by safety as a confirmation of the best efficacy in task accuracy.

Task 3: User's Secondary Interest Task

The search for evaluations' details was considered of secondary interest to the users (three voters out of 15), which led us set the Task 1 last screen as its initial stage. This kind of task execution as a consequence of others is a common scenario and its effects for multi-device design have great importance, especially in a task migration context. Table 6 presents the observed data.

Observed Variables	Treatments		
	Direct Migration	Linear	Overview
Efficacy (task accuracy)	17 (94%)	15 (84%)	17 (94%)
Efficiency (average execution time)*	24.07a	25.58a	12.39b
Easiness compared to the TelEduc desktop**	4ab	3b	4a
Efficiency compared to the TelEduc desktop*	3b	3ab	4a
Safety compared to the TelEduc desktop***	4a	3b	4a
Best treatment's choice	4 (22%)	4 (22%)	10 (56%)
Worst treatment's choice	8 (44%)	9 (50%)	1 (6%)

* Treatments with different letters in the same line diverge significantly for $p < 0.05$

** Treatments with different letters in the same line diverge for $p < 0.132$

*** Treatments with different letters in the same line diverge for $p < 0.158$

Table 6. Task 3 results.

According to Table 6, once again the Overview approach was able to overmatch the Linear transformation for every observed data. Both its measured values for efficacy and efficiency

[4] The experiment isolated task perception preventing users from activating any control mechanism inside any interface for all tasks (besides scroll bars). This procedure was crucial to identify that the Linear transformation's lower efficacy in Task 2 (33%) was due to a problem in the first stage of the user's mental model update cycle: hardness in identifying the need to switch students' names inside the combo box (see Fig. 10b). In order to guarantee a fair comparison between treatments for Task 2, we computed the following measures *for the Linear approach:* (1) the time taken by each subject to indicate the combo box activation as the first step to complete the task and (2) the smallest time to finish remaining steps (i.e. switch names inside the combo box, find and compare each grade with the user's grade, and count the total of greater grades). Thus, each subject's task execution time was a combination of both measures. Even benefitting the Linear approach with the increase of the smallest remaining time to each subject's partial time, Overview still proved to be more efficient.

were higher, indicating that users could perform the task faster and with better safety using an interface consistent to their previous experience. The Overview approach also overcame the efficiency obtained with the Direct Migration, indicating that the consistency on task perception shall be discussed together with the design principles compromised by the devices' relevant restrictions. Also, the subjective evaluation pointed the Overview as the best treatment to preserve the user's mental model by keeping a median four on every evaluated attribute under the seven point Likert scale. This might be the most important result for task migration contexts, in which perceptual changes could reduce devices' utility. Finally, the Overview interface was considered the best for Task 3 because of its efficiency. On the other side, the Linear transformation was the worst due to its layout differences compared to the TelEduc desktop interface, which confirms results from Task 2. These evaluations reveal the importance of consistency with the user's previous experience in order to address efficiency for multi-device applications.

User's Satisfaction

The last subjective evaluation aimed to compare interfaces by asking users to choose the one they liked most *despite the executed tasks* and to mention their reasons for that. Observed results highlight the importance of personalization: 12 subjects opted for the Linear approach; four chose Overview; and the remaining two decided for the Direct Migration.

As it can be noticed, although only the user's most relevant task had been improved with the Linear transformation (even without a significant difference to the Overview's average execution time), this approach was still considered the most attractive, contradicting some previous findings (MacKay et al., 2004; Lam & Baudisch, 2005; Roto et al., 2006). The main reason for such divergence is that our experiment didn't support comparisons to the Single Column automatic Linear transformation approach as on these authors' user studies. On the contrary, we decided to redesign TelEduc for the pocket PC to take the best from the device and also optimize the user's most relevant task. This way, we ended up with a more adapted and usable interface than the Single Column, in which no user-centered design decisions are taken. Thus, we argue that this experiment's design was able to make more fair comparisons because it took the best of each evaluated approach.

After carefully analyzing the questionnaire's answers, we perceived that the better usability present in the users' most relevant task with the Linear transformation was the major factor for its subjects' preference, as indicated by the following comments:

> *"…I consider this linear interface more functional than the actual TelEduc."* (Subject 1)
>
> *"...the linear approach makes activities easier than the TelEduc."* (Subject 15)
>
> *"The linear interface would be good even for the TelEduc desktop!"* (Subject 15 about Task 1)

Considerations like these raise questions about the TelEduc's usability as if the decisions taken for the pocket PC should also have been taken for the desktop, confirming once again the importance of consistency between both interfaces. Still, we had to explain the reason why subjects were more satisfied with an approach that wasn't able to reveal advantages in practice for the executed tasks. In this sense, the following factors may have contributed:

- *Low-risk decisions:* Eight out of the 12 Linear approach electors weren't able to execute one or more tasks with accuracy using this approach. However, every Overview and Direct Migration voter executed the three tasks correctly. We believe that, if the application domain had involved high risk decisions, no error would have been tolerated (*e.g.* money transferring, management of chemical residues, operation of high-cost machines, *etc.*);

- *Indifference to multi-device access:* Although the demographic questionnaire had revealed that 12 out of 18 subjects were interested in using mobile interfaces for the TelEduc, only five also wanted to access the system via desktop, which characterizes indifference over the multi-device access and inadequacy of task migration contexts for the considered domain and/or sample. Thus, the third level of Consistency Priorities was more appropriated to guarantee personalization of relevant tasks. Some comments in favor of the Linear approach confirm this assumption:

  "Because of fitting more information […] that I consider of my interest." (Subject 4)
  "...presents information in a more objective and intelligible way" (Subject 6)
  "Structure directed to the student individually." (Subject 12)
  "Because it shows individual information, less error inclined." (Subject 13)

As it can be noticed, although the Overview approach had been more adequate to execute tasks in general, subjects revealed a better satisfaction with the Linear transformation because of its task personalization. This observation corroborates that *the efficacy generated by the first two levels of the Consistency Priorities hierarchy concerning task perception and execution must be combined with the third level of personalization aiming better satisfaction and efficiency.* This combination may consider multiple use contexts by creating layout patterns to be chosen by the end user. This procedure can support both the sole and multi-device access in contexts of interchange and task migration.

6. Conclusion

The multi-device design methodology proposed in this chapter was based and supported by well established concepts from Philosophy and Psychology (definitions about logic and inductive reasoning), Connectionism laws (Thorndike, 1898), Cognitive learning theories (Hartley, 1998) and mental models (Young, 1983), as well as by recent findings from Neuroscience (Sohlberg & Mateer, 1989) and Human-Computer Interaction (MacKay et al., 2004; Lam & Baudisch, 2005; Roto et al., 2006; Hajdukiewicz, 2006). These theoretical foundations reinforce the hypothesis that interfaces of the same application must preserve perceptual characteristics and adopt a consistent behavior to execute tasks.

The experiment conducted on the Distance Learning domain also contributed with the following conclusions for multi-device design in contexts of task migration:

- *The Consistency Priorities (first two levels) preserve the user's mental model better than approaches maintaining full layout consistency (Direct Migration) or with a more dedicated design focus to the devices' characteristics (Linear):* This was verified via a subjective evaluation of the handheld interface built with our methodology, which revealed a significant similarity to the desktop version for easiness, efficiency and safety on tasks relevant to the users. This result was also significantly different to those obtained with the Direct Migration and Linear approaches, confirming their inability to attend the users' expectations;
- *The Consistency Priorities (first two levels) achieve similar efficacy and efficiency as the Linear's for tasks optimized in the latter:* Although the Linear interface was optimized for better efficiency with the user's most relevant task, our approach maintained similar efficacy and efficiency despite requiring more steps to execute. This fact reveals the

importance of consistency with the user's previous experience in contexts of task migration;

- *The Consistency Priorities (first two levels) enhance efficacy and efficiency compared to the Linear's for tasks not optimized in the latter:* Three times more subjects solved general tasks correctly using our approach contrasted to the Linear's and they took less than half of the Linear's time.

Although these results point the Consistency Priorities as a more adequate multi-device design approach for task migration, the Linear interface had higher preference by the subjects of the experiment due to its task personalization. This apparent contradiction can be explained by the fact that the experiment was realized in a context of task migration, but both the demographic questionnaire and the users' satisfaction evaluation made it clear that the sample wasn't interested in such context. Thus, while the first two levels of Consistency Priorities guaranteed better usability and preservation of the user's mental model, the personalization in the Linear approach had a great acceptance because the majority of the subjects preferred to access the application using only one device. This is in accordance with our initial claim that there is no multi-device approach capable to provide full usability in every context because the user may choose only one interface to access the application or interchange its use via many devices. Therefore, *it is necessary to combine approaches with different goals and suit the user according to the appropriate context.* In other words, *the efficacy generated by the first two levels of the Consistency Priorities hierarchy concerning task perception and execution must be combined with the third level of personalization aiming better satisfaction and efficiency.* This combination can be addressed with an active role for the user who shall specify the context of use in order to interact with the adequate interface pattern.

Results and implications obtained so far still leave open questions and draw lines of future research that might be pursued in follow-on work. Some of these questions are listed below:

- *Could the experiment results be extended to other domain applications besides e-learning?* We expect high-risk applications to reinforce our proposal of applying the first two levels of Consistency Priorities due to its better efficacy on task execution. Yet, applications with a restricted set of tasks and a clear demand for efficiency instead of accuracy may highlight the importance of personalization applied in the Linear approach. In both cases, combining approaches with different goals and suiting the user according to the appropriate context shall be perceived as a relevant design proposal;
- *Could the experiment results be extended to other samples?* The experiment's sample included only computer experts and even though the consistency in task perception and execution presented better results than the Linear approach (*e.g.* in Task 2, subjects didn't identify with good efficacy the need to switch students' names in the combo box of the Linear interface because this procedure wasn't in accordance with their previous experience). Thus, we expect that samples including computer novice users will highlight even more the importance of the first two levels of Consistency Priorities besides reducing the interest for task personalization (third level);
- *Could the experiment results be extended to contexts of sole device access?* If users first learn how to interact with a certain application using an interface *x* and only then opt for an interface *y*, we expect the transition between them to reveal similar results as those observed in our experiment. However, if users never need to accomplish any task with any of the application's interfaces besides with the only one they know, we expect

better results with the Linear transformation. Thus, we need to know how likely it is for users not to access an application using more than one available interface;

- *Once interfaces for task migration and sole device access were proposed, how could they be implemented by an automatic transformation approach in order to ease software maintenance?* We developed a prototype for contexts of interchange and task migration that automatically adapts the TelEduc desktop interface applying the decisions taken on the first two levels of the Consistency Priorities hierarchy. The adaptation process was similar to that described by Lam & Baudisch (2005), but including our proposed changes for text summarization, detailed view and also the ability to run the system with a pocket PC/smartphone web browser [5]. As for the sole device interface (applying also the third consistency priority), we didn't implement it for automatic adaptation because different types of personalization could make the adaptation very specific and vulnerable to small changes on the desktop interface;
- *How could the Consistency Priorities design process be automated?* The development of tools for task and actions' flow modeling integrated to the restriction analysis of target devices will be of great interest to both designer and developer. Most of all, the automatic identification of inconsistencies based on heuristics of interface analysis, and the solutions proposal based on the compromised design principles could dictate a new trend for the next generation of multi-device development environments.

We expect the arguments and conclusions presented herein to be useful as a support for user centered multi-device design. Thus, not only contexts of interchange and task migration shall be approached in a more adequate way, but also sole device access, in which users have an active role of personalization while choosing and/or customizing the interface.

7. References

Banavar, G.; Beck, J.; Gluzberg, E.; Munson, J.; Sussman, J. & Zukowski, D. (2000). Challenges: an application model for pervasive computing, *Proceedings of the 6th international conference on mobile computing and networking (MobiCom)*, pp. 266-274, ISBN 1-58113-197-6, Boston, MA, USA, August 2000, ACM Press.

Banerjee, S.; Gupta, A. & Basu, A. (2003). Online transcoding of web pages for mobile devices. *Proceedings of the 5th international conference on human-computer interaction with mobile devices and services (MobileHCI),* pp. 271-285, Udine, Italy, September 2003.

Baudisch, P.; Lee, B.; Hanna, L. (2004). Fishnet, a fisheye web browser with search term popouts: a comparative evaluation with overview and linear view. *Proceedings of the working conference on advanced visual interfaces (AVI),* pp. 133-140, ISBN 1-58113-867-9, Gallipoli, Italy, 2004. ACM Press.

Bergman, L. D.; Banavar, G.; Soroker, D. & Sussman, J. B. (2002). Combining handcrafting and automatic generation of user-interfaces for pervasive devices. *Proceedings of the 4th international conference on computer-aided design of user interfaces (CADUI),* pp. 155-166, Valenciennes, France, May 2002.

[5] Compatible with CSS, DHTML and Javascript (*e.g.* Opera Mobile: www.opera.com/products/mobile).

Buyukkokten, O.; Garcia-Molina, H.; Paepcke, A. & Winograd, T. (2000) Power browser: efficient web browsing for pdas. *Proceedings of the SIGCHI conference on human factors in computing systems (CHI)*, pp. 430-437, ISBN 1-58113-216-6, The Hague, The Netherlands, April 2000. ACM Press.

Cochran, W. G. & Cox, G. M. (1992). *Experimental designs.* Wiley & Sons, ISBN 978-0-471-54567-5, Canada, 2nd edition, 1992.

Denis, C. & Karsenty, L. (2004). Inter-usability of multi-device systems - A conceptual framework, In : *Multiple user interfaces: Cross-platform applications and context-aware interfaces*, pp. 374-385, Wiley & Sons, West Sussex, UK, 2004.

Dömel, P. (1995). Webmap: A graphical hypertext navigation tool. *Computer Networks and ISDN Systems*, v. 28, pp. 85-97, ISSN 0169-7552, 1995.

Coninx, K.; Luyten, K.; Vandervelpen, C.; den Bergh, J. V. & Creemers, B. (2003). Dygimes: dynamically generating interfaces for mobile computing devices and embedded systems. *Proceedings of the 5th international conference on human-computer interaction with mobile devices and services (MobileHCI),* pp. 256-270, Udine, Italy, September 2003.

Eisenstein, J.; Vanderdonckt, J. & Puerta, A. R. (2000). Adapting to mobile contexts with user interface modeling. *Proceedings of the 3rd IEEE workshop on mobile computing systems and applications (WMCSA)*, pp. 83-92, ISBN 0-76950-816-2, Monterey, CA, USA, December 2000. IEEE Computer Society.

Florins, M.; Trevisan, D. G. & Vanderdonckt, J. (2004). The continuity property in mixed reality and multiplatform systems: a comparative study. *Proceedings of the 6th international conference on computer-aided design of user interfaces (CADUI),* pp. 323-334, Funchal, Isle of Madeira, Portugal, January 2004.

Gaeremynck, Y.; Bergman, L. D. & Lau, T. (2003). More for less: model recovery from visual interfaces for multi-device application design. *Proceedings of the 8th international conference on intelligent user interfaces (IUI)*, pp. 69-76, ISBN 1-58113-586-6, Miami, FL, USA, January 2003. ACM Press.

Gajos, K. & Weld, D. S. (2004). Supple: automatically generating user interfaces. *Proceedings of the 9th international conference on intelligent user interfaces (IUI)*, pp. 93-100, ISBN 1-58113-815-6, Funchal, Portugal, January 2004. ACM Press.

Grudin, J. (1989). The case against user interface consistency. *Commun. ACM*, v. 32, n. 10, pp. 1164-1173, ISSN 0001-0782, 1989.

Hajdukiewicz, J. (2006). Interaction momentum - industrial application design and consistency across platforms. *Proceedings of the CHI 2006 workshop on the many faces of consistency in cross-platform design*, 2006.

Hartley, J. (1998). *Learning and Studying. A research perspective,* Routledge, ISBN 0-41516-851-1, London, UK, 1998.

Hodes, T. D.; Katz, R. H.; Servan-Schreiber, E. & Rowe, L. (1997). Composable ad-hoc mobile services for universal interaction. *Proceedings of the 3rd annual ACM/IEEE international conference on mobile computing and networking (MobiCom),* pp. 1-12, ISBN 0-89791-988-2, Budapest, Hungary, 1997. ACM Press.

Johnson-Laird, P. N. (1983). *Mental Models,* Cambridge University Press, ISBN 0-67456-882-6, Cambridge, 1983.

Jones, M.; Marsden, G.; Mohd-Nasir, N. & Buchanan, G. (1999). A site-based outliner for small screen web access. *Proceedings of W8 conference,* Toronto, USA, May 1999.

Lam, H. & Baudisch, P. (2005). Summary thumbnails: readable overviews for small screen web browsers, *Proceedings of the SIGCHI conference on human factors in computing systems (CHI)*, pp. 681-690, ISBN 1-58113-998-5, Portland, OR, USA, April 2005, ACM Press.

Lin, J. (2005). *Using design patterns and layers to support the early-stage design and prototyping of cross-device user interfaces.* PhD thesis, University of California, Berkeley, 2005.

MacKay, B.; Watters, C. R. & Duffy J. (2004). Web page transformation when switching devices, *Proceedings of the 6th international conference on human-computer interaction with mobile devices and services (MobileHCI)*, pp. 228-239, ISBN 3-540-23086-6, Glasgow, Scotland, September 2004, Springer.

Manktelow, K. & Jones J. (1987). Principles from the Psychology of Thinking and Mental Models, In : *Applying cognitive psychology to user-interface design,* pp. 83-117, John Wiley & Sons, ISBN 0-471-91184-4, New York, NY, USA, 1987.

Milic-Frayling, N. & Sommerer, R. (2002). Smartview: Enhanced document viewer for mobile devices. Technical Report MSR-TR-2002-114, Microsoft Research, November 2002.

Mori, G.; Patèrno, F. & Santoro, C. (2003). Tool support for designing nomadic applications. *Proceedings of the the 8th international conference on intelligent user interfaces (IUI),* pp. 141-148, ISBN 1-58113-586-6, Miami, FL, USA, January 2003. ACM Press.

Nation, D. A.; Plaisant, C.; Marchionini, G. & Komlodi, A. (1997). Visualizing web sites using a hierarchical table of contents browser: Webtoc. *Proceedings of the 3rd conference on human factors and the web*, Denver, Colorado, USA, June 1997.

Nichols, J.; Myers, B. A.; Higgins, M.; Hughes, J.; Harris, T. K.; Rosenfeld, R. & Pignol, M. (2002). Generating remote control interfaces for complex appliances. *Proceedings of the 15th annual ACM symposium on user interface software and technology (UIST)*, pp. 161-170, ISBN 1-58113-488-6, Paris, France, October 2002. ACM Press.

Nichols, J. (2006). *Automatically generating high-quality user interfaces for appliances,* PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania, December 2006.

Nichols, J.; Chau, D. H. & Myers, B. A. (2007). Demonstrating the viability of automatically generated user interfaces. *Proceedings of the SIGCHI conference on human factors in computing systems (CHI),* pp. 1283-1292, ISBN 978-1-59593-593-9, San Jose, CA, USA, April 2007. ACM Press.

Norman, D. A. (1988). *The psychology of everyday things.* Basic Books, ISBN 0-46506-709-3, New York, NY, USA, 1988.

Norman, D. A. (1993). *Things that make us smart: defending human attributes in the age of the machine.* Addison Wesley Publishing Company, ISBN 0201626950, 1993.

Paternò, F.; Mancini, C.; Meniconi, S. (1997). ConcurTaskTrees: a diagrammatic notation for specifying task models. *Proceedings of the IFIP TC13 international conference on human-computer interaction (INTERACT'97)*, pp. 362-369, ISBN 0-412-80950-8, Sydney, Australia, July 2007. Chapman & Hall.

Payne, S. J.; Squibb, H. R. & Howes, A. (1990). The nature of device models: the yoked state space hypothesis and some experiments with text editors. *Human-Computer Interaction*, v. 5, n. 4, pp. 415-444, 1990.

Ponnekanti, S.; Lee, B.; Fox, A.; Hanrahan, P. & Winograd, T. (2001). ICrafter: A service framework for ubiquitous computing environments. *Proceedings of the 3rd* *international conference on ubiquitous computing (Ubicomp)*, pp. 56-75, Atlanta, Georgia, USA, September 2001. Springer.

Pyla, P. S.; Tungare, M. & Pérez-Quiñones, M. (2006). Multiple user interfaces: Why consistency is not everything, and seamless task migration is key. *Proceedings of the CHI 2006 workshop on the many faces of consistency in cross-platform design.*

Rasmussen, J. (1987). Mental models and the control of action in complex environments. *Proceedings of the informatics and psychology workshop,* pp. 41-69, 1987.

Roto, V.; Popescu, A.; Koivisto, A. & Vartiainen, E. (2006). Minimap: a web page visualization method for mobile phones. *Proceedings of the SIGCHI conference on human factors in computing systems (CHI),* pp. 35-44, ISBN 1-59593-372-7, Montreal, Canada, April 2006, ACM Press.

Santo, G. D. & Zimeo, E. (2007). Reversing guis to ximl descriptions for the adaptation to heterogeneous devices. *Proceedings of the 22nd ACM symposium on Applied computing (SAC),* pp. 1456-1460, ISBN 1-59593-480-4, Seoul, Korea, March 2007. ACM Press.

Sharp, H.; Rogers, Y.; Preece, J. (2007). *Interaction Design: Beyond Human Computer Interaction.* 2nd. ed. Wiley & Sons, ISBN 0-470-01866-6, New York, USA, 2007.

Sohlberg, M. M. & Mateer, C. A. (1989). *Introduction to cognitive rehabilitation: theory and practice.* Guilford Press, ISBN 0-89862-738-9, New York, NY, USA, 1989.

Thorndike, E. L. (1898). *Animal intelligence: an experimental study of the associative processes in animals.* Psychological Review Monograph Supplement, v. 2, n. 4, pp. 1-109, 1898.

Wozny, L. A. (1989). The application of metaphor, analogy and conceptual models in computer systems. *Interacting with Computers,* v. 1, pp. 273-283, 1989.

Young, R. M. (1983). Surrogates and mappings: two kinds of conceptual models for interactive devices, In: *Mental models,* Gentner D. & Stevens A. (Ed.), pp. 35-42, Lawrence Erlbaum Associates, ISBN 0-89859-242-9, Hillsdale, NJ, USA, 1983.

7

Multifinger Haptic Interfaces for Collaborative Enviroments

Manuel Ferre, María Oyarzábal, Alexandre Campos and Mary Monroy
Universidad Politécnica de Madrid
Spain

1. Introduction

Haptic interfaces provide users with force information while they are interacting with virtual objects, allowing them to perform manipulation tasks and cooperate. Multi-finger haptic interfaces benefit from the use of several fingers, thereby a large number of degrees of freedom are processed, to improve interaction with virtual environments and increase the sense of immersion.

This chapter introduces a new two-finger haptic interface, known as MasterFinger-2. It improves haptic interaction; grasping objects can be easily reproduced by using this device. This interface is based on an open architecture which allows the control of each finger independently via Ethernet. In this sense, it also permits an easy development of cooperative tasks, where users interact directly with their fingers instead of using a tool.

MasterFinger-2 is based on a modular design in which each finger has its own mechanical structure and electronic controller. A finger is inserted into a thimble with 6 degrees of freedom, any position and orientation can be consequently achieved by each finger. Forces are reflected in any direction since there are three actuators per finger.

2. Overview of Haptic Devices

Haptic interfaces are devices which show tactile and force information to a user interacting with a real or virtual object (Tan, 1994). They allow the user to touch the objects and feel their mechanical properties, *e.g.* texture, hardness, shape etc. They can be also used to remotely manipulate objects, *i.e.* teleoperation.

The term "haptic interface" is frequently used to describe two types of interfaces, tactile devices and kinesthetic devices (Sciavicco & Siciliano, 2000), they differ in the kind of information exchanged with the user and the hardware used to build them. Tactile devices provide only tactile information to users, and none regarding kinesthetics. Kinesthetic devices usually provide information of reflected forces, although they can also provide tactile information.

Tactile information comes from the contact between the user finger and the device. Forces are produced by the device and are high enough to resist user's hand and arm movements. Kinesthetic devices are usually based on electric DC motors or other actuators exerting

forces on users. On the other hand, tactile devices are based on pins, vibratory elements or air injected to stimulate skin.
The contact interface between user and device in kinesthetic devices can adopt multiple configurations. It can be a joystick grasped by two fingers or it can provide more precise information by stimulating several fingers, *e.g.* exosqueletons. Tactile devices apply generally stimuli to the fingertips through matrices of pins, executed by piezoelectric crystals, servomotors, solenoids or pneumatic systems.

2.1. Tactile Devices

Nowadays, there are not many multifinger tactile interfaces available. EXOS Inc. commercializes Touchmaster; it can be used independently or with the Dexterous HandMaster. Xtensory Inc. commercializes the Tactool system. A completely different system is the Displace Temperature Sensing System (DTSS), commercialized by CM Research, which provides temperature information. Next table shows a summary of these interfaces.

Device	Company	Interface	Actuator	Stimuli	Tactile Sensation
CyberTouch (Immersion)	Immersion	5 fingers and the palm (6 vibrotactile stimulators)	Vibrotactile	Vibration 0-125Hz 1.2N peak-to-peak@ 125Hz	Contact with objects
TouchMaster (Exos, 1993)	EXOS	5 fingers and the palm	Magnetic	Vibration (0-200Hz)	Contact with objects
Tactool System	Xtensory	2 fingers	Pins	Impulse (30g) Vibration (20Hz)	Contact with objects
Displaced temperature Sensing System	CM Research	Through a thimble	Thermoelectric heat pump	Temperature change	Heating / Cooling

Table 1. Tactile Devices

2.2 Kinesthetic Interfaces

Compared to tactile devices, kinesthetic interfaces are generally bigger and heavier due to actuators' force requirements. These devices can couple to the hand by means of an exosqueleton, a glove, a thimble, a joystick, etc. In the following table we give a summary of some kinesthetic interfaces in which we can observe different ways of coupling the interface to the hand or fingers.

Device	Company	Degrees of Freedom	Main Features
PHANTOM (Massie & Salisbury,1994).	Sensable	6	Serial morphology First three DOF active and last three passive
SPIDAR-G (Kim et al., 2000)	Tokyo Institute of Technology	7	Based on thin steel cables that reflect forces to the end effectors. 3 DOF for translation, 3 DOF for rotation and 1 DOF for grasping
SARCOS (Sarcos)	Sarcos	7	Arm kinematics similar to human arm kinematics
VISHARD 10 (Ueberle et al., 2004)	Technical University of Munich	10	Hyperedundant system. Wide workspace
EXOS FORCE ARMMASTER (Exos, 1993)	Exos	5	Five DOF in the upper part of the arm. Two DOF in the lower part of the arm
CYBERGRASP (Immersion)	Immersion	5 DOF for force feedback (1 for each finger)	18 or 22 force sensors. Sensors to measure flexion and abduction
HIRO-II (Kawasaki et al., 2005).	Gifu University, Japan	6 in the arm 15 in the hand	Force and tactile sensation in all fingertips
MAGISTER-P (Sabater et al, 2007)	Miguel Hernández University, Spain	6	Parallel structure

Table 2. Kinesthetic Interfaces

3. MasterFinger Design

MasterFinger is a modular haptic interface where each finger is independently managed. All modules share mechanical structure and controller. Therefore, it is easy to scale the system from one to three fingers, or more. Next section shows the mechanical design of MasterFinger modules and describes the versions for two and three fingers.

3.1 Design of Masterfinger Module

Each finger is considered as an independent module with its own mechanical structure, controller and communications. Mechanical module design is based on a serial-parallel structure (Tsai, 1999) which confers it a wide workspace with a very low inertia.

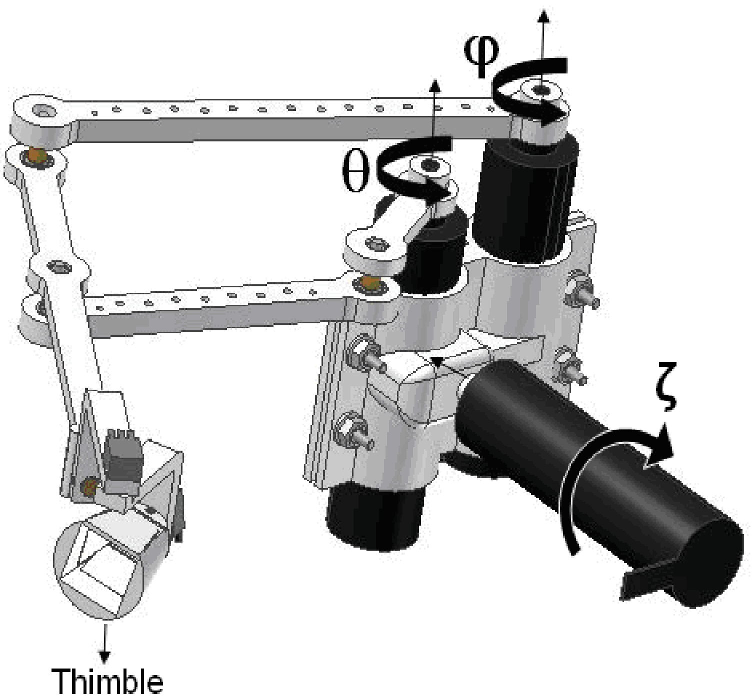


Fig. 1. Design of MasterFinger Modules

This configuration allows a comfortable manipulation since actuator inertia is mainly supported by the base. A module is made up of a six-degree of freedom mechanism and 3 actuators, as shown in Fig. 1. The second and third actuators are linked to a five-bar-structure (Tsai, 1999) providing a wide workspace area. This structure is linked to a thimble by a gimble with three-rotational degrees of freedom. The first degree of freedom allows vertical hand movements − approximately corresponding to the deviation movement ulna-radius in the wrist − while the second and third degrees of freedom are mainly related to finger movements. Figure 2 shows the five-bar-mechanism based on a parallel structure, *i.e.* second and third degrees of freedom.

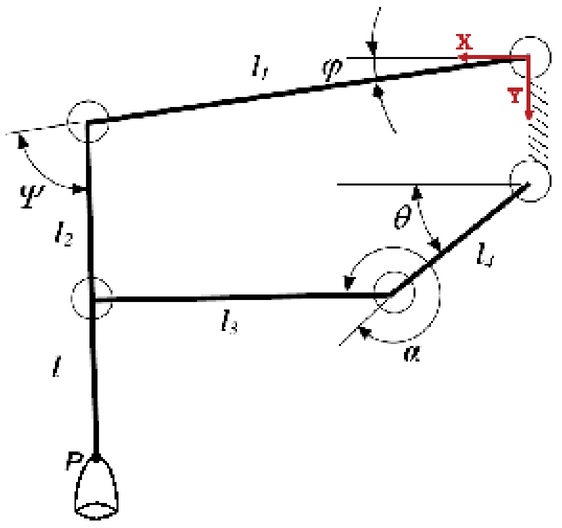


Fig. 2. Schematic view of the two first degrees of freedom.

Equations describing the five-bar-mechanism are the following:

$$x_P = l_1\cos(\varphi) + l\cos(\varphi + \psi) \tag{1}$$

$$y_P = l_1 sen(\varphi) + l sen(\varphi + \psi) \tag{2}$$

φ and θ angles are provided by the encoders, and *a* and ψ can be calculated as:

$$\psi = \arctan\left(\frac{B}{A}\right) - \arcsin\frac{b}{2l_2\sqrt{b - l_2^2 + l_3^2}} \tag{3}$$

where

$$B = -2l_1l_2 + 2l_2l_4\cos(\varphi - \theta) + 2dl_2\cos(\varphi) \tag{4}$$

$$A = 2l_2l_4\sin(\varphi - \theta) + 2dl_2\sin(\varphi) \tag{5}$$

$$a = l_1^2 + l_2^2 + d^2 - l_3^2 + l_4^2 \tag{6}$$

$$b = a - 2l_1l_4\cos(\varphi - \theta) - 2dl_1\cos(\varphi) + 2dl_4\cos(\theta) \tag{7}$$

and

$$\theta = -\arctan\left(\frac{B}{A}\right) - \arcsin\frac{b}{2l_2\sqrt{b - l_2^2 + l_3^2}} \tag{8}$$

therefore,

$$B = 2l_4 d - 2l_1 l_4 \cos(\varphi) - 2l_2 l_4 \cos(\varphi + \psi) \tag{9}$$

$$A = -2l_2 l_4 \sin(\varphi + \psi) - 2l_1 l_4 \sin(\varphi) \tag{10}$$

$$a = l_1^2 + l_2^2 + d^2 - l_3^2 + l_4^2 \tag{11}$$

$$b = a + 2l_1 l_2 \cos(\psi) - 2l_1 d \cos(\varphi) - 2l_2 d \cos(\varphi + \psi) \tag{12}$$

The Jacobian matrix allows formulating the differential model of joint velocities related to the end effector velocity, in Cartesian coordinates, and joint torques related to forces exerted at the end effector (Mark, 2006). Jacobian matrix is obtained from the following expression:

$$J = J_0 J_d \tag{13}$$

where

$$J_o = \begin{pmatrix} -l_1 \sin(\varphi) - l \sin(\varphi + \psi) & -l \sin(\varphi + \psi) \\ l_1 \cos(\varphi) + l \cos(\varphi + \psi) & l \cos(\varphi + \psi) \end{pmatrix} \tag{14}$$

$$J_d = \begin{pmatrix} 1 & 0 \\ \frac{\partial \psi}{\partial \varphi} & \frac{\partial \psi}{\partial \theta} \end{pmatrix} \tag{15}$$

The thimble orientation is measured by three encoders placed in the corresponding gimble joints. Fig 3 shows further thimble and gimble details. The thimble can be oriented in any direction in order to guarantee free movements of the finger. The three rotational axis of the gimble intersect on the user's finger tip. This geometrical configuration avoids torque reflection, *i.e.* only forces are reflected to the user's finger. The thimble has been developed to completely enclose the operator finger. The thimble includes four Flexiforce sensors by Tekscan Inc. These sensors are used to estimate normal and tangential forces exerted by the user. Normal forces are obtained from the sensor placed at the thimble bottom in contact with the finger tip. Tangential forces (Burdea 1996) are estimated from three sensors placed on the thimble inferior and lateral faces, respectively. Figure 3 gives two views of the thimble with these sensors.

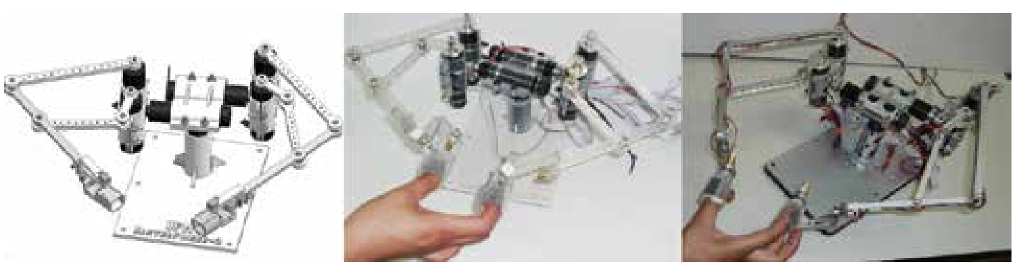
(a) (b) (c)

Fig. 3. Lateral and frontal view of thimble and sensors: a). CAD. (b). Resine prototype. (c)Aluminium prototype

All MasterFinger v1.0 components were initially built through a technique of rapid prototyping, stereo-lithography, using epoxy resin. The resin low weight allows an easy manipulation of the entire interface. However, the clearance from the material provokes problems regarding high precision. A second prototype has been built in aluminium aiming to obtain a better precision keeping low weight and inertia effects, which improves the user maniobrability. In order to obtain the reflected forces, three DC motors (Maxon RE 25, 10W) with a 225/16 reduction-planetary gear unit GP26 are used. These motors include also a 1000-pulse-per-revolution encoder providing motor orientation.

3.2 MasterFinger Architecture

MasterFinger-2 is made up of two modules, placed in such a way that the index and thumb fingers can handle it. It allows the user to interact with virtual environments in an easy and comfortable way for grasping tasks. Both modules are connected to the interface base with an additional joint to increase the workspace of this haptic interface. The first motor of both modules is on a horizontal plane; therefore, device inertia is significantly reduced. Figure 4 shows a general view of MasterFinger-2.



(a) (b) (c)

Fig. 4. Masterfinger-2: six degrees of freedom for each finger: (a). CAD model. (b). Resine prototype. (c).Aluminium prototype

It is advisable to notice that the module allows two different configurations, *i.e.* up-elbow and down-elbow, as shown in figure 4 (a). A workspace analysis was made aiming to compare these two options. According to figure 5, the up-elbow configuration has a bigger workspace than the down-elbow configuration. For this reason we block a joint in the five-bar mechanism in order to avoid down–elbow configurations. This workspace represents the volume where finger tips can be located, close to a 300 mm diameter sphere, hand movements correspond therefore to a wider space.

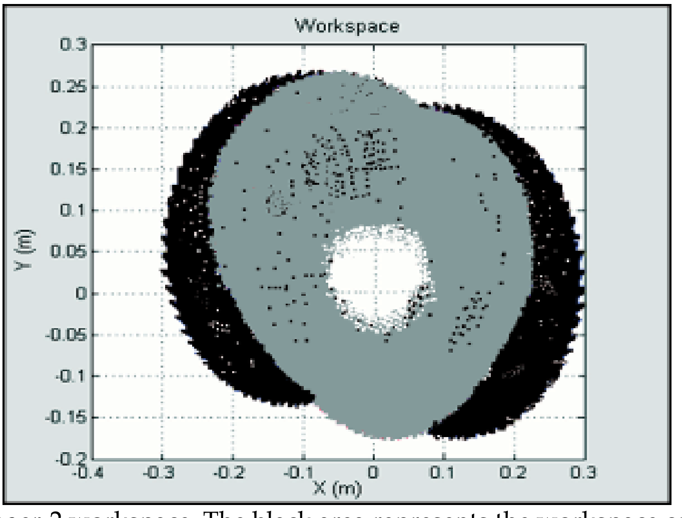


Fig. 5. MasterFinger-2 workspace. The black area represents the workspace covered by the up-elbow configuration and the grey area the one by the down-elbow configuration.

Additionally, technical features of MasterFinger-2 are described as follows. MF-2 weights approximately 2400 gr, so it can be easily transported to different locations. Each finger controller is provided with Ethernet access and each one uses a switch which works at 100 Mbits per second. UDP acts as the communication protocol; packets are transmitted at 200Hz between MasterFinger-2 and a computer that manages the environment simulation.

3.3 Prototype for a three-finger haptic interface

Some preliminary MasterFinger-3 designs have already been developed. They are currently under evaluation. MasterFinger-3 is a haptic interface for three fingers; thumb, index and either the middle or ring finger. This device will be made up of three modules which will be independently controlled. Figure 6 shows some designs developed so far around the MasterFinger-3 mechanical structure.

The design shown in figure 6a represents a MasterFinger-2 extension where the third module is attached to the common base of the haptic interface. The main advantage of this design is given by its reduced weight. Figure 6b shows the second design. This mechanism has a wide workspace, as its first degree of freedom is provided by a pulley system moving the device base. It also has an additional degree of freedom between the index and middle finger, known as "abduction movement".

Figure 6c shows the third design with the abduction movement between index and middle fingers too. It has a small wheel in the inferior part of the third module to better support the motor weight.

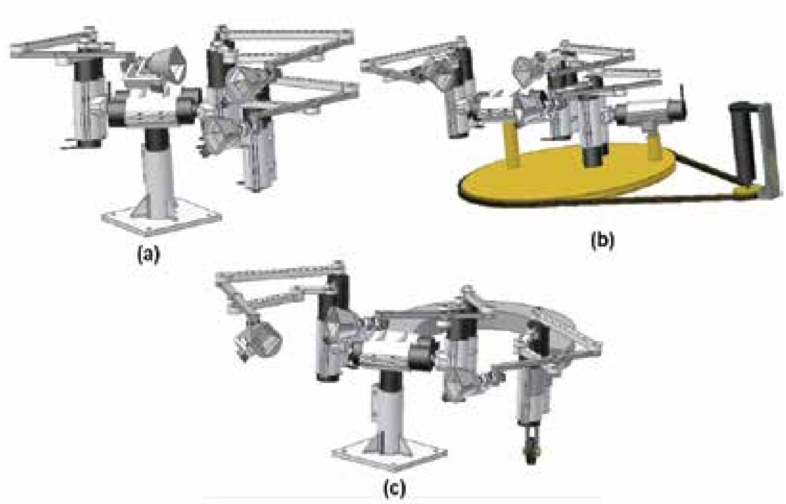


Fig. 6. Three different Masterfinger-3 designs

4. MasterFinger-2 Applications

The Masterfinger-2 has been designed to provide precise grasping using the thumb and index fingers. MasterFinger-2 is very suitable for cooperative tasks where two or more users are manipulating a virtual object. A networked architecture has been developed for this kind of application.

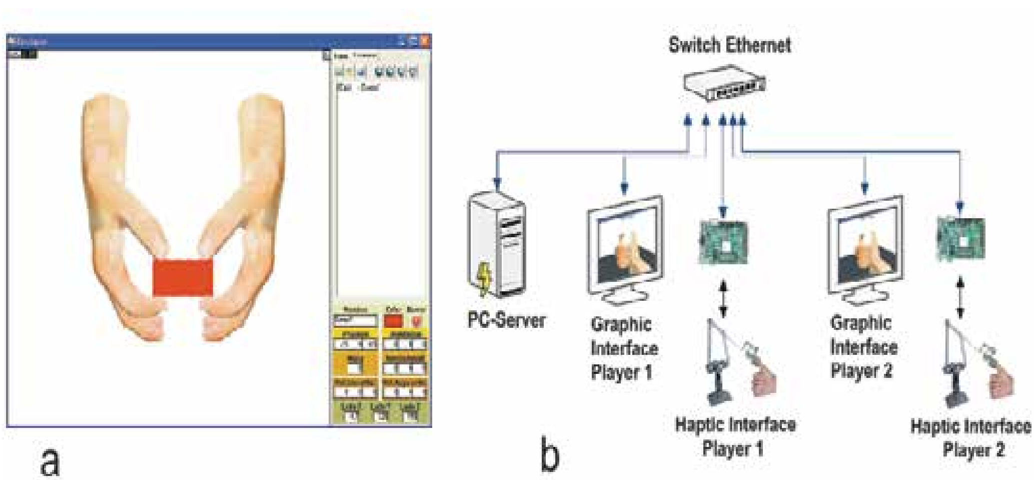


Fig. 7. a) Graphic Interface POP. b) Communication model

A computer is in charge of managing all scenario information. Haptic devices exchange continuously data with their controllers. A graphical display shows the object behaviour into the virtual environment. Figure 7a shows an example of this kind of virtual manipulation. This scenario can be used either by a user who manipulates a virtual objetc with both hands, or by two users manipulating the same object. Figure 7b shows the communication scheme. Haptic interfaces are linked to a controller connected to an Ethernet switch. Information is sent to a server that computes kinematics, evaluates an algorithm to detect contacts in the virtual world and controls the entire device. Once the server has all necessary data, it sends the corresponding commands to the haptic interfaces. Graphical information is also updated by the simulations given by the user hand movements.

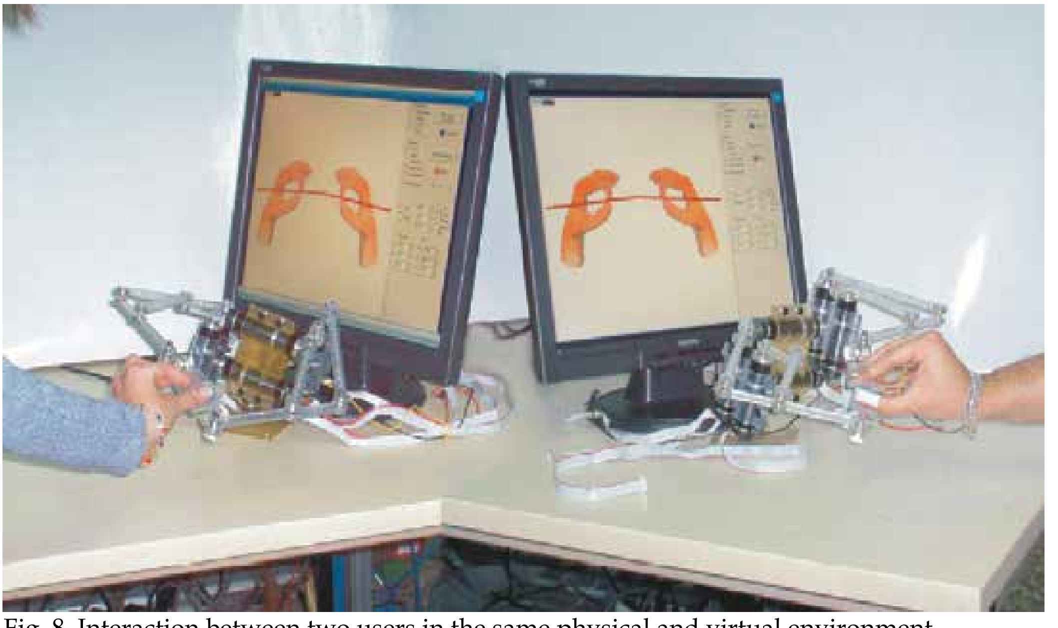


Fig. 8. Interaction between two users in the same physical and virtual environment

Figure 8 shows a further example of cooperative manipulation. In this case, two users are grasping the same virtual object. The objective is to manipulate a thin bar using thumb and index fingers.

5. Conclusions

The development of the MasterFinger haptic interface has demonstrated the multifinger haptic interaction relevance in the manipulation of objects and in the execution of cooperative tasks. The modular design of MasterFinger architecture allows this interface to easily scale up from 1 finger to 3 fingers. The MasrterFinger-2 shows a good behaviour as a haptic interface thanks to its low weight and inertia effects upon the user. It allows developing high realistic applications where one or more users are performing cooperative tasks.

Applications have proven the relevance of a multifinger device for properly grasping and manipulating virtual objects. It has required a distributed architecture to properly control the interaction in the virtual environment since many devices and processes, such as graphical displays, haptic devices and environment simulations are running at the same time. It represents a step forward for haptic applications since current environments are based on some devices linked to a stand alone computer. However, advanced developments for multifinger and multiuser haptic applications require a networked configuration in order to properly distribute processes.

6. References

Bouzit, M., Burdea, G.,Popescu, G., Boian, R.(2002). *The Rutger Master II - New Design Force-Feedback Glove*. IEEE Transactions on Mechatronics, vol. 7(2), (June 2002).

G. Burdea. (1996), *Force and Touch Feedback for Virtual Reality*: John Wiley & Sons, 1996, RES *(Ch. 6: Tactile Feedback Interfaces).* CyberGlove and Cybergrasp User's Guide, IMMERSION Corp. http://www.immersion.com/3d/docs/CGIIManual.pdf

Exos Co., (1993). TouchMaster and Force ArmMaster Specifications, Company Brochure, Woburn, MA.

H. Choset, K.M. Lynch, S. Hutchinson, G. Kantor, W. Burgard, L.E. Kavraki, and S. Thrun, *Principles of Robot Motion: Theory, Algorithms, and Implementation*. Cambridge, MA: MIT Press, 2005.

H. Kawasaki, T. Mouri, M. Osama, Y. Sugihashi, Y. Ohtuka, S. Ikenohata, K. Kigaku, V. Daniulaitis, K. Hamada and T. Suzuki. (2005) *Development of Five-Fingered Haptic Intereface: HIRO II,* Proc. of ICAT 2005, Christchurch, New Zealand, pp. 209-214, 2005 Immersion Corp., CyberTouch, CyberGlove and Cybergrasp User's Guides http://www.immersion.com/3d

J.M Sabater, R. Aracil and L. Payá . (2007) *A Novel Parallel Haptic Interface for Telerobotic System,* in Advances in Telerobotics, M. Ferre et al. Springer STAR series, pp. 45-60 2007

J.J. Craig (2005), *Introduction to Robotics: Mechanics and Control*, 3rd ed. Upper Saddle River, NJ: Pearson/Prentice-Hall, 2005.

L. Sciavicco and B. Siciliano. (2000), *Modelling and Control of Robot Manipulators*. New York: Springer, 2000.

Mark W. Spong, Seth Hutchinson, and M. Vidyasagar.(2006*). Robot Modeling and Control*,Wiley, 2006.

M. Ueberle, N. Mock, A. Peer, C. Michas, and M. Buss. (2004). *"Design and Control Concepts of a Hyper Redundant Haptic Interface for Interaction with Virtual Environments"*, Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, Workshop on Touch and Haptics, Sendai, Japan, 2004.

R. Murray, Z. Li, and S. Sastry. (1994), *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL: CRC Press, 1994. Sarcos Research Corporation. Sarcos Dextrous Arm. http://www.sarcos.com/ interspec_dexmaster.html

S. Kim, M. Ishii, Y. Koike, and M. Sato.(2000). *Design of a Tension BasedHaptic Interface: SPIDAR-G* Proceedings of World Multiconference on Systemics, Cybernetics, and Informatics: SCI 2000. pp 422-427, July 2000

Tan, H. Z., Srinivasan, M. A., Eberman, B., & Cheng, B. (1994). *Human factors for the design of force-reflecting haptic interfaces, Dynamic Systems and Control (ASME)*, 55(1).

T. H. Massie and J. K. Salisbury. (1994). *The PHANTOM Haptic Interface: A Device for Probing Virtual Objects. In: Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. Chicago (USA).

Tsai L-W.: Robot Analysis. (1999). *The Mechanics of Serial and Parallel Manipulators*. John Wiley & Sons, INCl, 1999.

8

Visual Information Presentation in Continuous Control Systems using Visual Enhancements

Jaekyu Park[1] and Sung Ha Park[2]
[1]*Korea Atomic Energy Research Institute,* [2]*Hannam University*
Korea

Every day, millions of people travel by land, water and air. Tracking and a continuous manual control are normally a critical part of any human-vehicle interaction. Also, because aircraft and ground vehicles are often moving at high speeds, the safety implications of transportation systems are tremendously important (Wickens et al., 1998).

In the case of transportation systems which are typical of tracking tasks, more than a million people are killed on the world's roads each year. For example, more than 40,000 people are killed on the roads of the United States each year. Traffic crashes also damage property, especially vehicles. By converting all these losses to monetary values, it is estimated that US traffic crashes in 2000 cost $231 billion, an amount greater than the Gross Domestic Product of all but a few countries (Evans, 1997). According to a report by the Ministry of Construction and Transportation and the Korea Transport Institute, there were 222,158 accidents in 2004, resulting in 7,013 deaths and 347,661 injured, including all traffic modes, in Korea. As a result, total accident costs added up to 14.5 trillion won which amounted to about 1.86% of Korea's 2004 GDP (Ministry of Construction & Transportation, 2005; Korea Transport Institute, 2006).

The types of systems and mechanisms people control in their jobs and everyday lives vary considerably from simple light switches to complex power plants and aircraft. Whatever the nature of a system, the basic human functions involved in its control remain the same. A human being receives information, processes it, selects an action, and executes the action. The action taken then serves as a control input to a system. In the case of most systems, there typically is some form of a feedback to a person regarding the effects of an action taken. In particular, because tracking tasks which are present in all aspects of vehicle control, including driving an automobile, piloting a plane, or steering and maintaining a balance on a bicycle require a continuous control of something, they are tasks that often involve complex information processing and decision-making activities to determine a proper control of a system, and these tasks are greatly influenced by the displays and dynamics of the system being controlled (Sanders & McCormick, 1993).

In this regard, effective interfaces in various visual displays which are vehicle information systems and aircraft cockpit displays inside an aircraft control room can be important factors to reduce the cognitive workload on human beings during a tracking task like a transportation system.

In this chapter, we describe the fundamental principles about a visual display design and investigate a visual enhancement that is influenced by the performance of visual tasks. Also, we address a representation of the visual information in a continuous control system through a case study related with a visual enhancement.

1. Introduction

1.1 Visual Information & Visual Cognitive Load

The accidents of a human machine system occur by not only the carelessness of human beings who are operators of a system and human causes such as the selection of an inappropriate behavior and action but also factors which are against a human being's ability. These accidents caused by these factors can be prevented by 'engineerging changes'. The engineering changes introduce a kind of fully automated system or redesign system in order to exclude people from implemanting inevitablily dangerous actions themself or to perform a task that is beyond a human being's ability (Lee, 1998). For example, an airbag to protect a drive or a passenger from the impact of a crash and a collision avoidance system to automatically stop a vehicle in the event of a collision are included in these engineering endeavors. However, many of the engineering changes sometimes cause additional problems for a human operator execution because of technical limits or costs etc. Therefore the mechanical-engineering elements and human factors must be harmonized properly rather than apply an engineering approach that does not consider a human operator's performance capacity and limitation. This approach is called an ergonomics approach and basically pursues this goal : proper harmony of human factors and mechanical factors, and is often called a human-centered design because the fundamental point of this approach focuses on human beings.

One of the important aspects for considering an ergonomics approach is the interaction methods of a human-machine: a bilateral relationship is established by efficiently making a `Connection' between a human being and a machine within one system. For example, all information that is related to a digression from a normal driving practise must be transmitted to a human's sensory organs through the dashboard in order to drive safely and efficiently in the given road situation, and the driver's efforts for modifying a deviation must be transmitted to the vehicle again. If we think deeply about the interaction of a driver and a vehicle, this interaction may be regarded as a human-machine interaction.

The human being obtains information and controls the system through this interaction. The information which is given to a human being is transmitted by the five senses: sight, hearing, touch, smell, and taste. Especially, the best method for transferring information to humans is visual information. This visual information transfers easily and quickly from a simultaneous perception of a large amount of information to humans. And most of the information among the human senses is inputted through the eyes (Dul & Weerdmeester, 2001). Especially, a situation to the effect that 90% of the information required for driving is visual, is common (Sivak, 1996).

However, this visual information is a burdening cognitive workload by gradually offering various and complex information. So, many researches for a utility of visual information and information that represent's a form of this information are performed. For example, if a driver is burdened by abundant or complex contents from navigation information, the performance of driving is worsened because the driver uses information offered by the car navigation system, evaluates it and makes a reasonable decision. Therefore other navigation informations should be limited as much as possible except navigation information considering positively necessary on the situation. The problem is deciding what information is necessary. Proper visual information can make a task such as driving and piloting better without requiring much capacity of display for a human information processing (Dingus et al., 1989).

1.2 Motor Control

Human operator should convert perceptual environment information in most systems. This behavior should be immediately after perceiving a stimulus. If this selection process of behavior is not accomplished rightly, a human error can be caused. A criterion for selecting a behavior depends on the level of automation. Nowadays, many systems have been automated, but a human operator at least should manually recover systems when an error has occurred in these systems. Therefore, it is necessary to regard a behavior selection through a motor control in order to configure a system by considering a human. Researches of human performance using a motor control have studied two aspects: skill approach and dynamic system approach. Not only do these two-aspects use different experiment subjects and analysis subjects, but also the environment for applying results from these experiments and analyzes is different. A skill approach mainly treats analog motor behavior. The behavior of this type is called an 'open-loop' because it is not necessary to treat visual feedback from the view of a human information processing model. In contrast to the skill approach, a dynamic system approach mainly treats a human’s ability which controls or tracks dynamic systems in order to adjust a particular spatiotemporal locus when there is an environmental uncertainty (Poulton, 1974; Wickens, 1986). Most transport controls fit into this category and are called a 'closed-loop' because these controls should treat feedback. These deal with a discontinuous control and a continuous control each, because of these two principles of a control. An open-loop control focuses on 'Fitt's law' through speed-accuracy trade-offs and helps to forecast information for a discontinuous control. In contrast, a closed-loop control offers information for forecasting a continuous control by describing how a human operator controls a physical system.

1.3 Tracking Task

Nowadays, most of the controls, from a simple control in daily life to a complex control in a complex system such as a nuclear power plant, have a closed-loop property. Especially, when facing a complex human situation and a complexity of human-machine system concerns from researches related to a perception movement skill or movement activity of a human, to engineering researches related to a tracking are increasing. This change in domain results from the great influence of three nonhuman elements on the performance of an operator.
(1) The dynamics of the system itself: how it responds in time to the guidance forces applied
(2) The input to the operator (the desired trajectory of the system)
(3) The display, the means by which operator perceives the information concerning the desired and actual state of the system
These elements interact with many of the human operator’s limitations to impose difficulties for a tracking in the real world. These limits in particular influence an operator’s ability to

track: processing time, information transmission rate, predictive capabilities, processing resources, and compatibility. A human-centered design will be accomplished through considering this aspect in a human-machine system design. Especially, from the aspect of information for estimating a control, a limitation of an operator will be a complement because visual information can transfer easy to humans.

2. Visual Display Design

2.1 Principles of Display Design

A display is a product to play the role of a interface so that visual information which is transferred by a system is cogitated by humans. Therefore, we have to consider a point for a vision, processing visual information and the relationship between human sensory and display properties because the display has an interfacing role between human and machine in order to transmit information. However, only one display tool can not harmonize with all tasks because the properties of a human user who performs a task are various. Main parameters which are essentially responsible for an optimum corresponding physical form of a display and something requiring a task a are series of principles about human perception and information processing. These principles are based on all the merits and demerits that human perception and information processing has (Wickens & Hollands, 2000; Boff et al., 1986), and whether the best display has occurred which depends on how well the result of the information analysis applies these principles.
The ergonomics principles for designing a display consist of four categories: principles of perception, principles of mental model, principles based on attention, and principles of memory. These principles include as follows (Table 1).

Category of principles	Case of principles
Principles of perception	Absolute judgement limits Top-down processing Redundancy gain Discriminability
Principles of mental model	Principle of pictorial realism Principle of moving part Ecological interface design
Principles base on attention	Information access cost Proximity compatibility principle
Principles of memory	Principle of predictive aiding Principle of knowledge in the world Principle of consistency

Table 1. Ergonomics principles related to display design

2.2 Visual information presentation

The methods for presenting visual information based on the principles of a display are various. However, as previously mentioned, it is necessary to consider the presentation methods which are used for a task because of the properties of the tasks and human beings. Especially, the presentation method is restrictive in a continuous control system which we are going to deal with in this chapter. Hence, it is important to find effective methods for reducing the visual load in order to offer visual information. This representative method of visual information is described through a visual enhancement in this chapter.

3. Visual enhancements

Several geometric scaling, enhancement techniques may assist users in their performance and be utilized to improve the interpretability of displays. And these enhancement techniques provide information on the magnitude of the errors which occur when observers are required to make directional judgments using perspective displays or 3D perspective displays. Visual enhancement techniques used in ergonomics are as follows.

3.1 Geometric scaling

The geometric scaling techniques have been applied from three aspects for enhacing visual information. One geometric scaling technique that may be applied to displays is that of a magnification (Wickens et al., 1989b). Repeated observations have been made that objects on a display seen as to be smaller or closer together than they really are (Meehan, 1992; Meehan & Triggs, 1988; Roscoe et al., 1981). As a result, these objects are perceived as being farther away from the observer than they really are. Another geometric scaling technique that has been applied to displays is an amplification of the vertical dimension of a display relative to the horizontal dimension (McGreevy & Ellis, 1991). The horizontal and vertical dimensions of an aviation display are usually asymmetrical. Finally, the technique of nonlinear scaling of an object size in relation to a distance may also be enforced in displays (Wickens et al., 1989b). As a result of the size-distance invariance relationship, images of objects that are very far away will appear as very small on the display.

3.2 Symbolic enhancements

Several symbolic enhancements which enhanced the effectiveness of a display have been used in a display design for an air traffic control in order to transfer spatial information (McGreevy & Ellis, 1985, 1991). The addition of a grid surface or ground plane to a display produced a marked improvement in the perception of the depth. The regular lines of the grid also served as an indicator of the horizontal distance between the objects in the display. A line which connected each aircraft to its true position on the ground plane made the relationship between each aircraft and the grid considerably clearer.

3.3 Visual cue for depth perception

The designer of a display faces a problem which is an appropriate implementation of monocular cues in a display so that it provides a user with an accurate sense three-dimensionality. Concerns that need to be considered include the number of monocular cues that should be selected and which cues to represent. There are usually monocular cues in the natural world such as: (1) light (luminance and brightness effects, aerial perspective, shadows and highlights, colour, texture gradients), (2) occlusion or interposition, (3) object size (size-distance invariance, size by occlusion, familiar size), (4) height in the visual field, and (5) motion (motion perspective, object perception). For example, in a perspective display various combinations of monocular cues may be utilized to create a perception of a depth. So, we have to consider how these cues interact with each other to create a perspective image.

3.4 Frame of Reference

The frame of reference that is provided to a viewer is also an important consideration in various display designs (Andre et al., 1991; Aretz, 1991; Barfield et al., 1992; Baty et al., 1974; Ellis et al., 1985; Harwood & Wickens, 1991; Olmos, Liang & Wickens, 1997; Rate & Wickens, 1993; Wickens et al., 1989b; Wickens et al., 1994, 1996; Wickens & Prevett, 1995). For example, in an egocentric display, the symbol representing ownship remains stationary while the flight environment moves around it. It has been proposed that the frame of reference that is implemented should be compatible with a viewer's mental model of their movement through the environment (Artez, 1991; Barfield et al., 1995b; Wickens et al., 1989a). Several studies have shown that this mental model may depend on whether a viewer is performing local guidance or global awareness functions.

3.5 Visual Momentum

Visual momentum refers to the visual landmarks that film editors generate to reduce a visual inconsistency among several scenes when editing a film (Park & Woldstad, 2006). The concept of a visual momentum has been expanded to demonstrate the design features of a display system and applied to integrate information among different displays (Woods, 1984; Aretz, 1991). These visual momentums provide perceptual landmarks to help human operators to maintain a cognitive representation among multiple displays.

4. Case Study

This case study was performed to investigate the effects of visual enhancements on the performance of continuous pursuit tracking tasks. Indicator displays with varying visual enhancements were presented on a CRT monitor. Human operators performed manual tracking tasks by controlling the cursor position with a mouse to pursue the motion of a horizontal bar on the indicator display. Quantitative assessments of different display conditions were made by using tracking errors and a modified Cooper-Harper rating as performance measures.

4.1 Experimental Design

We used a within-subject factorial design with three levels of visual enhancement and three levels of task difficulty as independent measures. Three visual enhancements were none visual enhancement (None), a shaded reference bar (Shade), and a translucent reference bar (Shade with line). The shaded reference bar and translucent reference bar were virtual cues overlaid on the horizontal bar of the indicator display, as shown in Fig. 1.

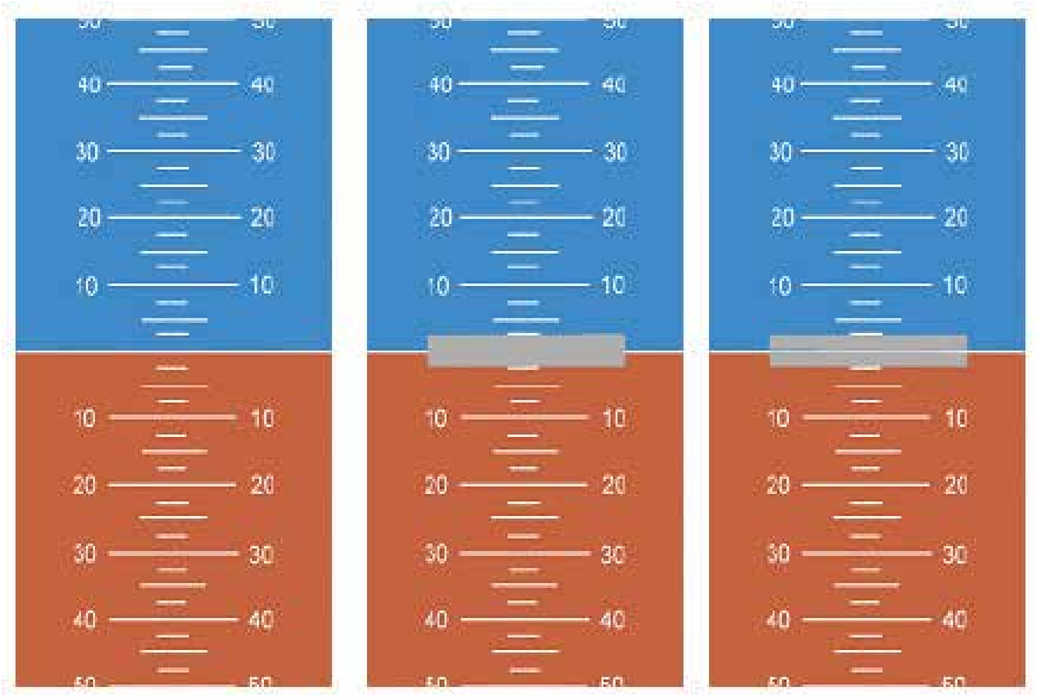


Fig. 1. Indicator displays and visual enhancement cues used in the experiment - From left, none visual enhancement(*None*), a shaded reference bar(*Shade*), and a translucent reference bar(*Shade with line*)

Three levels of a task difficulty were manipulated by changing the speed of the target (i.e., the horizontal bar on the indicator display). The difficulty of a task was adjusted by means of varying a subjects' workload. A preliminary study was conducted in order to tune the difficulty of a task. It was found that reliable changes in the difficulty level could be achieved by varying the speed of the target. As a result, three difficulty levels that were controlled by the speed of the horizontal bar were selected. The average speeds of the target for the low (Low), medium (Medium), and high difficulty (High) levels were 80, 100 and 120pixels/second, respectively. Dependent measures included tracking errors and subjective ratings of a workload. A tracking error was defined as the total number of pixels between the target and the cursor during the task. The order of the task condition within the blocks was counter-balanced across the subjects, in order to minimize the effect of learning.

4.2 Experimental procedure

Upon arrival for the experiment, participants were instructed to practice the tracking task with all the display configurations. Following an initial practice, participants completed the experimental tasks for the data collection. Participant's tracking data was measured for 60 second/condition. After completing each task, they rated their subjective workload using the modified Cooper-Harper rating scale. They were allowed to rest between trials, if necessary.

4.3 Results

The ANOVA results for the tracking errors showed significant main-effects of a visual enhancement, $F(2, 9)=13.663$, $p=0.0002$ and task difficulty, $F(2, 9)=8.619$, $p=0.0024$ (Table 2). The interaction of the visual enhancement and the task difficulty was not significant, $p=0.1161$.

Source	DF	SS	MS	F-Value	Pr>F
Subject	9	851.571	94.619		
Visual enhancement	2	78.742	39.371	13.663	0.0002*
Visual enhancement ×Subject	18	51.869	2.882		
Task difficulties	2	72.219	36.110	8.619	0.0024*
Task difficulties ×Subject	18	75.416	4.190		
Visual enhancement × Task difficulties	4	9.930	2.483	1.995	0.1161
Residual	36	44.799	1.244		

**: significant at α=0.05*

Table 2. ANOVA results for a visual enhancement and a task difficulty.

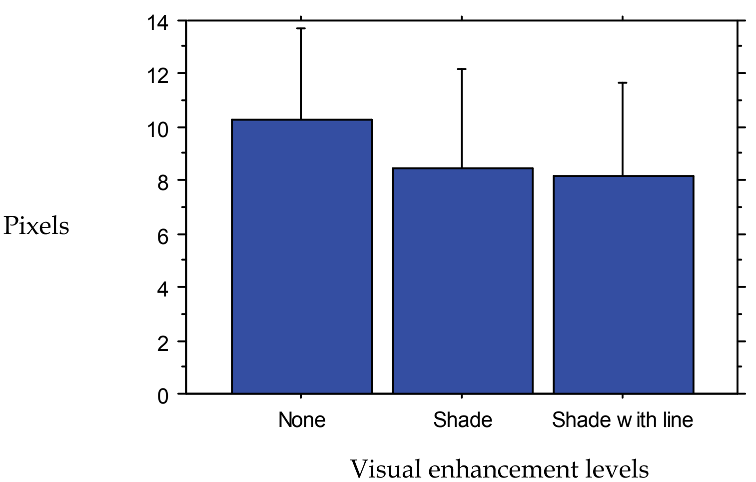


Fig. 2. The means of tracking error for the three visual enhancement conditions (*Unit: pixel*)

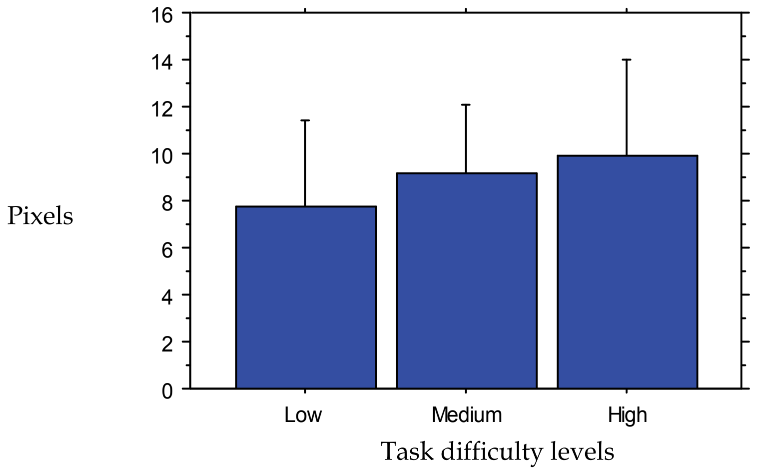


Fig. 3. The means of tracking error for the three task difficulty levels (*Unit: pixel*)

Student-Newman-Keuls comparisons of the means indicated that the none visual enhancement condition (None) resulted in the largest tracking errors and was significantly different from the shaded reference bar (Shade) and the translucent reference bar (Shade with line). The difference between the shaded reference bar (Shade) and the translucent reference bar (Shade with line) was not significant. The results imply that the shaded reference bar (Shade) and the translucent reference bar (Shade with line) were significant for improving a tracking performance. The tracking task employed in our study requires a frequent use of focused attention. We believe that the visual enhancement cues play an important role in augmenting visual information on a target location. Fig. 2 shows the mean tracking errors for the visual enhancement conditions.

Results of the mean comparisons also revealed that the largest tracking errors were committed in a highly difficult condition (High), followed by, in order, a medium difficulty (Medium), and a low difficulty condition (Low). Fig. 3 shows the mean tracking errors for the task difficulty conditions.

Source	DF	SS	MS	F-Value	Pr>F
Subject	9	71.883	7.981		
Visual enhancement	2	38.756	19.378	4.622	0.0240*
Visual enhancement ×Subject	18	75.467	4.193		
Task difficulties	2	42.022	21.011	11.278	0.0007*
Task difficulties ×Subject	18	33.533	1.863		
Visual enhancement × Task difficulties	4	3.644	0.911	1.528	0.2148
Residual	36	21.467	0.596		

**: significant at α=0.05*

Table 3. ANOVA results for the subjective workload.

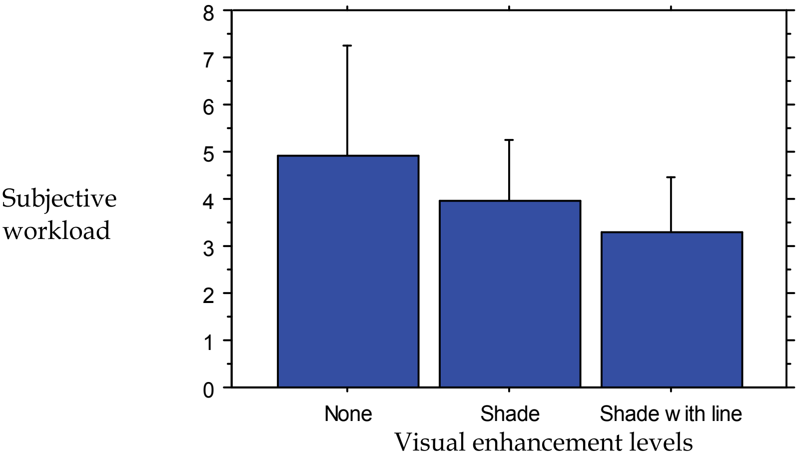


Fig. 4. The means of the subjective workload for the visual enhancement levels

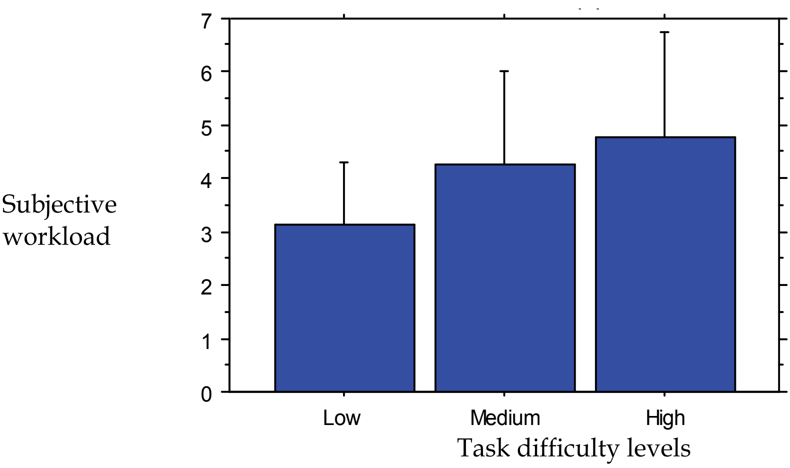


Fig. 5. The means of the subjective workload for the task difficulty levels

The ANOVA results for the subjective ratings of the workload also showed significant main-effects of a visual enhancement, $F(2, 9)=4.622$, $p=0.024$ and task difficulty, $F(2, 9)=11.278$, $p=0.0007$ (Table 3). The interaction of the visual enhancement and the task difficulty was not significant, $p=0.2148$. Student-Newman-Keuls comparisons of the means indicated that the translucent reference bar (Shade with line) was superior to the none visual enhancement (None). However, the difference between the shaded reference bar (Shade) and the none visual enhancement (None) was not significant.

For the task difficulty, performing the task with a highly difficulty condition (High) was judged to be more difficult than performing the task with medium (Medium) or low difficulty conditions (Low).

As previous results have mentioned, the results of ANOVA showed that the performance and subjective workload were significantly affected by the types of visual enhancements and task difficulties. Also, the results of a pair-wise analysis showed that the amount of deviation between the mouse pointer and the horizontal bar moving on an indicator were reduced by tendering visual enhancement cues. Particularly, the performance and subjective ratings were significantly improved in the case of providing a shaded reference bar (Shade) and a translucent reference bar (Shade with line). From the results of comparing the means for each level of the task difficulty, as the task difficulty increased, the degree of a deviation between the mouse pointer and the moving horizontal bar of the indicator were gradually increased. The low velocity (LOW) of a task difficulty was significantly different from the medium velocity (Medium) and high velocity (High). This results support previous findings that virtual cues can be utilized to provide additional visual information for the tasks requiring considerable attention such as a tracking task (Hardy & Lewis, 2004; Park & Koo, 2004).

5. Conclusion

This chapter was intended to identify and quantify the effects of visual enhancement cues on the performance of continuous control tasks such as tracking tasks. Also, we investigated the types and utilities of visual enhancements as visual aids that improve a performance and offer spatial information. Especially, we have indentified that various visual enhancements improve not only a performance but also the possibility of an error through a case study.

The findings of this chapter are applicable to the design of a head-mounted display (HUD) in the context of virtual environments. These findings can also be used as guidelines for designing visual displays for a continuous control system accompanied with a high speed manipulation such as those found in automobile and aircraft systems. Especially, the results of this case study could be applied to design the guidance for the information representation in an information system based HUD such as a Smart car which is an IVIS (In Vehicle Information System) developed by GM motors and Carnegie Mellon University.

In this chapter, when the continuous control tasks were performed through visual enhancements, it was assumed that the participants received visual cues from the same point of view. However, it didn't consider factors such as a depth perception and a pattern recognition of the subjects who were the main recipients of the visual information. Further studies are needed with considerations on the cognitive properties.

6. References

Andre, A.D., Wickens, C.D., Moorman, L. & Boschelli, M.M. (1991). Display Formatting Techniques for Improving Situation Awareness in the Aircraft Cockpit. *International Journal of Aviation Psychology*, Vol. 1, No. 3, pp. 205-218, ISSN: 1050-8414

Aretz, A.J. (1991). The design of electronic map displays. *Human Factors*, Vol. 33, No. 1, pp. 85-101, ISSN: 0018-7208

Barfield, W., Rosenberg, C., Han, S.H. & Furness, T. (1992). *A God's eye (exocentric) versus pilot's eye (egocentric) frame-of-reference for enhanced situation awareness*. Interactive Computer Graphics and Human Factors Laboratory, University of Washington, Seattle, WA

Barfield, W., Rosenberg, C. & Furness III, T.A. (1995). Situation Awareness as a Function of Frame of Reference, Computer-Graphics Eyepoint Elevation, and Geometric Field of View. *International Journal of Aviation Psychology*, Vol. 5, No. 3, pp. 233-256, ISSN: 1050-8414

Baty, D.L., Wempe, T.E. & Huff, E.M. (1974). A study of aircraft map display location and orientation. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 4, No. 6, pp. 560-568, ISSN: 0018-9472

Boff, K.R., Kaufman, L. & Thomas, J. (1986). *Handbook of Perception and Human Performance: Sensory Processes and Perception, Cognitive Processes and Performance*. John Wiley & Sons, ISBN: 0471829560, New York

Dingus, T.A., Antin, J.F., Hulse, M.C. & Wierwille, W.W. (1989). Attentional Demand Requirements of an Automobile Moving-Map Navigation System, *Transportation Research Part A: General*, Vol. 23, No. 4, pp. 301-315, ISSN: 0191-2607

Dul, Jan & Weerdmeester, B.A. (2001). *Ergonomics for Beginners: A Quick Reference Guide*, (2nd Ed.), Taylor & Francis, ISBN: 0748408258, London

Ellis, S.R., Tyler, M., Kim, W.S., McGreevy, M.W. & Stark, L. (1985). Visual enhancements for perspective displays: perspective parameters. *Proceedings of the International Conference on Systems Man and Cybernetics*, pp. 815-818, Tucson, Arizona, November, 1985, Institute of Electrical and Electronics Engineers, New York

Evans, L. (1997). A Crash Course in Traffic Safety. In: *1997 Medical and Health Annual*, E. Bernstein, pp. 126-139, Encyelopedia Britannica, ISBN: 0852296320, Chicago

Hardy, G.H. & Lewis, E.K. (2004). Cockpit Display of Traffic and Wake Information for Closely Spaced Parallel Approaches. *Proceedings of AIAA Guidance, Navigation, and Control Conference and Exhibit*, AIAA-2004-5106, Providence, RI

Harwood, K. & Wickens, C.D. (1991). Frames of Reference for Helicopter Electronic Maps: The Relevance of Spatial Cognition and Componential Analysis. *International Journal of Aviation Psychology*, Vol. 1, No. 1, pp. 5-23, ISSN: 1050-8414

Korea Transport Institute. (2006). *2003 Traffic Accident Estimate Costs in Korea*. Korea Transport Institute

Lee, J.S. (1998). The Human Factors Considerations on Design of In-Vehicle Navigation System, *'98 Pusan kyungnam Automotive Techno Center Workshop*, Pusan kyungnam Automotive Techno Center, Busan, February, 1998

McGreevy, M.W. & Ellis, S.R. (1985). *A perspective format for cockpit traffic displays. Tech. Report (Tech. Report NASA-TP85126),* NASA Ames Research Center, Aerospace Human Factors Research Division, Moffett Field, California

McGreevy, M.W. & Ellis, S.R. (1991). *Format and basic geometry of a perspective. display of air traffic for the cockpit (Tech. Memorandum 86680),* NASA Ames Research Center, Aerospace Human Factors Research Division, Moffett Field, California

Meehan, J.W. & Triggs, T.J. (1988). Magnification effects with imaging displays depend on scene content and viewing condition. *Human Factors*, Vol. 30, No. 4, pp. 487-494, ISSN: 0018-7208

Meehan, J.W. (1992). Imaging-display micropsia, experience, and ocular dominance. *In Society for Information Display International Symposium Digest of Technical Papers,* Vol. 23, pp. 301-303, ISSN: 0003-966X

Ministry of Construction & Transportation. (2005). *2005 Transportation safety enforcement plan*. Ministry of Construction & Transportation

Olmos, O., Liang, C.C. & Wickens, C.D. (1997). Electronic Map Evaluation in Simulated Visual Meteorological Conditions. *International Journal of Aviation Psychology*, Vol. 7, No. 1, pp. 37-66, ISSN: 1050-8414

Park, S.H. & Koo, J.M. (2004). Effects of Tele-Robotic Task Characteristics on the Choice of Visual Display Dimensionality. *Journal of the Ergonomics Society of Korea*, Vol. 23, No. 2, pp. 25-35, ISSN: 1229-1684

Park, S.H. & Woldstad, J.C. (2006). Design of visual displays for teleoperation, In: *International Encyclopedia of Ergonomics and Human Factors*, W. Karwowski, (2nd Ed.), pp. 1579-1583, Taylor & Francis, ISBN: 041530430X, London

Poulton, E.C. (1974). *Tracking skill and manual control*, Academic Press, ISBN: 0125635508, New York

Rate, C. & Wickens, C.D. (1993). *Map dimensionality and frame of reference for terminal area navigation displays: Where do we go from here? (Tech. Report No. ARL-93-5/NASA-93-1)*, Institute of Aviation, Aviation Research Laboratory, University of Illinois, Savoy

Roscoe, S.N., Corl, L. & Jensen, R.S. (1981). Flight display dynamics revisited. *Human Factors*, Vol. 23, No. 3, pp. 341-353, ISSN: 0018-7208

Sanders, M.M. & McCormick, E.J. (1993). *Human Factors in Engineering & Design, 7th ed.*, McGraw-Hill, ISBN: 007054901X, New York

Sivak M. (1996), The information that drivers use: is it indeed 90% visual?. *Perception* Vol. 25, No. 9, pp. 1081–1089, ISSN: 0301-0066

Wickens, C.D. (1986). The effects of control dynamics on performance. In: *Handbook of perception and performance*, K. Boff, L. Kaufman, & J. Thomas(Eds.), Vol. II, pp. 39-1-39-60, John Wiley & Sons, ISBN: 0471829560, New York

Wickens, C.D., Haskell, I., & Harte, K. (1989a). Ergonomic design for perspective flightpath displays. *IEEE Control Systems Magazine*, Vol. 9, No. 4, pp. 3-8, ISSN: 0272-1708

Wickens, C.D., Todd, S. & Seidler, K. (1989b). *Three-dimensional displays: perception, implementation and applications (Tech. Report No. CSERIAC SOAR 89-001)*, Crew System Ergonomics Information Analysis Center

Wickens, C.D., Liang, C.C., Prevett, T. & Olmos, O. (1994). *Egocentric and exocentric displays for terminal area navigation (Tech. Report No. ARL-94-1/NASA-94-1),* Institute of Aviation, Aviation Research Lab, University of Illinois, Savoy

Wickens, C.D. & Prevett, T. (1995). Exploring the dimensions of egocentricity in aircraft navigation displays: influences on local guidance and global situation awareness. *Journal of Experimental Psychology: Applied*, Vol. 1, No. 2, pp. 110-135, ISSN: 1076-898X

Wickens, C.D., Liang, C.C., Prevett, T., & Olmos, O. (1996). Electronic maps for terminal area navigation: effects of frame of reference and dimensionality. *International Journal of Aviation Psychology*, Vol. 6, No. 3, pp. 241-247, ISSN: 1050-8414

Wickens, C.D., & Hollands. J.G. (2000). *Engineering Psychology and Human Performance. 3rd ed.*, Prentice-Hall, ISBN: 0321047117, Upper Saddle River, NJ

Wickens, C.D., Lee, J.D., Liu, Y. & Gordon, S.E. (2003). *An Introduction to Human Factors Engineering*, (2nd Ed.), Prentice Hall. ISBN: 0131837362, Newark, NJ

Wierwille, W.W. & Casali, J.G. (1983). A Validated Rating Scale for Global Mental Workload Measurement Applications. *Proceedings of the Human Factors Society 27th Annual Meeting,* pp. 129–133, Norkfolk, October, 1983, Human Factors and Ergonomics Society, Santa Monica, California

Woods, D.D. (1984). Visual Momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies*, Vol. 21, pp. 229-244, ISSN: 0020-7373

9

Contact-free Stress Monitoring for User's Divided Attention

Dvijesh Shastri, Avinash Wesley, Ioannis Pavlidis
Computational Physiology Lab, Department of Computer Science, University of Houston, USA

1. Introduction

During many occasions, computer users and machine users in general have to perform more than one task simultaneously. When a combination of events demands critical decisions and rapid actions, the subject's alertness is raised. If alertness persists, it is likely to reach beyond certain acceptable levels and ultimately transform into stress. Stress due to the subject's divided attention may lead to degradation of his/her performance for one or more simultaneous tasks.

In this study, we describe research that aims to quantify stress levels of subjects due to divided attention. Such cases often arise in software use and beyond and take a serious toll on performance and emotion. The proposed method is based on the thermal signature of the face. We use the supraorbital skin temperature as the physiological variable of interest. Because of higher measurement sensitivity and its contact-free nature, facial thermal stress monitoring has been an increasingly popular approach (Puri et al., 2005, a. Pavlidis et al., 2002, b. Pavlidis et al., 2002). Contact sensor based physiological measurement methods restrict subjects' motion and increase their awareness of being monitored (Yamaguchi et al. 2006, Yamakoshi et al. 2007, Healey et al 2005). Therefore, it is not a very effective way for continuous physiological monitoring.

Although concurrent execution of multiple tasks is part of human life, no sufficient research has been done to understand its effects on human emotional states and performance. The purpose of this study is to evaluate a subject's emotional states and effect on performance while executing parallel tasks. We use simulated driving and concomitant cell phone conversation in our experimental design. This is a quintessential divided attention example in man-machine interaction with which most people are familiar with.

The results of our research show that the simultaneous performance of dual tasks increases blood flow in the supraorbital vessels and frontalis muscle. A change in blood flow alters heat dissipation from the supraorbital region and thus, it can be monitored through a thermal camera. This work opens a new area of research in non-contact stress monitoring for divided attention situations.

2. Methodology

During concurrent dual tasks performance, we have observed considerable skin temperature increase in the supraorbital region of all 11 subjects. This elevated temperature is the results of increased blood flow to the supraorbital region in order to supply energy for the increased mental activities. This finding matches our previous reporting that user stress is correlated with the increased blood flow in the frontal vessel of the supraorbital region (Puri et al., 2005). In the past, we used the periorbital region to quantify stress during startle response and polygraph examination (a. Pavlidis et al., 2002, b. Pavlidis et al., 2002, Tsiamyrtzis et al. 2006). As oppose to our requirement, that is to monitor sustained stress during the divided attention situations, the periorbital region is used to quantify instantaneous stress. Moreover, the users' continuous moving eyes during simulated driving prevent us from using the periorbital measurement in this study. As a result, we focus our attention to the skin temperature of the supraorbital region and its involvement in determining sustained stress.

Unlike the periorbital area, which accommodates a wide range of temperature values, the supraorbital area is plateau in nature. When tracks this plateau region, thin ranged feature points of the tracker compromise its stability to a certain degree (Dowdall et. al 2007). As a consequence, the tracker is shifted from its target position repeatedly. Therefore, we select a larger region of interest for the tracking algorithm but compute the mean thermal footprint of an appropriate subset of the region (Pogreška! Izvor reference nije pronađen.).

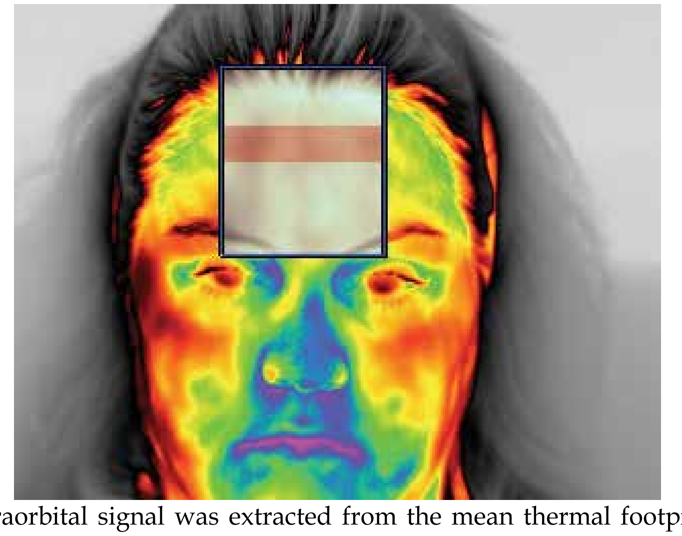


Fig. 1. The supraorbital signal was extracted from the mean thermal footprint of the pink colored region inside the rectangle.

For every subject, we select a Region of Interest (ROI) that covers the supraorbital area (see Figure 1). We compute the mean temperature of the ROI for every frame in the thermal clip. We, thus, produce a 1D supraorbital temperature signal from the 2D thermal data. However, due to imperfections in the tissue tracking and systemic noise, the measurement from this area carries substantial noise, which we suppress to a large degree by a Fast Fourier

Transformation (FFT) based noise cleaning approach (Tsiamyrtzis et al. 2006). Finally, we model the global trend of the noise-cleaned signal by fitting a linear polynomial to each experimental segment. Figure 2 illustrates the raw temperature signal, the noise cleaned signal, and the linear segment fitting. The slope value describes the temperature evolution of each segment.

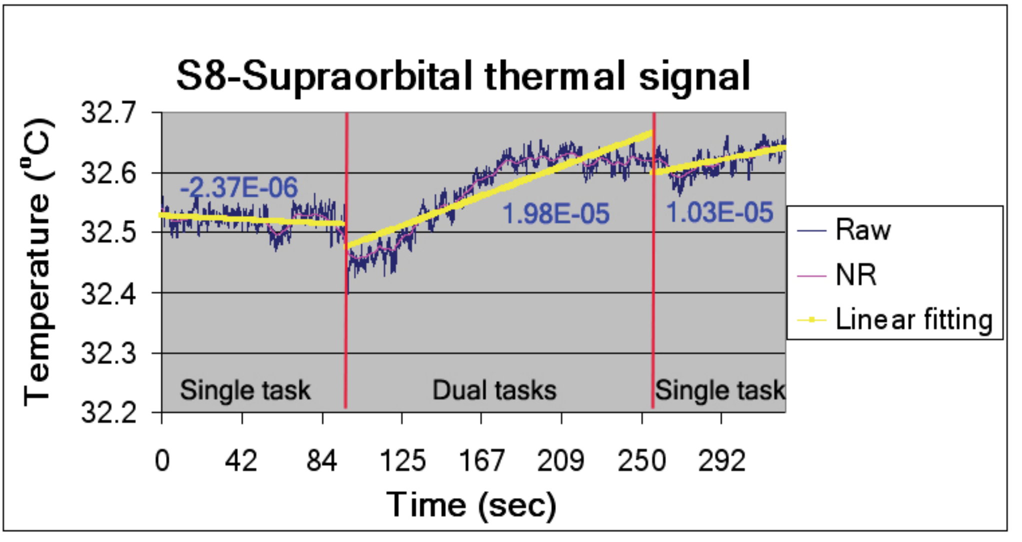


Fig. 2. The supraorbital raw temperature signal (marked in blue color), noise reduced (NR) temperature signal (marked in pink color) and linear fitting (marked in yellow color). Slope values of the respective segments are showed in blue colored text.

3. Experimental Design

We used a high quality Thermal Imaging (TI) system for the data collection. The centerpiece of the TI system is a ThermoVision SC6000 Mid-Wave Infrared (MWIR) camera (FLIR Systems) (NEDT=$0.025°C). For each subject, we recorded 3 thermal clips: while the subject resting, playing the driving simulation game, and the cooling off period. Thus, we collected 11 subjects x 3 clips/subject = 33 thermal clips.

The data set features subjects of both genders, different races, and with varying physical characteristics. The subjects were placed 6 feet away from the thermal camera (Fig. 3). We used a XBOX-360 game console and the *Test Drive: Unlimited* game to simulate real life driving. The subjects were asked to follow all traffic signs, drive normally and not to race during the experiment. They were given an opportunity to test drive before the experiment begun to facilitate themselves with the simulated driving setup. After the test drive, the subjects were asked to relax for 5 minutes before the experiments begun. This helps to isolate effects of other stress factors that the subjects may have carried from the past events. The subjects' facial thermal signature was recorded during this relaxation period. We called it *baseline* segment.

Next, the subjects were asked to play the driving simulation game. This part of the experiment lasted for approximately 5 minutes. After around a minute of the simulated driving (the *initial single task* segment), we made a cell phone call to the subjects and played a set of prerecorded questions in the following order:

Instruction: Please do not hang up until you are told so.

Q1: Are the lights ON in the room, yes or no?
Q2: Are you a male or female?
Q3: Who won the American civil war, the north or the south?
Q4: What is 11 + 21?
Q5: How many letter 'e' are in the word *experiment*?
Q6: I am the son of a mom whose mother in law's son hit. How am I related to the other son?
Q7: My grandma's son hit his son. How are the sons related?
Q8: A man is injured in 1958 and died in 1956. How is that possible?
Q9: What is 27 + 14?

Instruction: You may now hang up the phone and pay attention to the game.

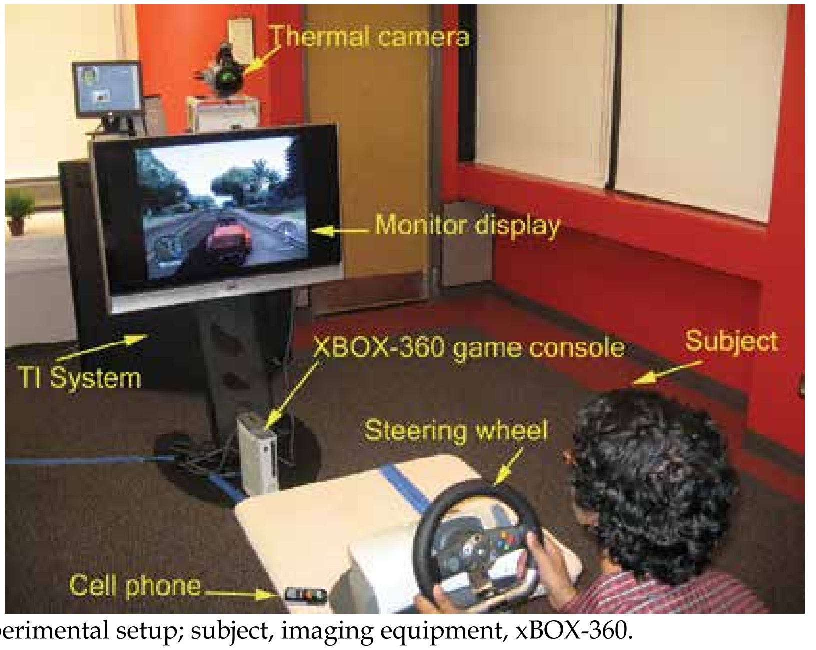


Fig. 3. Experimental setup; subject, imaging equipment, xBOX-360.

The question set was a combination of basic, logical, simple math and ambiguous questions. The order of the questions was designed to build-up emotional pressure on the subjects. Additional pressure was achieved by repeating one more time every question that was incorrectly answered. The subjects were supposed to drive while talking on the cell phone (the *dual task* segment). At the end of the phone conversation, subjects put the phone down and continued driving till the end of the experiment (the *latter single task* segment).

Finally, the subjects relaxed for 5 minutes. The purpose of this so-called *cool-off* segment was to monitor physiological changes after the simulated driving experiment.

4. Experimental Results

We used the slope as the thermal stress indicator. For each subject, we compute the slope value for every experimental segment as described in the Methodology section. Fig. 4 shows the mean slope values of the various segments for the entire data set (mean subject). The graph clearly indicates that the temperature increase during the concurrent dual tasks is the highest among all segments. Since the temperature increase is correlated to blood flow, the results indicate that more blood flows to the supraorbital region during the dual task action. With the exception of subject-6 (S6), the dual task segment of all subjects has a higher temperature gradient than its corresponding baseline and initial single task. This validates our assumption that the user's divided attention while critical tasks performance increases the user's stress level.

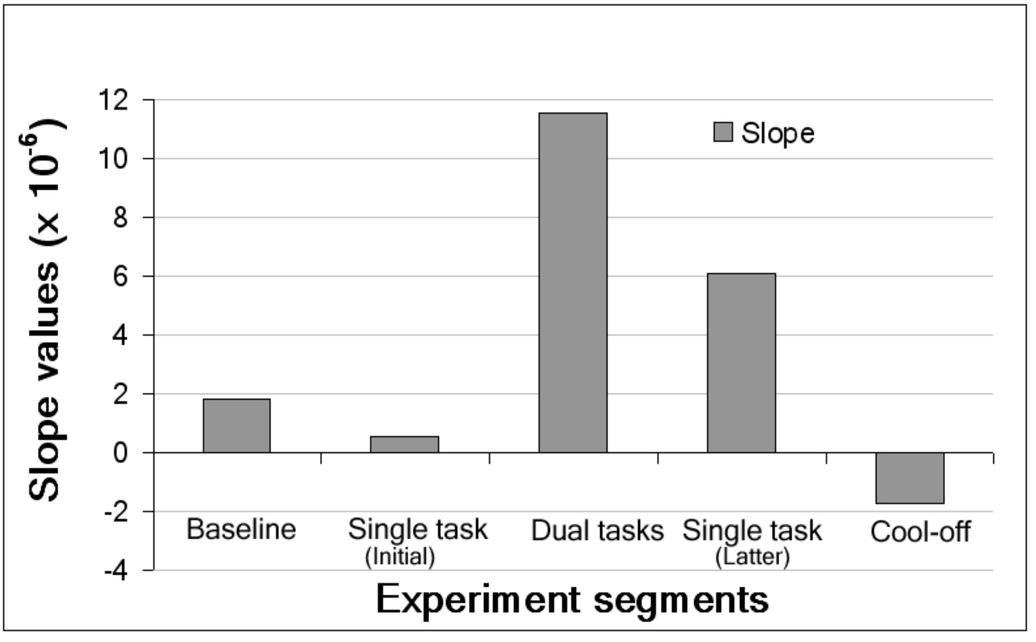


Fig. 4. Mean slope value of the experimental segments. The temperature increase is the highest during the dual tasks performance.

While performing the single task second time (Latter), most of the subjects carried the mental stress from the past dual tasks period. On the other hand, as the ablove graph illustrates, the subjects experienced less mental stress for the same single task performed first time (initial) as it was conducted right after the relaxation period (baseline) where they got an opportunity to isolate their mental stress due to the past events. Most of the subjects admitted during the post-debriefing session that they were thinking about their dual tasks performance while performing the latter single task. Therefore, the subjects experienced higher mental stress during the latter single task as compared to the initial single task. Figure 4 confirms this finding, mathematically. It illustrates that the rate of temperature change of the latter single task is higher than that of the initial single task.

The rationale behind higher slope value of the baseline segment as compare to the initial single task segment is that many subjects played the driving simulation game first time ever during the test drive. Thus, the baseline segment reflects the stress level they acquired during the test drive period, which was right before the baseline segment. The test drive helped isolating anxiety of performing a task that was never performed before. Therefore, the initial single task represents stress due to conducting the single challenging task only, i.e. driving the simulated car.

S6 is an interesting subject. The supraorbital temperature of the subject increased almost half of the dual task period and then decreased during the remaining period (Figure 5). We found the temperature decrease on the supraorbital was due to emotional perspiration. The cause for the emotional perspiration during divided attention is unknown to us at this point of our research. More experiments are required to reveal the full picture and currently we are pursuing it.

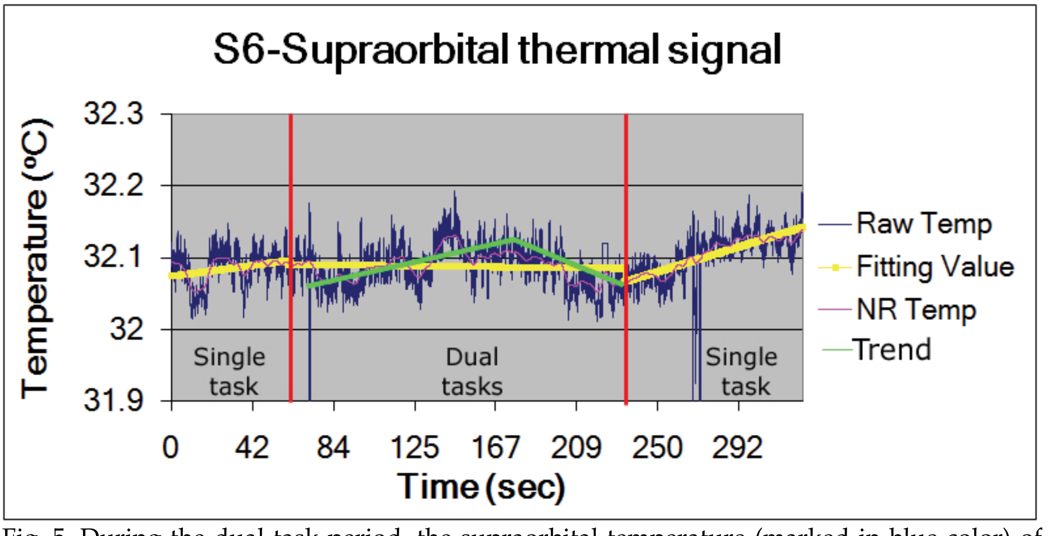


Fig. 5. During the dual task period, the supraorbital temperature (marked in blue color) of S6 shows ascending global trend in the first half and then descending global trend in the second half of the period as marked in green color. Therefore, the linear fitting to the segment is approximately a horizontal line.

In all cases, the rate of temperature change of the cool-off segment is much slower than that of the dual task segment. In most cases, the rate of the temperature change of the cool-off segment is slower than that of the initial single task, and the latter single segments. This illustrates that the subjects indeed felt relaxed after 5 minutes of intense mental activity.

Performance of the drivers degraded during the dual task segment, as measured by the point system of the simulator. This was inversely proportional to the average stress level measured through the supraorbital channel.

5. Conclusion

This research work demonstrates the feasibility of stress quantification in situations where the attention of the user is divided. Psychologically kosher (i.e., unobtrusive) quantification of stress and its correlation to user performance and emotion are of singular importance in man-machine interaction. We have also proved that talking on a cell phone during simulated driving increases the supraorbital skin temperature significantly. This finding clearly demonstrates that concurrent performance of two critical tasks increases user's stress level. Thus, we can safely claim that the proposed system is capable of reflecting the user's stress in divided attention situations. A feedback system can be devlope that alerts the users about their mental status based on the facial thermal signature . The initial experiment with a small dataset shows a lot of promise. More multi-faceted and extensive experiments, however, are necessary to understand the complete picture.

The proposed system can be use to monitor physiological behavior during critical multitasking activities. The potential use of our system is to understand the vehicle drivers' emotional states in order to reduce traffic accidents. The major bottelnack of this system is very high cost of the thermal camera, which prevents the system to be a practial applciation. We hope the cost of the thermal camera will reduce in the near future.

6. Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 0414754, entitled "Interacting with Human Physiology."

7. References

Dowdall, J., Pavlidis, I., and Tsiamyrtzis, P. (2007) . Coalitional tracking. *Computer Vision and Image Understanding 106*, 2-3, (2007), 205-219, ISSN: 1077-3142

FLIR Systems, 70 Castilian Dr., Goleta, California 93117, http://www.flir.com

Healey, J.A., and Picard, R.W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems 6*, 2 (6 2005), 156-166. ISSN: 1524-9050

a. Pavlidis, I., Eberhardt, N.L., and Levine, J. (2002). Human behavior: Seeing through the face of deception. *Nature 415*, 6867 (01 2002), 35.

b. Pavlidis, I., and Levine, J. (2002) Thermal image analysis for polygraph testing. *IEEE Engineering in Medicine and Biology Magazine 21*, 6 (11-12, 2002), 56-64, ISSN: 0739-5175

Puri, C., Olson, L., Pavlidis, I., Levine, J., and Starren, J. (2005). StressCam: non-contact measurement of users' emotional states through thermal imaging. *CHI' 05 extended* *abstracts on Human factors in computing systems*, pp. 1725–1728, ISBN: 1-59593-002-7 ACM press, Portland, OR, USA

Tsiamyrtzis, P., Dowdall, J., Shastri, D., Pavlidis, I., Frank, M., and Ekman, P. (2006). Imaging facial physiology for the detection of deceit. *International Journal of Computer Vision 71*, 2, (10, 2006), 197–214, ISSN: 0920-5691 (Print) 1573-1405 (Online)

Yamaguchi, M., Wakasugi, J., and Sakakima, J. (2006). Evaluation of Driver Stress using Biomarker in Motor-vehicle Driving Simulator. *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* pp. 1834-1837, ISBN 14244-0033-3, 30 Aug – 3 Sep. 2006, New York City, USA

Yamakoshi, T., Yamakoshi , K., Tanaka, S., Nogawa , M., Shibata , M., Sawada , Y., Rolfe , P., and Hirose , Y. (2007) A Preliminary Study on Driver's Stress Index Using a New Method Based on Differential Skin Temperature Measurement. *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* pp. 722 – 725, ISBN: 978-1-4244-0787-3, 22-26 Aug. 2007, France

10

Design Optimization of Pressure Sensing Floor for Multimodal Human-Computer Interaction

Sankar Rangarajan, Assegid Kidané, Gang Qian, Stjepan Rajko
Arts, Media and Engineering Program
Arizona State University Tempe, AZ 85287
USA

1. Introduction

Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use. Humans communicate with each other, intentionally or unintentionally, using various interpersonal communication modes such as static and dynamic full-body, limb, and hand gestures, facial expressions, speech and sounds, and haptics, just to name a few. It is natural to design human-computer interaction systems with which users can communicate using these interpersonal communication modes. To this end, multimodal human-computer interaction (MMHCI) systems are receiving increasing attention recently. An overview of the recent advances of MMHCI can be found in (Jaimes and Sebe, 2007). Our research mainly focuses on movement analysis based on visual and pressure sensing for movement based MMHCI, which read the movement of user(s), and respond accordingly through real-time visual and audio feedback. Such movement based MMHCI systems have immediate applications in a number of areas with significant impact on our daily lives, including biomedical, e.g. rehabilitation of stroke patients (Chen, et al., 2006), culture and arts, e.g. studying patterns and cues in complex dance performances, and interactive dance performances (Qian, et al., 2004), K-12 education, e.g. collaborative and embodied learning (Birchfield, et al., 2006), sports (e.g. analyzing and improving athletic performance based on weight distributions), and security (e.g. movement based smart surveillance systems), just to name a few.

Movement based MMHCI mainly deals with looking at dynamic characteristics of a person or a group of people such as joint angles, position of body parts, force and torque associated with limb movements, instantaneous velocity, acceleration and direction of body motion. In order to enable such a system to understand the user's movement robustly and accurately, it is important to augment the user's environment with novel sensors for accurate detection and estimation of the above movement qualities. It is worthy to understand that all the above movement qualities have underlying shape and/or effort attached which forms vital degrees of freedom for sensing modalities. Optical motion capture systems have become the obvious choice of researchers and technologists today for visual sensing of movement. However visual sensing alone is not sufficient for holistic inference of human movement since it can comprehend only shapes e.g. joint angles, orientation of body parts associated with human movement and give no clue about effort. Also visual sensing suffers from occlusion. Haptic sensing such as pressure sensing becomes inevitable for the above reasons as it aids to understand and comprehend the motivation driven physical effort attached to every movement and thereby exploring the inherent nature of the human body as a powerful communication medium.

Taking all the above factors into account, multimodal movement based human computer interaction system has been envisioned using both the pressure sensing floor (haptic) and motion capture system (visual) in order to perform holistic human movement analysis. The motion capture system that we use is commercially available and has been purchased for our research. However the pressure sensing floor is an in-house system developed specifically to address the research problem which thereby forms the core focus of this chapter. In this chapter, we present the system level description of pressure sensing floor followed by a discussion on hardware and software developments. Then we discuss the design methodologies for integration of the floor system with the marker based motion capture system as a first step towards the creation of an integrated multimodal environment.

2. Problem Statement

Pressure sensing system design targeting human computer interaction applications should confirm to certain requirements. In order to meet the sensing needs of such an application several design challenges need to be overcome. Firstly, the pressure sensing system should have a large sensing area to allow for unconstrained movement in the capture space. Secondly, high sensor densities are required for precise pressure localization and detailed analysis of pressure patterns. Thirdly high frame rate and low sensing latency are indeed critical for real time human computer interaction to capture rapidly changing human activities. It is worth mentioning here that there is a performance trade off between frame rate/ sensing latency and sensing area/sensor densities. Large sensing area with high sensor densities results in large number of sensors for scanning and data acquisition thereby decreasing the maximum achievable frame rate and increasing the sensing latency. Hence the performance optimization of the pressure sensing system to ensure large sensing area, high sensor resolution at reasonably good frame rate and low sensing latency is a major challenge. Fourthly, in many cases, there are only few users and large portion of the sensing space is not active at all. So proper data compression scheme to avoid network congestion and effectively utilize the given bandwidth poses a challenge. Fifthly, smart sensing systems should be inevitably equipped with context aware capabilities to sense the state of the environment and users and make a perception regarding the context of the environment or the user. Reliable person location tracking by clustering and tracking of active disjoint floor regions forms a vital part of perceiving context and emerges as a major implementation challenge. Finally, to allow movement based human computer interaction using multiple communication channels, such as gesture, pose and pressure distributions, the pressure sensing floor needs to be integrable with other sensing modalities to create a smart multimodal environment. Fast and accurate alignment of floor sensing data in space and time with other sensing modalities is another challenge. Furthermore, a need exists for a design of a modular and scalable system to allow for easy expansion and reconfiguration to suit external environments of different shapes and sizes.

In related prior work, various pressure sensing systems had been developed to capture and view pressure information associated with human movement across a floor. A detailed performance comparison study of those existing pressure sensing systems in terms of the above mentioned desired features are listed in Table 1.

Sensor System	University	Year	Sensing method	Sensing area (Sq.feet)	Frame rate (Hertz)	Sensor density (sensor/sq.inch)	Data resolution (Number of bits)	Integrability	Modular	Portable
MIT Magic Carpet	MIT Media Labs	1997	Piezoelectric wires	60	60	0.06	8	Yes	No	Yes
LiteFoot	University of Limerick Ireland	1997	Optical proximity sensors	42.25	100	0.3	NA	No	No	No
ORL Active floor	Oracle Research Lab	1997	Load cells	10.76	500	0.01	16	No	No	No
High resolution pressure sensor distributed floor	University of Tokyo, Japan	2002	Binary switch	43	15	10.57	1	No	Yes	No
Z Tiles	University of Limerick Ireland & MIT Media Lab	2004	FSR	NA	100	0.5	12	No	Yes	Yes
Floor Sensor system	University of Southampton, UK	2005	Binary switch	15.68	22	1.3	1	No	No	No
AME Floor I	Arizona State University	2004-05	FSR	9	10	0.44	8	Yes	No	No
AME Floor II	Arizona State University	2005-06	FSR	60	33	6.25	8	Yes	Yes	No
AME Floor - III	Arizona State University	2006-07	FSR	180	43	6.25	8	Yes	Yes	No

Color coding: 1, 2, 3

NA – NOT AVAILABLE

Table 1. Performance comparison table of existing pressure sensing systems

The ranking in each dimension (column) is color-coded such that the best system is in dark green, the second best in lighter green and the third in very light green. MIT magic carpet (Paradiso et al., 1997) and Litefoot (Griffith & Fernström, 1998) had fairly large sensing area and frame rate but were limited by poor sensor densities. ORL active floor (Addlesee et al., 1997) used load cells which lack the capability of detailed pressure measurement and cannot be used for applications requiring high sensor densities. High resolution pressure sensor distributed floor (Morishita et al., 2002) has the best sensor density so far but was a binary floor (poor data resolution) that just detects presence or absence of pressure and does not give any measurement of pressure values on an analog scale. Z-tiles floor space (Richardson et al., 2004) utilized a modular design, had high frame rate and data resolution but again suffers from low sensor density. Floor sensor system (Middleton et al., 2005) is a low cost design but again a binary floor with poor data resolution. Also most of the sensing systems except MIT magic carpet (Paradiso et al., 1997) were stand alone systems and lacked the capability to be integrated in a multimodal environment which is vital requirement for our application. In-shoe sensors (Paradiso et al.,

2000) have also been considered for force and pressure measurements but they have a limited scope of foot pressure measurement only. Also in-shoe systems tend to alter the subject's pressure application due to foot orientations by close contact.

It is quite obvious that all the sensing systems listed above have at least one serious limitation rendering it unsuitable to meet our application goals. It is worth mentioning that two generations of pressure sensing floor systems were developed with very similar goals as ours at the Arts, Media and Engineering (AME) Program at Arizona State University, namely, AME Floor I (Kidané et al., 2004) and AME Floor II (Srinivasan et al., 2005) listed at the bottom of the table. It is apparent from the comparison table that the second generation did see pronounced feature improvements over the first generation. AME floor I (Kidané et al., 2004) was a smaller prototype floor with 256 force sensing resistors arranged in less dense sensor matrix. During tests (Kidané et al., 2004) , it was found that there were large zones of no pressure detection during several activities. Also the scan rate was low deeming it unsuitable for real time human-computer interaction applications. These shortcomings were addressed by AME floor II (Srinivasan et al., 2005) with high sensor densities and high frame rate. Although AME Floor II (Srinivasan et al., 2005) showed significant advances and extended capabilities over AME floor I (Kidané et al., 2004) , it covered only a fraction of the sensing area required for our application, showed high sensing latency and lacked user friendliness. Also it showed preliminary multimodal integrable capabilities in temporal domain only and not spatial domain.

To fully address these issues, we have developed an improved, ingenious and in-house pressure sensing floor system (AME Floor-III) described in this chapter and listed in the last row of Table 1. AME Floor-III system is characterized by large sensing area, higher frame rate, smaller latency, enhanced user friendliness, spatial and temporal integrability with motion capture system to create a multimodal environment, modular/scalable design thereby matching our ideal pressure sensing demands for real time movement based human computer interaction. Comparison with other systems reveals that our proposed system in this chapter ranks among the top three in most of the dimensions of the performance criteria. Although there are four systems with frame rates higher than ours, the sensing area and sensor resolutions of these systems are much lower than our system. This chapter is an extension of our previous paper (Rangarajan, et al, 2007a) based on (Rangarajan, 2007b).

3. Pressure Sensing Floor Overview

This section provides essential information on pressure sensors, modular design approach used in building the large area pressure sensing floor. Later this section dives in deeper to explain the specifics of the embedded floor hardware and floor control software. Floor control hardware used in AME Floor-II (Srinivasan et al., 2005) has been retained in AME Floor-III but however the microcontroller firmware has been optimized to achieve high frame rate and reduced latency. Hardware overview given in this section creates a solid foundation to explain the optimization techniques in section 4.

3.1 Pressure Sensors: Force Sensing Resistors

Force sensing resistors have been used as individual sensor entities for AME Floor-III system. They are made up of pressure sensitive polymer between conductive traces on sheets of Mylar. As the name implies, these sensors exhibit a change in resistance when pressure or force is applied on them. The value of resistance is of the order of mega ohms under no pressure and drops to few kilo ohms when pressure is applied. Each pressure sensor element has an approximate sensing area of 6 mm x 6 mm and measure 10 mm x 10 mm including the non-sensing area. Such a small size paves way for dense aggregation of sensors in the sensing space thereby resulting in higher sensor densities. It is important to note that the force sensing resistors does not give very accurate measurements of pressure or force applied as there may be 15% to 20% variation between each other. Also they suffer from a property called creep or drift where the measured resistance values tend to slowly vary when subjected to constant pressure over a long period of time thereby inducing an error in pressure measurements. However force sensing resistors can be used very effectively for relative pressure measurements and acquiring pressure distribution data which serves the purpose of wide variety of applications such as medicine for diagnosis of various gait pathologies, automotive, robotics and interactive arts applications.

Fig. 1. Sub-floor steel framework (top left), surface floor wooden framework (top right), Complete view of AME Floor-III after assembly (bottom).

3.2 Pressure Sensing Mat

Force sensing resistors are generally available in several shapes and sizes like sensor pads, two dimensional sensor array matrix, continuous force sensing strips or several other forms depending on the application. Pressure sensing mat is a dense aggregation of force sensing resistors forming a two dimensional sensor array matrix. Tekscan 5315 pressure mat consisting of 2016 force sensing resistors arranged in grid of 42 rows x 48 columns have been used for AME Floor-III design.The dimension of each pressure mat is approximately 62 cm x 53 cm with an active area of 48.8 cm x 42.7 cm .The sensor mat is rated at 30 pounds per square inch (PSI). There are 2016 sensors in an active area of 322.98 square inch giving sensor densities of about 6.25 sensors/square inch.

3.3 Pressure Sensing Panel

Pressure sensing panel is constructed with eight such pressure sensing mats (Srinivasan, P., 2006). Eight Tekscan-5315 mats are arranged in 4 rows x 2 columns mounted on a wooden floor frame as shown in Fig. 1 (top right). Each pressure sensing mat has a non-sensing zone at the borders surrounding the active area. The pressure sensing mats are so laid and affixed on the floor panel in such a way that the active area of one mat overlaps the inactive area of another thereby avoiding such inactive zones (Srinivasan, P., 2006). Each pressure sensing mat has a connection tab where the pressure data of all the sensors collectively arrive. This connection tab passes through a slit on the front side of the panel and is back-folded to interface with hardware control board. Thus each panel has eight hardware control boards (one for each pressure sensing mat) mounted on the back side.

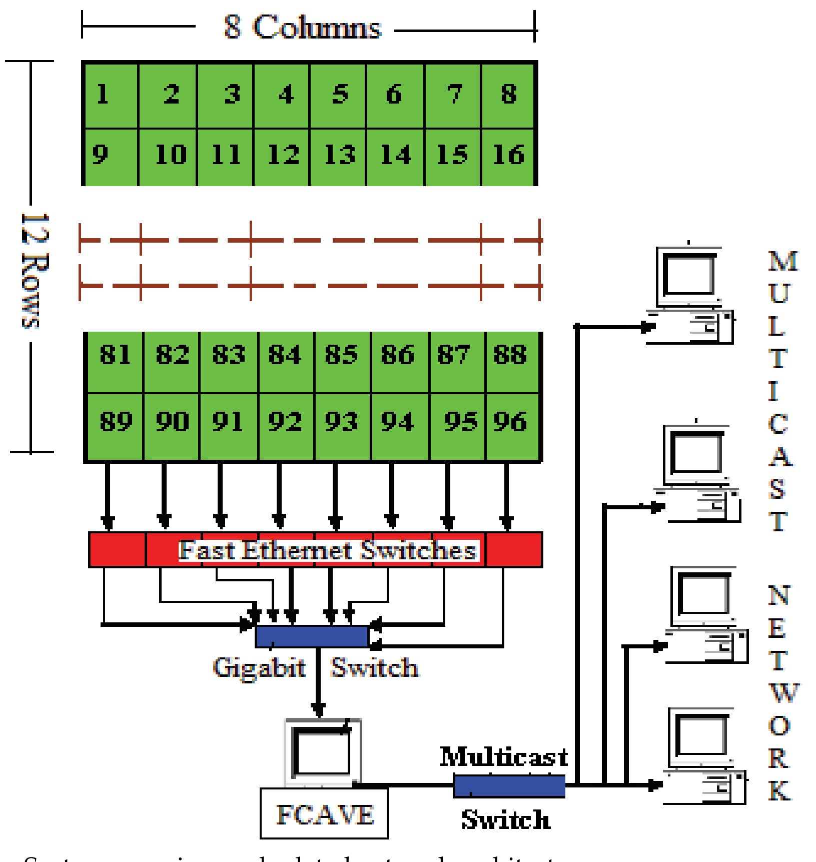


Fig. 2. Floor System overview and related network architecture

3.4 Large Area Pressure Floor

AME Floor-III is constructed by assembling 12 such pressure sensing panels (explained above) in 3 x 4 panel matrix. Thus the entire floor consists of a total of 96 networked pressure sensing mats assembled in 12 rows x 8 columns as shown in Fig. 2 and spanning a total sensing area of 180 square feet (15 feet x 12 feet). Such a modular design ensures large sensing area while still maintaining smaller frames for ease of use and installation. Also modularity in design paves way for creation of floor of different shapes and sizes (walkways, dance floor) and easy reconfiguration to suit external environments. The related network architecture used in AME Floor-III is illustrated in Fig. 2. All the 96 pressure sensing units are assigned static IP addresses and they form a local private network. Each and every pressure sensing unit has an associated hardware control board with an ethernet interface. There are two layers of network switches as shown in the Fig. 2. Multiple switches in multiple layers are deployed to share the network load and ensure sufficient leeway so that network switches are not operating to its rated full capacity which in turn increases performance and life time. All the twelve pressure units in one column are connected to a single fast ethernet switch on the first layer by means of ethernet cables. In a similar fashion, all pressure sensing units in 8 columns communicate with the fast ethernet network switch of their respective columns. The output port of eight fast ethernet switches is wired to the gigabit switch on the second layer. The output of the gigabit switch communicates with the host computer running the floor control software viz. Floor Control and Visualization Engine (FCAVE). FCAVE collects the pressure data arriving from 96 different IP's on 96 different ports and uses the source IP to identify and index the pressure data pertaining to different mats. The software further assembles all the 96 data packets (arriving from 96 mats) based on their location to create one large floor packet for each frame and sends it out to a multicast network. By this arrangement several ends users listening to the multicast network get access to the pressure data.

The mechanical design and installation of AME Floor-III is implemented in three layers namely the sub floor framework, surface floor (shown in Fig 1) and marley layer. The sub floor framework forming the bottom most layer is constructed using long steel rails welded to form a grid like structure and mounted on wooden blocks. This layer serves as a raised pedestal for the entire floor giving an elevation of approximately 4 inches above ground and provides the required spring and resilience to prevent injuries due to user activity like falling, jumping etc. Also such a raised installation paves way for all the necessary interconnect, ethernet wiring, power distribution and cabling to be housed beneath the floor in a neat and coherent fashion. The surface floor is made of a solid wooden framework and made to rest on the sub-floor layer. This layer forms the solid rigid structure supporting the users on the floor system. The pressure sensing mats and the hardware control circuitry for data collection are affixed to the surface floor structure on the frontal and dorsal side respectively. The third and the topmost layer is sheet of marley which is a vinyl surface, covering the entire area of the floor. The marley serves two main purposes. Firstly it aids in protection of the sensor matrix which are easily susceptible to damage by sharp and pointed objects and thereby increasing the longevity of the sensors. Secondly, it provides the necessary friction and contact grip for the subjects thereby preventing slips and fall injuries. Marley surface is generally preferred over standard wood or tiled surface structures for better movement control and less slipperiness.

3.5 Floor Control Hardware

The hardware control circuitry used in AME Floor–II (Srinivasan, P., 2006) has been retained in AME Floor –III but the microcontroller firmware has been optimized in AME Floor-III to achieve a higher frame rate. The floor hardware (Srinivasan, P., 2006) comprises of microcontroller, multiplexers, A/D converter and ethernet enabled rabbit controller which are all wired together on a hardware control board and collectively termed as 'mat based controller'. The block diagram of the floor hardware (Srinivasan, P., 2006) is shown in Fig 3.

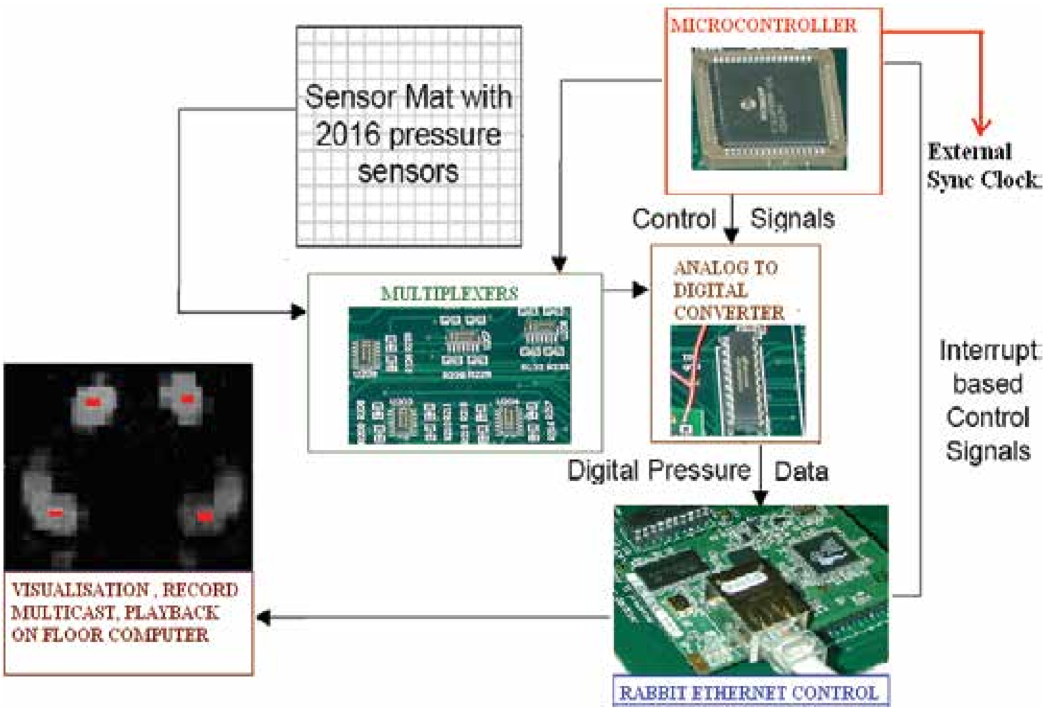


Fig. 3. System level block diagram of the mat-based controller

The microcontroller (PIC18F6585) forms the heart of the mat-based controller which generates the timing and control signals for all the components on the hardware control board to coordinate and sequence their operation of scanning sensors and reading pressure values. It has programmable capabilities to synchronize the sensor scan based on an internal timer or from an external clock signal. The latter has been currently implemented whereby the scan of all 2016 sensors on a single mat are synchronized with the external clock from the motion capture system. This implementation paves way for temporal synchronization of AME Floor-III and motion capture system for multimodal sensing. At the onset of falling edge of the synchronization clock, the microcontroller initiates a sequential scanning process of 2016 sensors arranged in 42 x 48 matrix. The pressure sensors (force sensing resistors) indicate a change in resistance when pressure is applied. This change in resistance is converted to a proportional analog voltage by a simple resistor divider network. Signal multiplexing has been implemented using a bank of six CD74HC4067 16-to-1 multiplexers to read the pressure voltage signals. Three multiplexers are used for the row lines and three for the column lines of the sensor matrix and each input line of the multiplexer is wired to a single pressure sensor output. The microcontroller streams out the multiplexer selects signals in a sequence to read the pressure values from sensor 1 to sensor 2016 one at a time. A bunch of gain control operational amplifiers follows the multiplexers which performs the necessary amplification and signal conditioning of the analog voltage. The outputs of the operational amplifiers are fed to a high speed ADC0820 converter which coverts the sensed analog voltage to 8 bit digital pressure value on the reception of the read/start- conversion signal from the microcontroller. Digitized pressure data is transferred to RCM 3200 module (Rabbit controller) through an input port after interrupt enabled handshaking signals with the microcontroller. The RCM3200 module contains Rabbit 3000 processor running at 44.2 MHz and 10/100 Base-T Ethernet connectivity. It is worth mentioning again that the multiplexer select signals, ADC read signal and rabbit interrupt signals are all generated by the microcontroller which are the major control signals used to synchronize/sequence the operation of the components on the hardware control board. Rabbit units are assigned unique static IP addresses. The Rabbit module collects 8 bit digital pressure data of all 2016 sensors and assembles them to create a pressure packet pertaining to that mat. It attaches a frame number at the end of each pressure packet and sends it out onto the network to the host computer running the floor control software through an array of switches. The host computer software listens to the IP addresses and port numbers to which the rabbit has been programmed to send the pressure data thereby collecting pressure data for further processing.

3.6 Floor Control and Visualization Engine (FCAVE)

Floor Control and Visualization Engine (FCAVE) software developed at the host computer has an interactive graphical user interface with various control buttons and indicators (Fig. 4) and it is programmed to respond dynamically to user input. This software receives the raw pressure data packet for each mat separately, assembles the data of all 96 mats, assigns an incremental frame number and creates floor data frame which is ready for processing.

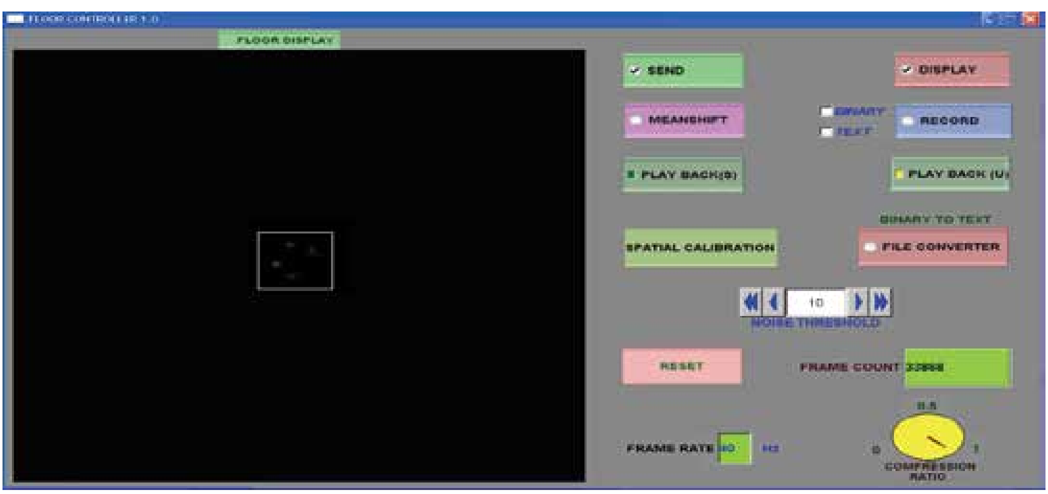


Fig. 4. Graphical user interface of Floor Control and Visualization Engine

FCAVE software has two operating modes namely 'live mode' and 'playback mode'. As the name implies, real time data collection and processing is done in the 'live mode' whereas offline data processing from a recorded pressure data file is usually done in the 'playback mode'. Furthermore playback can be done in synchronous and asynchronous ways. Synchronous playback streams the recorded pressure data synchronous with the motion capture playback stream. Asynchronous playback streams the recorded pressure data at the desired frame rate without any synchronization with motion capture system. FCAVE also offers various other controls like multicast pressure data to users on network, grayscale display of pressure information, set noise filter value, perform mean shift tracking of pressure clusters ,frame counter reset , record to file etc. FCAVE software development paved way for enhanced user-friendliness (with a lot of features as shown in Fig. 4), efficient data compression and mean shift tracking of active, disjoint pressure clusters in real time.

4. Hardware and Software Developments

This section mainly talks about hardware improvements done on AME Floor-II (Srinivasan, P., 2006) and new software developments to result in AME Floor-III. AME Floor II (Srinivasan, P., 2006) operated at frequency of 33 Hz and also suffered from significant latency of 200 milliseconds. Latency experiments are done to measure and quantify the latency along the data path and further optimizing them for latency reduction. Hardware optimizations in AME Floor-III eventually lead to increased frame rate (33 Hz to 43 Hz) , reduced mean latency (200 ms to 25 ms) and improved real time performance over its precursor AME Floor-II (Srinivasan, P., 2006).New software developments like data compression and mean shift tracking have imparted context aware capabilities to the system. This section elaborates on the hardware optimization techniques used to reduce latency and increase frame rate and new software developments namely data compression and mean shift.

4.1 Optimization of System Latency

Small latency is critical for real time sensing systems used in human-computer interaction applications. Latency is defined as the time lag between the time instant of the true event and the time instant the pressure data pertaining to the true event arrives at the end users on a multicast network. The overall system latency is the sum of two components namely intrinsic latency and extrinsic latency. Intrinsic latency is defined as the latency induced by the sensor scanning process. Each sensing unit has a pressure mat with 2016 sensors and an associated mat based controller for pressure data collection and signal conditioning. All sensors are scanned sequentially from sensor 1 to sensor 2016 to read the pressure values. There is an inherent delay for the scanning process to complete and pressure packet to be produced. This delay is called as the intrinsic latency which is present due to lag in various hardware components on the mat based controller. The microcontroller generates the sensor scan signals and the scan routine incorporates all the hardware component delays. Thus the total execution time of the microcontroller scan routine T_{scan} determines the frame rate F ($F = 1/T_{scan}$) of the system. After a complete mat scan of 2016 sensors, the pressure data packet for that mat is produced. Extrinsic latency is defined as the time taken for such a pressure data packet to reach the end users on the multicast network and it accounts for the network transmission delay and FCAVE software delay. Due to sequential scanning process, the intrinsic latency is direct function of the active sensor location given by a sensor address (An active sensor would be one that has pressure applied on it and sensors are addressed sequentially from 1 to 2016). A mathematical relationship is first established which gives an expected range of the intrinsic latency values based on the system scan rate and active sensor location. From this theoretical model, it becomes apparent on what latency distribution to expect when pressure is applied on a particular sensor location and later latency experiments are done to verify the same. The following section presents the mathematical relationship between intrinsic latency, frame rate and active sensor location.

4.1.1 Theoretical approach – relationship of intrinsic latency, active sensor location and frame rate

Let's assume that the system is running at a frame rate F and the time taken for one complete scan cycle of N sensors (N = 2016 in our case) is T_{scan}. Pressure sensors applied with active load are defined as active sensors. Let L be the address of such an active sensor. The intrinsic latency related to this sensor at L needs to be determined. Let U be the address of the sensor currently being scanned at the time instant when the pressure application occurs on sensor L. Let X_L and X_U be time elapsed since the start of the scan until the sensor L and sensor U are reached respectively by the scan routine, i.e.

$$X_L = 1/\mathrm{N} \times L \times T_{scan} \tag{1}$$

$$X_U = 1/\mathrm{N} \times U \times T_{scan} \tag{2}$$

According to the relationship between X_L and X_U, there are two different cases to be considered which are pictorially represented in Fig 5.

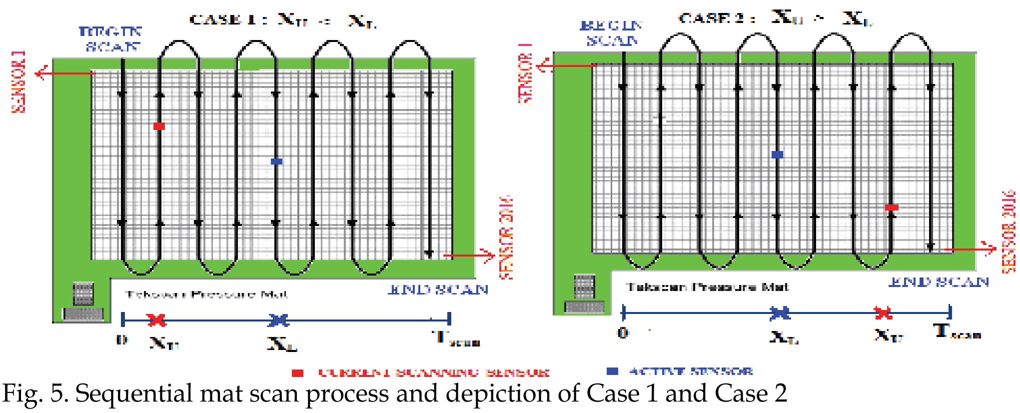


Fig. 5. Sequential mat scan process and depiction of Case 1 and Case 2

Case 1: $X_U \leq X_L$, pressure applied on sensel L is registered in the current scan cycle.
Case 2: $X_U > X_L$, pressure applied on sensel L is registered in the next scan cycle.
Hence, given L, the intrinsic latency τ caused by system scan is a function of X_U,

$$\tau(X_U) = \begin{cases} T_{scan} - X_U, \; when\, 0 \le X_U \le X_L \\ 2T_{scan} - X_U, \; when X_L < X_U < T_{scan} \end{cases} \quad (3)$$

Since X_U assumes a uniform distribution in $[0, T_{scan}]$ it can be easily shown that τ is uniformed distributed in the range given below

$$T_{scan} - X_L < \tau \le 2T_{scan} - X_L \quad (4)$$

Therefore, the mean intrinsic latency for the sensel at L is given by

$$\tau_m = 1.5T_{scan} - X_L \quad (5)$$

Thus the mean intrinsic latency is a direct function of T_{scan} and active sensor location X_L. Furthermore, since L can also be treated as a uniform random variable between 1 and N, the mean average intrinsic latency of all sensels on a mat is given by

$$E\{\tau_m\} = 1.5T_{scan} - E\{X_L\} = T_{scan} \quad (6)$$

Equation 6 clearly reveals that the intrinsic latency depends on the frame rate F ($1/T_{scan}$) and active sensor location X_L. As expected, the intrinsic latency decreases as the frame rate is increased. Equation 6 implies that as the active sensor location becomes closer and closer relative to the end of mat, the intrinsic latency decreases linearly. This can be justified by the sequential nature of the scanning process. Latency experiments have been conducted (explained in the following section) to verify the above statement and check the validity of Equation 6.

4.1.2 Experimental approach for measurements of system latency

The experimental set up shown in Fig 6 is used to measure both intrinsic and extrinsic latency and thereby overall system latency.

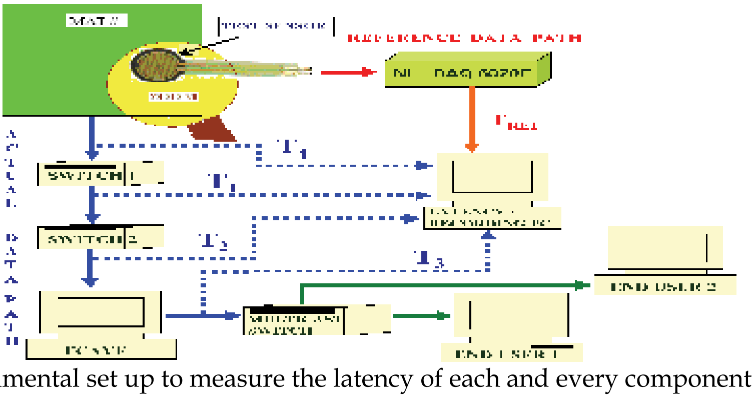


Fig. 6. Experimental set up to measure the latency of each and every component along the data path.

In order to measure the latency, a time reference is required which gives the time instant of the true event (e.g. pressure strike). The time of arrival of the actual data packet pertaining to the true event is then recorded and the displacement in time between the time reference and arrival time gives the measure of latency. A single physical test sensor (shown under zoom lens in Fig 6), National Instruments data acquisition hardware (NI-DAQ 6020E) and Labview application are used to get the time reference of the true event. The test sensor is placed directly above the sensor on the mat on which the pressure is going to be applied (active sensor). The Labview application is programmed to read the incoming data from two input ports namely reference data port and actual data port. The test sensor output feeds to the NI-DAQ hardware and in turn to the reference data port of the Labview software application to create the reference data path. The components in the reference data path are chosen to be relatively fast and responsive to give a solid reference for accurate latency measurements. Actual data path is the normal data flow through the switches and floor control software given by the system architecture. Actual data port of the Labview software is connected to fetch the data anywhere along the actual data path as shown in Fig 6. Pressure applied by a swift strike on the test sensor is the event used in the experiment. Sensor beneath the test sensor suffers the event at the same instant of time as that of the test sensor and hence the test sensor can be used a reference. When an event occurs, two different channels (reference data path and actual data path) carry information about the same event to the Labview software application. Labview software reads the data from the test sensor arriving at the reference data port and records the arrival time. Under the assumption that the transmission delays along the reference data path are at negligible levels, the reference time stamp gives the time instant of the true event. Also the active sensor on the pressure mat transmits the event through the actual data path to the actual data port of the Labview application. Labview records the arrival time of the actual data packet as well. The time displacement between the actual data arrival time and reference data arrival time is computed by the Labview application as a true value of latency.

Different read out points namely (T_1, T_2, T_3, T_4) are taken to measure and quantify the latency at each and every point along the data path. Intrinsic latency is obtained from T_4 and T_{REF} values. Extrinsic latency is mainly caused by the various components in the data path like the two network switches and host computer running the floor control software. T_1,T_2 and T_3 measurements are used to quantify the latency added by switch 1, switch 2 and floor control software respectively using the formulas given below.

$$T_{INL} = T_4 - T_{REF} \tag{7}$$

$$LATENCY_{SW1} = T_1 - T_{INL} - T_{REF} \tag{8}$$

$$LATENCY_{SW2} = T_2 - T_1 \tag{9}$$

$$LATENCY_{FCAVE} = T_3 - T_2 \tag{10}$$

Extrinsic latency measurements done on AME Floor-II (Srinivasan, P., 2006) revealed a major contribution of 167 milliseconds ($LATENCY_{FCAVE}$) from floor control software and negligible additions by the network switches ($LATENCY_{SW1}$ & $LATENCY_{SW2}$).

The floor computer was then upgraded to dual processor, dual core and multithreading techniques were used to improve the real time performance of floor control software and reduce extrinsic latency to negligible levels in AME Floor-III.

Having reduced the extrinsic latency to negligible levels, focus is shifted to intrinsic latency reduction. This experimental set up is further used to get empirical measurements of intrinsic latency T_{INL} and validate the mathematical model derived in section 4.1.1. Pressure is applied on a set of fixed sensor locations on the mat and the mean system latency is computed for over 100 trials for the floor system running at 40 Hz. Equation 6 gives the theoretical estimate of the mean latency given the active sensor location and frame rate. Fig 7 shows the correlation between the mean latency values computed from theoretical and practical data sets when the system is running at 40Hz. The offset between theoretical and practical values is mainly due to the DAQ polling frequency by Labview application. It is found that Labview application polls the data acquisition card (NI-DAQ 6020E) at 5 millisecond intervals (DAQ polling error) on an average. Hence the time reference T_{REF} is delayed from the true value by a time period *t*, where *t* is a random variable (*0 ≤ t ≤ DAQ* polling error).This explains why the practical value of latency is less than the theoretical value by an offset '*t*'. In other words the offset or mean error between the theoretical and practical data sets should always be less than or equal to the DAQ polling error which is proved by means of Table 2.

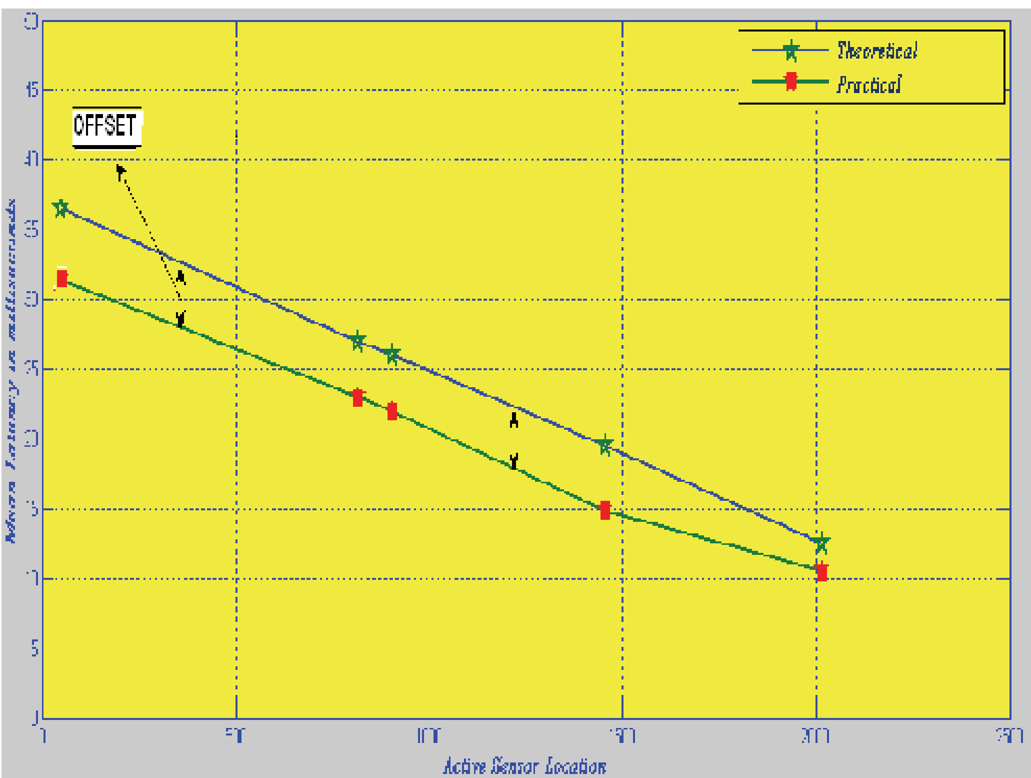


Fig. 7. Plot of mean latency (ms) vs. active sensor location for theoretical and practical data.

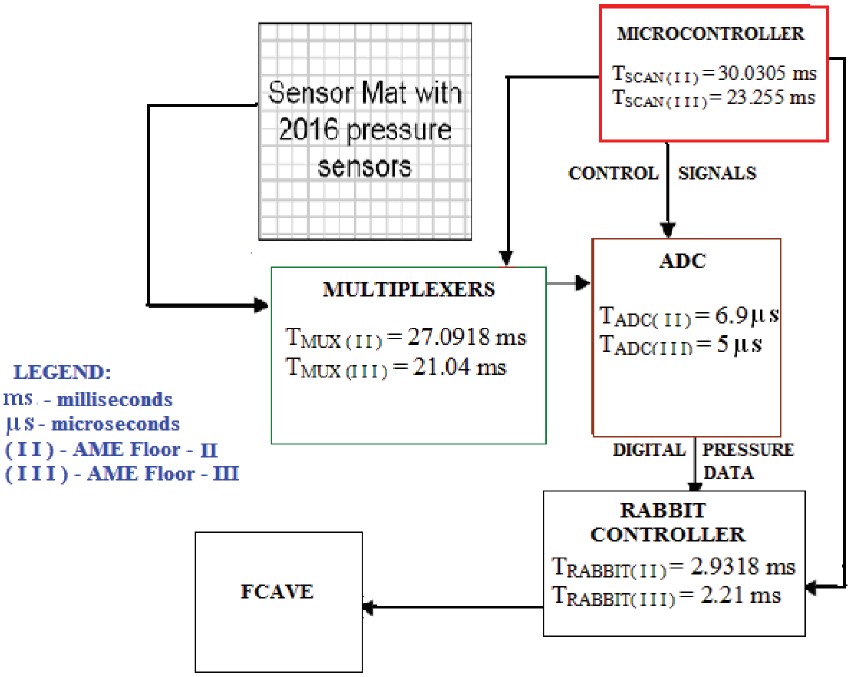
Active Sensor Location (0-2015)	Theoretical Mean	Practical Mean (> 75 Trials)	Mean Error	DAQ Polling error
48	36. 5	31. 36	5.14	6
815	27	23	4	4
905	26	22	4	5
1455	19. 5	14. 88	4.62	5
2015	12. 5	10. 53	1.97	4

Table 2. Mean error between theoretical and practical data sets for different sensor locations

Equation 6 states that we can minimize intrinsic latency by minimizing T_{scan}, or equivalently maximizing the frame rate ($F = 1/T_{scan}$). Hence efforts were invested to increase the frame rate and reduce intrinsic latency which is described in section 4.2.

4.2 Maximization of Frame Rate

Frame rate of floor system is determined by the speed of hardware components on the hardware control board. Every hardware component has certain delay or lag associated with it. The microcontroller scan routine incorporates all the hardware component delays and accordingly generates the control signals. The sum of all hardware component delays gives minimum T_{scan} required whose reciprocal gives the maximum achievable frame rate.Fig 8 shows the block diagram of floor hardware (Srinivasan, P., 2006) annotated with delay values for each hardware component explaining how we had achieved a maximum frame rate of 43 Hz in AME Floor- III from an old value of 33 Hz in AME Floor-II prototype (Srinivasan, P., 2006). It is important to note that suffix (II) on Fig 8 refers to AME Floor-II whereas suffix (III) refers to AME Floor-III system. The block diagram quantifies the time savings obtained on each hardware component in the current system relative to AME Floor-II. These time savings and hence increase in frame rate are obtained by doing a more refined timing analysis on each hardware component to determine their operational delay and accordingly generating the timing and control signals from the microcontroller. Section 4.2.1 enumerates the technique used to increase the frame rate from 33 Hz to 43 Hz.



$$T_{SCAN(II)} = T_{MUX(II)} + T_{ADC(II)} + T_{RABBIT(II)} = 30.0305 \text{ ms } (F_{(II)} = 1/T_{SCAN(II)} = 33 \text{ Hz})$$
$$T_{SCAN(III)} = T_{MUX(III)} + T_{ADC(III)} + T_{RABBIT(III)} = 23.255 \text{ ms } (F_{(III)} = 1/T_{SCAN(III)} = 43 \text{ Hz})$$

Fig. 8. Block diagram of Floor hardware annotated with hardware component delays

4.2.1 Frame rate increase technique

It is apparent from Fig 8 that the time savings obtained in the A/D converter, multiplexer and rabbit controller lead to an overall increase in frame rate from 33 Hz to 43 Hz. The operational delay of the A/D converter and rabbit controllers are determined by trial and error procedure. Repeated iterations are done with different delay values (for A/D and rabbit) in the microcontroller routine and the least delay for correct operation is then determined. Major time saving is obtained in the multiplexers by non-uniform multiplexing technique. It is important to note that the time taken for each sensor to be scanned and pressure value to be read is not uniform for all 2016 sensors on the pressure sensing mat. The reason behind the above statement can be explained with the aid of Fig 9. The floor control hardware includes three 16 x 1 row multiplexers and three 16 x 1 column multiplexers. Each input line of the multiplexer is wired to single sensor output. The microcontroller generates the multiplexer enable signal to enable a particular row and column multiplexer. Soon afterwards, the multiplexer select signals are also sent out by the microcontroller to read a particular input line. Additional instructions are required in the microcontroller scan routine when there is a switch from one multiplexer to another. The sensors wired to the first input line of multiplexer accounts for such a switch thereby taking more time to complete. For example, sensor 17 on the mat requires a switch from row mux-1 to row mux-2 which is achieved by additional instructions and hence longer completion time. Sensor 1345 takes even longer time to complete because it requires two switches namely column mux-2 to column mux-3 and row mux-3 to row mux-1.

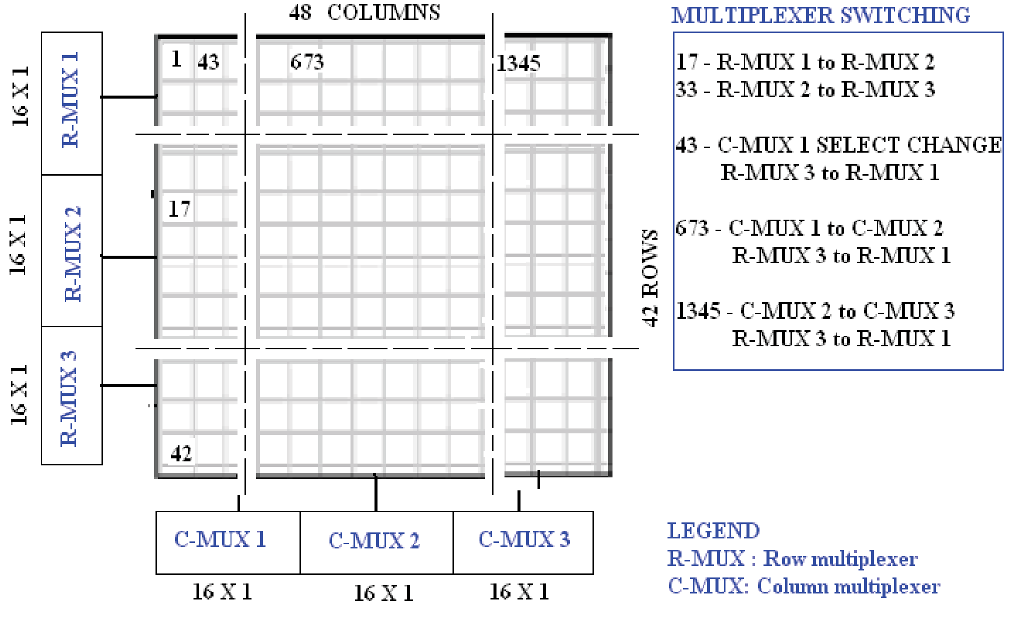


Fig. 9. The arrangement of multiplexers for a pressure sensing mat and sensors that require longer scanning time due to multiplexer switching.

AME Floor-II (Srinivasan, P., 2006) resorted to a uniform multiplexing technique whereby the time taken to scan each sensor was made uniform throughout the pressure sensing mat. The worst case or longest sensor scanning delay was determined and all sensors were scanned uniformly with that delay value. This was achieved by incorporating additional delays even for sensors that could have finished scanning in lesser time. In AME Floor-III all these extra delays were removed and the multiplexing is made non-uniform. Currently, each sensor is scanned at the fastest rate possible which eventually leads to significant savings in multiplexer to finish operation and thereby increase in frame rate.

4.3 Data Compression

Each pressure mat has 2016 sensors and each sensor in turn sends one byte of pressure data at 43 Hz. Thus each mat data packet size adds to 2017 bytes which includes 2016 bytes of pressure data and one byte of frame number. The data volume from the entire floor comprising of 96 mats running at 43 Hz is a whopping 8.4 MB/sec. Usually, except a small area where the subject is in contact with the pressure sensing floor, most of the sensors do not have any load acting on them. Consequently a large proportion of the sensor data are null values of pressure or noise serving no interest to applications. Also there has been slight random noise observed in few sensors because of the nature of the sensing material which reports small values of pressure. Hence a simple but elegant compression algorithm is implemented by the floor control software to filter out all pressure values below the chosen noise threshold and pack only “active” sensor values and their addresses (location on floor system matrix) to be sent out to the end users on the network. Compression ratio as high as

0.9 is observed under normal case with five subjects which proves significant data volume reduction on the network.

It is known that compressed data packet comprise of only active sensor values and their address whereas the uncompressed data packet comprise of all sensor values (arranged in a sequence) and no address information since its address is implied by its location in the data packet. Thus the compression algorithm adds an additional overhead of sensor address which works well for low user activity with less active sensors. However as the user activity on the floor increase or when large numbers of sensors are active, the packet size also grows and a point is reached when compressed data volume exceeds uncompressed data volume. It is determined that this breakeven point is generally high and beyond bounds for normal usage. Hence the algorithm works well for most of the situations.

4.4 Mean Shift Tracking of Pressure Clusters

Context awareness is the vital part of any smart environment. Perceiving context means sensing the state of the environment and users and it can be done with regard to a person or an activity. This may involve a variety of tasks such as person recognition, person location tracking, activity detection, activity recognition, activity learning etc. The primary step to accomplish the above tasks is to develop an efficient tracking procedure that shall ascertain the person location on the floor and also shift in the pressure gradient. The latter may lead to the study of various pressure patterns tied to each and every user activity. A mean shift algorithm is used to achieve the above mentioned goal. Mean shift is a simple iterative procedure that shifts each pressure data point to the average of the pressure data points in the neighbourhood.

4.4.1 Mean shift: an introduction

Mean shift is the process of repetitively shifting the centre t to the sample mean. The sample mean of samples S under a kernel $K(x)$ centred at t, with sample weights $w(s)$, can be found using this equation:

$$m(t) \equiv \frac{\sum_{s \in S} K(s-t)w(s)s}{\sum_{s \in S} K(s-t)w(s)} \tag{11}$$

where $m(t)$ is the new sample mean (Cheng., 1995). It's proven (Comaniciu et al., 2000) that if the kernel $K(x)$ has a convex and monotonically decreasing profile $k(||x||^2)$, then the centre t will converge onto a single point. The kernel used in our tracking algorithm is the truncated Gaussian kernel which is the combination of the flat kernel and Gaussian kernel. The truncated Gaussian kernel is given by

$$(F_\lambda \; G_\beta)(x) = \begin{cases} e^{-\beta \|x\|^2} & if \; \|x\| \leq \lambda \\ 0 & if \; \|x\| > \lambda \end{cases} \tag{12}$$

where λ is the radius of the Gaussian kernel and β is the Gaussian kernel coefficient.

4.4.2 Clustering and tracking algorithm

The algorithm is iterated for every frame of pressure data. Each and every frame of pressure data contains information about the location of pressure and value of pressure at that location. The pressure values constitute the weights and pressure location constitutes the data points that need to be iterated using the mean shift algorithm. The full algorithm for finding and tracking the pressure clusters is given below.

1) For the first frame of pressure data or new cluster formation , cluster centres and the data points are one and the same i.e. the centre set T is the same as the data set S, and both evolve with each iteration using the mean shift formula in equation 11 and truncated Gaussian in equation 12. Data points are clustered through the blurring process (Cheng, 1995) using the observed pressure data as the weight used in eqn. 11. Once the process has converged, the data set will be tightly packed into clusters, with all of the data points located closely to the centre of that cluster. (The process is said to the converged either after the maximum number of iterations defined by the algorithm or earlier when the mean shift of centres becomes less than the convergence threshold) After convergence, each cluster has a 'centre' and 'label' associated with it. All data points not associated with any cluster centre are classified as orphan pressure points.

2) For every subsequent pressure data frame, centres from the previous frame are updated through the mean shift algorithm (eqn. 11) using current observed pressure values as weights and checked for convergence. In practice, entirely new data points resulting in new cluster centres (new labels) can occur which is computed in step (3).

3) Calculate the number of orphan pressure points. If the number of orphan pressure points exceeds a chosen threshold then repeat step (1) to find new cluster centres. Orphan pressure points fewer than the chosen threshold are discarded.

4) Perform mean shift using the new set of cluster centres (repeat steps 2 & 3).

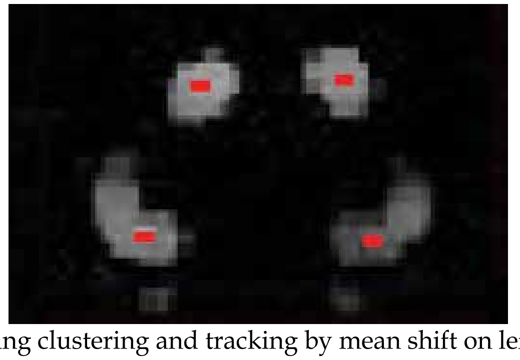


Fig. 10. Snapshot showing clustering and tracking by mean shift on left foot and right foot. Two pressure clusters are formed for each foot (one for heel and one for toe) and cluster centres are depicted by red dots.

5. System Integration for Multimodal Sensing

This section presents the system integration of AME Floor-III and motion capture system to create a multimodal environment for holistic movement sensing. Multimodal systems have always proved to be robust and effective than unimodal systems because it provides wide varieties of information for better realization of the subject movement in capture space. In a multimodal system, users have the flexibility to interact with the environment through multiple communication channels e.g. gesture, voice and pressure distribution paving way for increased expressive power and user friendliness. Multimodal systems provide high redundancy of content information and hence high reliability. Also the weakness of one modality is offset by the strength of the other. In this manner, multiple sensing modalities possessing symbiotic relationships are found to be very effective for human computer interaction. Hence after the completion of pressure sensing floor, efforts have been put in to integrate it with the motion capture system to create a smart multimodal environment.

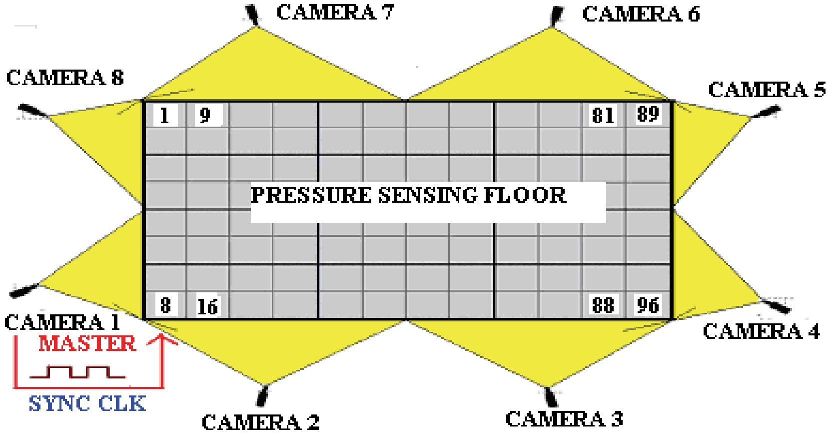


Fig. 11. Common capture volume of the floor and motion capture system.

A common capture volume (12′ by 15′) is first created within the sensing capabilities of the floor and motion capture system. The motion capture cameras are arranged around a capture volume and the floor forms a part of the capture volume as shown in Fig. 11. The location of the floor with respect to the coverage area of the cameras is important when pressure data about some movement needs to be interpreted with the marker. The pressure floor and motion capture system are integrated with respect to time and spatial domains. A subject moving in the capture space is sensed by both systems and they give information about the location and activity of the subject. Motion capture data contains the 3D location coordinates of the markers in physical space whereas the pressure data contains the pressure values and 2D location. Both sensing systems have independent coordinate set and hence spatial alignment by means of coordinate transformation becomes essential to ascertain the location of the subject in common capture space. Also any activity done by the subject is being detected by both systems simultaneously and hence both sensing modalities must operate synchronously. Thus time synchronization and spatial alignment are critical for two data sets to be highly correlated for holistic inference.

5.1 Temporal Alignment

Temporal alignment is defined as the process of synchronization of both sensing modalities so that both systems record an event in the common capture volume at the same time instant. Perfect temporal alignment leads to a holistic inference on the time of occurrence of the event. Temporal Alignment of the floor and motion capture system is achieved by means of a common sync clock. This sync clock is generated by the master camera of the motion capture system and is used to trigger the scan of the floor. This sync clock is used by the motion capture system to control the camera shutters. The clock signal feeds as an external signal to the micro-controller (in the local mast based controller) to initiate floor sensor scan. In this way, the scan of the floor and the camera image capture are synchronized in time domain if both are operating at the same frame rate or frame rate multiples of one another. The maximum achievable frame rate of the motion capture system and AME Floor-III are not equal. The motion capture system is capable of running at higher frequencies than the floor. Running the motion capture system at the same frequency of the floor results in a situation where the full sensitivity of the motion capture system is not utilized. Thus for temporal alignment, motion capture system is always set to run at multiples of the floor frequency. The common sync clock runs at the frequency of the motion capture system and that clock is down sampled by a factor to generate the scan frequency (or frame rate) of the floor. Currently the motion capture system is set to run at 120 Hz and the floor at 40 Hz. The frequency set for the floor system should be less than the maximum achievable floor frame rate i.e. 43 Hz. This arrangement generates 3 motion capture data frames for every single pressure date frame. So the motion capture data frames are down-sampled (redundant frames are ignored) to create an equal number of floor and motion capture data frames for comparison purposes. All data frames are referenced by means of frame numbers to track the same event detected by both systems. The time of occurrence of the event (relative to the start of data capture) can be computed from the frame number of the data pertaining to the event and frame rate of the sensing modality.

A frame alignment experiment is conducted to verify the temporal alignment of the AME Floor-III and motion capture system. The motion capture system is set to run at 120 Hz and AME Floor-III at 40 Hz. A predefined start up procedure is resorted to ensure the start of both sensing modalities at the same time instant. A mallet with a single marker on its head is banged on a single pressure sensor of the floor from a fixed height. The motion capture system tracks the movement of the marker on the mallet whereas the pressure sensing system monitors the pressure value on the single pressure sensor. The vertical coordinate (Y- coordinate) of the marker given by the motion capture system and pressure value on that sensor given by the pressure sensing floor are monitored over time. Ideally the pressure value on that sensor should peak when the marker coordinate is at the lowest position (ground level). In other words, the pressure peak should occur at the same time instant when the vertical height of the marker is at its lowest value. Since the motion capture system is running at three times the frequency of AME Floor-III, motion capture frames are down-sampled to create an equal number of floor frames for comparison and check of frame alignment between the two data sets. Fig 12 gives the time-sampled plot of the pressure sensor values (green dots) and marker vertical height (pink dots) from data captured during the experiment. It can be seen that the ‘first’ pressure peak detected by the pressure floor and the ‘first’ lowest marker height detected by the motion capture system occurs at frame number 46 (after down-sampling of motion capture data frames). It is clear that the results obtained agree with our expectation thereby demonstrating a perfect temporal alignment between the floor and motion capture system.

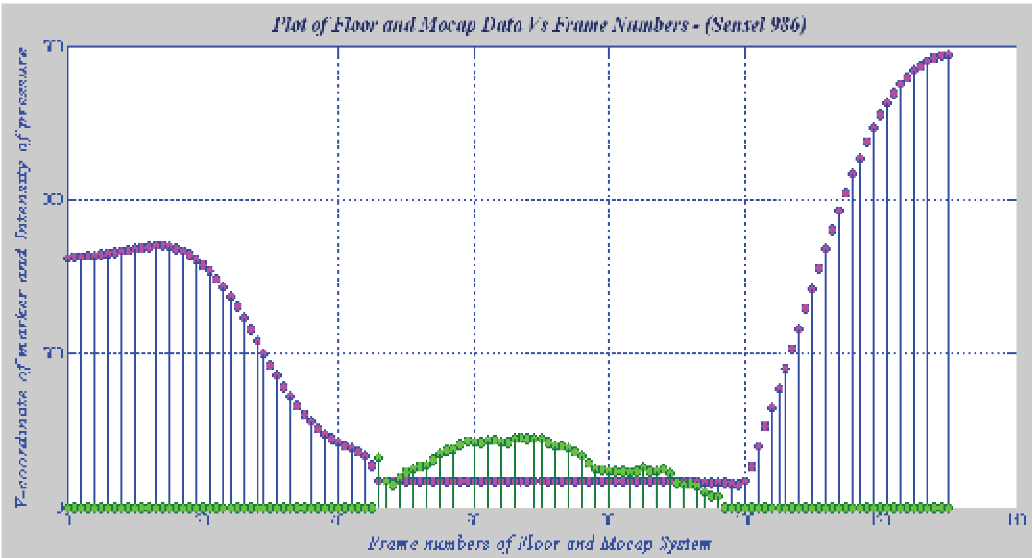


Fig. 12. Plot of pressure and marker data v.s. floor frame numbers showing perfect temporal alignment.

5.2 Spatial Alignment

Spatial alignment (in this context) is defined as the process of determining the transformation parameters for conversion of a spatial coordinate in one coordinate system to an equivalent coordinate in another coordinate system for a holistic inference on the location of the subject. The floor coordinate system is a two dimensional system in sensor units whereas the motion capture coordinate system is a three dimensional system in mm units. Hence it is essential to implement coordinate transformation between the floor and motion capture system so that we can view the events in one coordinate space for ease of inference and visualization

A spatial calibration procedure is in order to align the floor and motion capture system in physical space. Firstly the motion capture system is calibrated and stabilized. Three reflective markers are placed on the floor as shown in Figure 13. Origin marker is placed on the first sensor of mat 19, x-axis marker on the first sensor of mat 23 and z-axis marker on the first sensor of mat 75. Three points inside the floor are chosen so that they are well within the coverage areas of the cameras. The positional coordinates of these three markers are then gathered which in turn denotes the position of the floor in the motion capture coordinate space. Using this information, three co-ordinate transformation parameters namely rotation, translation and scaling are computed. These parameters constitute the coordinate transformation matrix which is then applied to each and every floor coordinate to get the respective coordinate in the motion capture system. The converse also can be computed to view the data in the floor coordinate space alone. Spatial alignment computations are done by floor control software in real time. The theory and math behind the calculation of scaling, translation and rotation parameters is explained in the following sections.

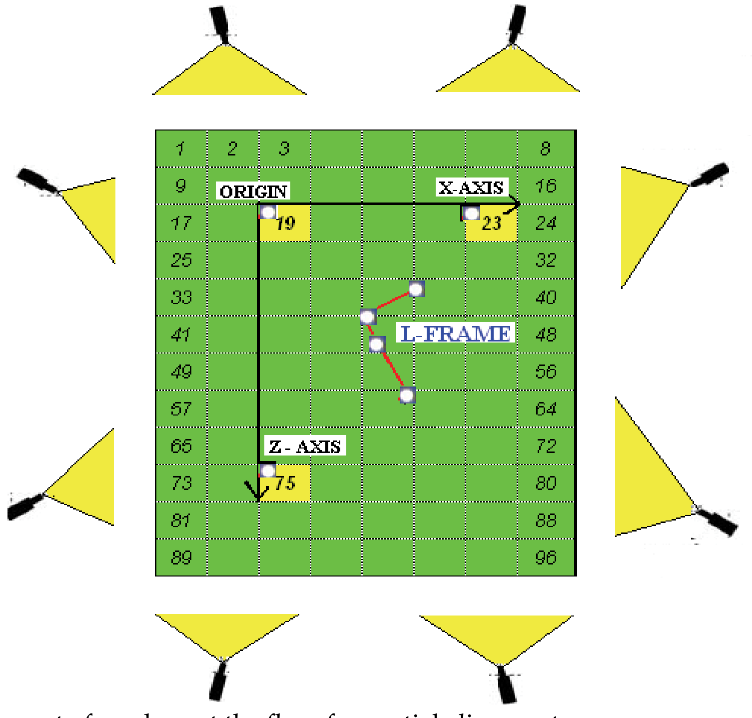


Fig. 13. Placement of markers at the floor for spatial alignment

Let the positional coordinates of the origin, x-axis and z-axis markers obtained from the motion capture system be represented as $(x_1, 0, z_1)$, $(x_2, 0, z_2)$, $(x_3, 0, z_3)$ respectively. The y-coordinate is always at zero because the floor is parallel with the X-Z plane and intercepts at the zero point on the y-axis of the motion capture coordinate space. Using the above positional information the scaling, rotation and translation parameters are computed as follows.

5.2.1 Scaling

Scaling parameter in X-direction (S_X) is computed from the positional coordinates of origin and x-axis markers. Scaling factor (S_X) is obtained when the distance between origin point and x-axis point is divided by the number of sensors in between the two points. It is expressed in mm/sensel.

$$S_X = \frac{\sqrt{(x_2 - x_1)^2 + (z_2 - z_1)^2}}{48 \times 4} \tag{13}$$

Scaling parameter in Z-direction (S_Z) can be derived similarly.

$$S_Z = \frac{\sqrt{(x_3 - x_1)^2 + (z_3 - z_1)^2}}{42 \times 7} \tag{14}$$

It is worth mentioning here that ideally $S_X = S_Z$ since the sensors are uniformly distributed over the entire area.

5.2.2 Rotation

The angle of rotation from the floor coordinate system to the motion capture coordinate system is computed as follows. Vector OX is computed from the positional coordinates of origin and x-axis markers.

$$\overrightarrow{OX} = (x_2 - x_1, z_2 - z_1) \tag{15}$$

The rotation angle θ is given by atan2 (z_2-z_1, x_2-x_1), which is the counter-clockwise angle in radians between the x-axis of motion capture coordinate system and the vector OX (x-axis) of the floor coordinate system.

5.2.3 Translation

Since the origin marker is placed two mats from the top and two mats from the left of the actual floor boundary, the translation T_{SX} = *-96* (since 2 mats x 48 columns/mat = 96 columns) and T_{SZ} = *-84* (since 2 mats x 42 rows/mat = 84 rows). Thus the translation in 'mm' units is obtained by multiplying with their respective scaling factors.

$$T_{MX} = -96 \times S_X \qquad T_{MZ} = -84 \times S_Z \tag{16}$$

The actual translation parameters (T_X, T_Z) are then calculated by rotating the above parameters by an angle θ and adding to the origin vector.

$$(T_X, T_Z) = (x_1, z_1) + (T_{MX}, T_{MZ}) \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{17}$$

5.2.4 Coordinate transformation equations

The coordinate transformation parameters namely scaling (S_X, S_Z), rotation (θ) and translation (T_X, T_Z) are computed from the above equations. Now given the actual floor coordinate (X_F, Z_F) of a point and the coordinate transformation parameters, (X_M, Z_M) the coordinate of the point in the motion capture coordinate space is given by

$$\begin{bmatrix} X_M \\ Z_M \\ 1 \end{bmatrix} = \begin{bmatrix} S_X \cos\theta & -S_Y \sin\theta & T_X \\ S_X \sin\theta & S_Y \cos\theta & T_Z \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_F \\ Z_F \\ 1 \end{bmatrix} \tag{18}$$

After the implementation of spatial alignment, experimental data are collected and shown in Table 3. Pressure is applied on the sensors in the neighbourhood of the origin of the motion capture system including the sensor situated at the origin. Sensor 0 of mat 36 corresponds to the origin of the motion capture coordinate system. Transformed coordinates of this sensor gives a value (1.94395, 0, -3.53732) which reveals a good accuracy of spatial alignment. Each sensor has a total area of 10 mm x 10 mm thereby explaining the reason for this offset.

FRAME #	X_F	Z_F	VALUE	MAT #	SENSOR INDEX	X_M	Y_M	Z_M
38939	192	167	29	28	41	1.01874	0	-13.6887
38939	192	168	27	36	0	1.94395	0	-3.53732
38939	192	169	22	36	1	2.86916	0	6.61411

Table 3. Tabulation of collected pressure data in an around the origin of the motion capture coordinate space. Transformed floor coordinates in the motion capture coordinate space are also shown in last three columns.

6. Applications in Multimodal Movement Sensing and Analysis

6.1 Balance Analysis

Falling is one of the major health concerns for elderly people and incidence of falls is high for persons aged over 75. Hence an efficient fall detection system is necessary to detect potential situations of fall and signal the user of an impending fall or alert for assistance after the person is immobilized by fall. The state of body balance is the feature of interest in fall detection systems. The state of body balance is characterized by centre of gravity (COG) and centre of pressure (COP). COG is computed from the motion capture data by assigning weight to each marker and computing the weighted mean. If the weight of each marker represents the weight of the body mass around that marker, the weighted mean is a good approximation of the centre of gravity. Similarly the COP is the weighted mean of all the pressure data points. The subject's overall state of balance is determined by the relative positions of the COG and COP. If the COG is directly above the COP, the subject is in a state of balance.As COP and COG moves away from each other, the subject slowly transitions into a state of off-balance. Thus it is obvious that time synchronization and spatial alignment of both sensing systems are critical for such an exercise. Since feelings of balance are visceral in human beings, such a quantitative approach paves way to tie the behavior of the system to a sensation/feeling that is very internal and apparent to the user and thereby complementing human computer interaction.

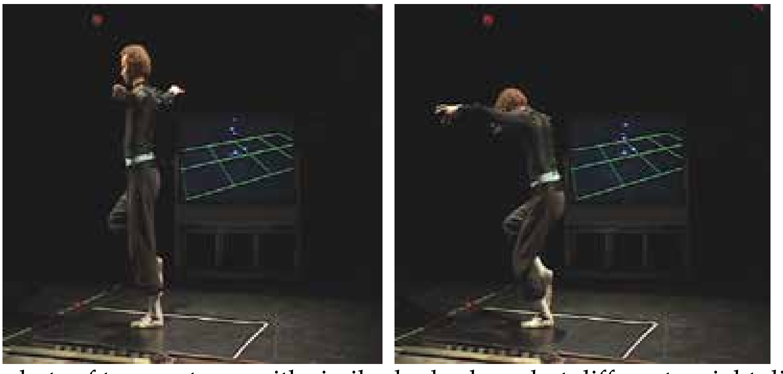


Fig. 14. Snapshots of two gestures with similar body shape but different weight distribution

6.2 Gesture Recognition

This multimodal sensing system has also been used to drive a gesture recognition system that uses both kinematics and pressure distribution to recognize gestures. Such a gesture recognition system can distinguish gestures that have similar body shapes but different weight distributions as shown in Figure 14. These two gestures are recognized as one and the same by marker based motion capture system due to similar body shape. Hence pressure sensing becomes vital to distinguish between such gestures. The ability of the gesture recognition system to read and analyze both body kinematics and pressure distributions encourages users to communicate with computers in expressive ways.

7. Conclusions and Future Work

We have successfully designed, developed and deployed a pressure sensing floor system with a higher frame rate, less latency, high sensor resolution, large sensing area that can provide us with real time data about the location and amount of pressure exerted on the floor. The floor has been integrated and synchronized with the marker based motion capture system to create a smart environment for movement based human computer interaction. Future direction shall be towards extending the context aware capabilities of the floor system. An algorithm that can clearly distinguish between the left foot and the right foot shall find extensive usage of the floor to numerous applications. Shape descriptors such as Fourier and Hu moments to distinguish left foot and right foot on the basis of shape come in handy for such an analysis. Such intelligence to the floor to recognize and distinguish the left and right foot shall pave way for recognizing gestures with varying foot contact. The above work may be further extended to make a distinction of the heel and toe of a particular foot as well. This shall find extensive use in diagnosis of various gait pathologies as most disorders are reflected by abnormal pressure patterns localized to either the toe or heel. AME Floor-III as it stands now is not portable. Further work is also being done in the design of the interfacing hardware to make a portable system. The creation of a wireless pressure sensing system is a possible alternative towards a portable system. Integration of other sensing modalities such as audio-based sensing (microphone arrays), a wireless EMG system into the existing multimodal framework to create a very powerful tool for movement based human computer interaction is another major challenge in the future. The establishment of statistical models and machine learning techniques to model the underlying relationships of human movement information sensed by the system are also being investigated.

8. Acknowledgement

The authors thank Mr. David Lorig of Arts, Media and Engineering Program, Arizona State University for building the floor supporting frame and installing the floor. This chapter is based upon work partly supported by U.S. National Science Foundation on CISE-RI no. 0403428 and IGERT no. 0504647. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation (NSF).

9. References

Addlesee, M.; Jones, A., Livesey, F. & Samaria, F, (1997). The ORL active floor. *IEEE Personal Communications*, 97, 35-41.

Birchfield, D.; Ciufo, T., Minyard, G., Qian, G., Savenye, W., Sundaram, H., Thornburg, H. & Todd, C. (2006). SMALLab: a mediated platform for education. *Proceedings of the 33rd International Conference and Exhibition on Computer Graphics and Interactive Techniques* in conjunction with SIGGRAPH, Boston, MA, 2006.

Chen, Y.; Huang, H., Xu, W., Wallis, R. I. , Sundaram, H., Rikakis, T., He, J., Ingalls, T. & Olson, L. The design of a real-time, multimodal feedback system for stroke patient rehabilitation. (2006). *Proceedings of ACM Multimedia*

Cheng, Y. (1995). Mean shift, mode seeking and clustering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 8, pp.790-799.

Comaniciu, D.; Ramesh, V. & Meer, P. (2000). Real-time tracking of non-rigid objects using mean shift, *IEEE Conf. Computer Vision and Pattern Recognition*, pp.142-149.

Griffith, N. & Fernström, M. (1998). LiteFoot:A floor space for recording dance and controlling media.*Proceedings of International Computer Music Conference*, pp.475-481.

Jaimes, A. & Sebe, N. (2007). Multimodal Human-Computer Interaction: A Survey, *Computer Vision and Image Understanding*, Vol. 108, No. 1-2, pp. 116-134.

Kidané, A.; Rodriguez, A., Cifdaloz, O. & Harikrishnan, V. (2004). ISAfloor: A high resolution floor sensor with 3D visualization and multimedia interface capability. Technical Report, AME Program, Arizona State University.

Middleton, L., Buss, A. A., Bazin, A. I. and Nixon, M. S. (2005) A floor sensor system for gait recognition. *Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pp 171-176., Buffalo, New York, USA, 17-18 October, 2005

Morishita, H.; Fukui, R. & Sato, T. (2002) .High resolution pressure sensor distributed floor for future human-robot symbiosis environments. *International Conference on Intelligent Robots and Systems*, pp. 7803-7398.

Paradiso, J.; Hsiao, K., Benbasat, A. & Teegarden, Z. (2000). Design and implementation of Expressive Footwear, *IBM Systems Journal*, Vol. 39, No. 3 - 4, pp. 511-52.

Paradiso, J.; Abler, C., Hsiao, K. & Reynolds, M. (1997). The Magic Carpet: Physical sensing for immersive environments. *ACM Press*, *Ext. Abstracts CHI*, pp. 277- 278.

Qian, G.; Guo, F., Ingalls, T., Olson, L., James, J. & Rikakis, T. (2004). A gesture-driven multimodal interactive dance system. *Proceedings of IEEE International Conference on Multimedia and Expo*, Taibei, Taiwan, China, 2004

Richardson, B.; Leydon, K., Fernström, M. & Paradiso, J. (2004). Z-Tiles: building blocks for modular, pressure-sensing floorspaces. *Extended Abstracts of the conference on Human factors and computing systems*, pp. 1529-1532.

Rangarajan, S., ; Kidané, A., Qian, G., Rajko, S. & Birchfield D. (2007a). Design of a pressure sensing floor for movement based human computer interaction. *Proceedings of European Conference on Smart Sensing and Context (EuroSSC)*, pp. 46-61. Lake District, UK, LNCS 4793, Springer-Verlag Berlin Heidelberg.

Rangarajan, S. (2007b). Design optimization and multimodal integration of a large area pressure sensing floor, Master's Thesis, Arizona State University

Srinivasan, P., (2006). Design of a Large area pressure sensing floor – A thesis presented for the requirements Master of Science Degree, Arizona State University, 06.

Srinivasan, P.; Birchfield, D., Qian, G. & Kidané, A., (2005). A Pressure Sensing Floor for Interactive Media Applications. *Proceedings of ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, pp. 278-281

Improving Target Acquisitions through Utilizing Pen Pressure

Xiangshi Ren[1], Jibin Yin[1], Shengdong Zhao[2], & Yang Li[3]
[1] *Kochi University of Technology, Kochi 782-8502*
Japan
ren.xiangshi@kochi-tech.ac.jp
[2] *University of Toronto, Toronto, Ontario*
Canada
sszhao@dgp.toronto.edu
[3] *University of Washington, Seattle, WA 98195-2350*
USA
yangli@cs.washington.edu

1. Introduction

Target selection via pointing is a fundamental task in graphical user interfaces (GUIs). A large corpus of work has been proposed to improve mouse-based pointing performance by manipulating control display (CD) parameters (Blanch et al., 2004; Grossman & Balakrishnan, 2005; Guiard et al., 2004; Kabbash & Buxton, 1995; Worden et al., 1997) in desktop environments.

Compared with mouse-based desktop GUIs, pen-based interfaces have a number of different characteristics. First, pen-based interfaces typically use absolute pointing via a direct input device (i.e., a pen), which is very different from indirect input, such as using a mouse. Second, in addition to the 2D position (x, y) values, many pen-based devices offer additional sensory properties (such as pen pressure values) that can be useful for interaction. Third, many pen-based interfaces have limited display space and input footprint. As the amount of information displayed on the screen increases, users have to select smaller targets. This is especially obvious in mobile products, such as personal digital assistants (PDAs), pen-based mobile phones, and other mobile pen-based applications. Compared with the extensive studies carried out for mouse-based pointing, more empirical studies are needed to determine how we can improve pen-input usage and efficiency.

Although previous studies have intended to exploit novel pen-based selection techniques, such as Slide Touch (Ren & Moriya, 2000), Drag-and-pop (Baudisch et al., 2003), Bubble Radar (Aliakseyeu et al., 2006) and Beam Cursor (Yin & Ren, 2006), these techniques were mostly designed for situations where targets are sparsely distributed across a display space. When targets are smaller and densely packed, the benefit of these techniques tends to be diminished or become unavailable.

Recently, an increasing amount of work has explored the use of pen pressure, which is available on pen devices (such as most Tablet PCs or Wacom tablets), as the third input dimension for interaction design (Herot & Weinzapfel, 1978; Li et al., 2005; Ramos et al., 2004; Ramos et al., 2003; Ramos & Balakrishnan, 2005), in addition to the 2D x-y coordinates. However, little attention has been paid to using pen pressure to improve target selection tasks. This study, therefore, investigates the possibility of improving the performance of target acquisition tasks for pen-based environments by taking advantage of pen pressure potentials. This chapter presents the Adaptive Hybrid Cursor, the first interaction technique that employs pen pressure for target selection. The Adaptive Hybrid Cursor can automatically adapt the selection cursor as well as the target space based on pen-pressure.

There are three fundamental elements in a selection task: a cursor, a target, and a selection background (including a void space). We explored how pen pressure can be employed to improve target acquisition tasks by varying these three elements. The background plays an important role in many applications but its use was often overlooked in previous work. For example, numerous functionalities have been designed to associate with the background in Windows and Mac desktops, from basic but important functions such as selecting and deselecting, to re-arranging desktop icons and also to more complex operations such as changing certain properties of applications. A background also serves as a visual storage space for future elements. Furthermore, group selection techniques (such as rectangular or lasso techniques) would be awkward to operate without being able to select an empty space. The famous quote from the ancient Chinese philosopher, Lao Tze, says, "the usefulness of the wheel, cup and house is actually based on their emptiness". Without the ability to select the background, many applications become difficult to use.

The Adaptive Hybrid Cursor has the following design characteristics:

(1) This technique takes advantages of pressure-sensitive input devices. Pressure is used to control the zoom ratio of interface contents. To achieve a steady zoom control by pressure, an optimal pressure mapping function is employed.

(2) This technique improves performance by manipulating all three components of target selection: the background, the target and the cursor. Such technique design allows quick and accurate small target selections, even for targets that are arranged tightly.

(3) This technique employs an adaptive strategy for target selections, in which two selection mechanisms are coupled: (i) Zooming *Cursor* method and (ii) Zooming *Target, Cursor* and *Background*. With the adaptive strategy, the best mechanism is invoked according to information on the size and layout density of a desired target.

(4) This technique provides easy cancellation by reversing the pressure value without having to use an extra mode-switch button.

In evaluations of this technique, the two selection mechanisms of this technique are thoroughly examined by formal experiments. Subjects performed 2-dimension selection tasks with different densities and sizes of targets. The researchers found that the technique indicated benefits in selecting small targets with high densities. This technique could be implemented on devices capable of sensing pressure like tablet computers or some other pen-based devices.

In this chapter, we first review the related work. Next we describe the design of our new technique. We then present the evaluation of the Adaptive Hybrid Cursor under various target acquisition conditions. We conclude with a discussion of our results and directions for future work.

2. Related Work

In this section, we discuss related work regarding both target selection techniques and pen pressure.

2.1 Previous Work on Selection Techniques

Target selection tasks can be modelled by Fitts' law (Fitts, 1954; MacKenzie & Buxton, 1992). One common form of Fitts' law is $MT=a+b\log_2(A/W+1)$, which states that the time (*MT*) to acquire a target with width *W* and distance (or amplitude) *A* from the cursor can be predicted (where *a* and *b* are empirically determined constants, and the term inside the log function is called *Index of Difficulty* or *ID*). Obviously, target acquisition performance can be improved by increasing *W*, decreasing *A*, or both.

The width of a target is usually defined by the space it occupies on the screen. The *effective target width (EW)* may be defined as the analogous size of a target in motor space. In standard pointing, the effective target width matches the visual width. However, the effective width can be increased either for the cursor (Grossman & Balakrishnan, 2005; Kabbash & Buxton, 1995; Worden et al., 1997) or the target (Cockburn & Brock, 2006; McGuffin & Balakrishnan, 2002; Zhai et al., 2003) to achieve the same effect. Most previous studies have shown the effectiveness of their proposal only for single isolated target (McGuffin & Balakrishnan, 2002; Zhai et al., 2003), while they have not been shown to work well when multiple targets are present in close proximity (Cockburn & Brock, 2006; Guiard et al. 2004; McGuffin & Balakrishnan, 2002; Zhai et al., 2003). The state of the art in this category is Bubble Cursor (Grossman & Balakrishnan, 2005), a mouse-based technique that allows selection of discrete targets by using a Voronoi diagram to associate void space with nearby targets. Bubble Cursor works well even in a normal-density multiple-target environment except for the limitations mentioned in the discussion section of this paper.

There is also a large body of work that is intended to improve selection performance by decreasing A. They either bring the target much closer to the cursor such as Drag-and-pop developed by Baudisch et al. (2003), and 'vacuum filtering' introduced by Bezerianos & Balakrishnan (2005), or jump the cursor directly to the target, such as with the object pointing technique (Guiard et al. 2004). Overall, the performance of techniques aiming to decrease A is largely affected by the number of distracting targets between the starting position and the target. They tend to work well on large displays where targets are further away or in low density environments with few distracting targets. These techniques become less effective with high or normal density environments in regular or smaller size displays such as Tablet PCs or PDAs.

Some have tried to improve pointing and selection by dynamically adjusting the Control Display gain. The gain is increased on the approach to the target and decreased while inside the target thus increasing and decreasing the motor space at critical moments in the selection process. TractorBeam (Parker et al., 2005) is a hybrid point-touch technique that aids selection by expanding the cursor or the target, or by snapping to the target. Worden et al. (1997) implemented 'Sticky Icons' by decreasing the mouse control-display gain when the cursor enters the icon. Blanch et al. (2004) showed that performance could be predicted using Fitts' law, based on the resulting larger *W* and smaller *A* in the motor space. The common problems for these techniques occur when multiple small targets are presented in close proximity, as the intervening targets will slow the cursor down as it travels to its destination target.

An interesting special case here is a technique which is used on large displays to help reach targets that are beyond the arm's reach (Aliakseyeu et al., 2006; Baudisch et al., 2003; Bezerianos & Balakrishnan, 2005; Collomb et al., 2005 ; Nacenta et al., 2005), e.g., RadarView (Nacenta et al., 2005). However, since RadarView decreases both *A* and *W* proportionally, the *ID* is unchanged. The benefit of RadarView is only demonstrated on larger displays where users can operate on RadarView to save the extra movement required to reach a distant target i.e. one that is beyond arm's reach. Bubble Radar (Aliakseyeu et al., 2006) combines RadarView and Bubble Cursor by first placing the objects within reach, and then applying Bubble Cursor to increase selection performance. Bubble Radar also tried to address the background selection problem of Bubble Cursor by using a button switch controlled by the non-dominant hand, however, since Bubble Radar is virtually another Bubble Cursor, its advantage is likely to diminish in a high density environment.

2.2 Related Work on Pressure

There has been less work done on pressure than on pointing-based target acquisition characteristics. Studies on pressure can be roughly divided into two categories. One category investigates the general capabilities of humans interacting with computers using pressure. For example, Herot & Weinzapfel (1978) investigated the human ability of the finger to apply pressure and torque to a computer screen. Buxton (1990) studied the use of touch-sensitive technologies and the possibilities for interaction they suggest. Ramos et al. (2004) explored the human ability to vary pen-tip pressure as an additional channel of control information. The other category of study is where researchers build pressure enabled applications or techniques. For instance, Ramos & Balakrishnan (2003) demonstrated a system called LEAN and a set of novel interaction techniques for the fluid navigation, segmentation and annotation of digital video. Ramos & Balakrishnan (2005) designed Zlider widget. Li et al. (2005) investigated using pressure as a possible means to delimitate the input phases in pen-based interactions. Although these works opened the door to establish pressure as a research avenue, we are unaware of any work which addressed the issue of applying pressure into discrete target acquisition. We attempt to investigate this issue in this paper.

3. Adaptive Hybrid Cursor Design

A few previous studies have shown that a reasonable manipulation of targets, cursors and context can enhance target acquisition. However, the tradeoff between the "original" state of these three elements and the "manipulation" state needs to be considered in technical design. Our approach is to employ pen-pressure which is an available parameter in some pen based devices and can be used to easily produce a continuous value or a discrete state. Pen-pressure has the potential to affect selection implementation. Based on this idea we designed the Adaptive Hybrid Cursor technique.

Adaptive Hybrid Cursor includes two states. It first determines whether it should zoom its contexts (target and background) and/or cursor according to the initial location of the cursor and the information regarding the position of targets. If the condition is not suited to the adaptive strategy, Adaptive Hybrid Cursor initiates the Zoom Cursor technique described in Section 3.1 (see Fig. 1). If the condition satisfies the adaptive strategy criteria, Adaptive Hybrid Cursor begins to zoom the targets, the cursor and background based on the pressure described in Section 3.2 (see Fig. 2).

3.1 Zoom Cursor Technique (State 1)

One possibly fruitful direction open to the examination of pressure-enhanced target acquisition is to use pen pressure to enlarge the cursor size. Based on this intuition, we designed Zoom Cursor, a technique that allows a user to enlarge the cursor size by pressing the pen tip harder on a tablet or a touch-sensitive screen (see Fig. 1).
As determined in previous studies (Barrett et al., 1996), the degree of pen pressure perceived by human users is not consistent with that sensed by digital instruments. For example, at a low spectrum of pen pressure, the sensed pressure value increases much faster than users would expect. Previous work has used a sigmoid transfer function to achieve the effects produced by pressure. In our experiments we also employed the sigmoid transfer function. The application of pressure is comprised of an initial "dead zone", slow response at low pressure levels (too sensitive for users to distinguish and control), smooth transitions at median pressure levels and quick responses at high pressure levels (users often confirm pre-selection by imposing heavy pressure on a pen-tip). We employed a piecewise linear function to approximate the pressure mapping.

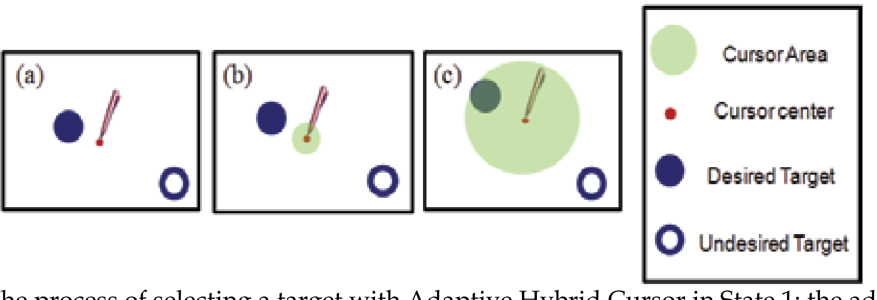


Fig. 1. The process of selecting a target with Adaptive Hybrid Cursor in State 1: the adaptive hybrid cursor employs the Zoom Cursor technique which changes the size of the cursor when targets are big in a low density environment. (a) the pen-tip lands on the screen; (b) pressure value is used to zoom the cursor. (c) pressure and location of the cursor are adjusted to make the zoomed cursor interact with the desired target. The desired target is selected by quickly lifting the pen-tip. Note that the same legend is used for Fig. 2.

If pressure causes the cursor to become too large, then more than one target might be included, and this may confuse the user. To overcome this problem, a basic principle should be specified so that when enlarging the cursor, only one target will be included at one time. Therefore, a maximum size for the cursor should be determined according to the current position of the cursor and the layout of targets. This will help to ensure that an enlarged cursor cannot include more than one target. Note that the maximum size of the cursor is dynamically changed based on the proximity of surrounding targets. We follow the algorithm used to set the radius of the cursor in Bubble Cursor. We also use a circular-shaped cursor and we allow only one target to be selected each time.
To describe the algorithm in an environment with targets *T1, T2, ..., Tn* we used the following definitions:

Minimum Distance i (MinDi): The length of the shortest line connecting the center of the Zoom Cursor and any point on the border of *Ti*.
Maximum Distance i (MaxDi): The length of the longest line connecting the center of Zoom Cursor and any point on the border of *Ti*.
A simplified version of the algorithm is as follows:
Calculate the Minimum Distance to each target: MinD1, MinD2,…, MinDn
Calculate the Maximum Distance to each target: MaxD1, MaxD2,…, MaxDn
Set maximum radius of Zoom Cursor = the second minimum value (MinD1, MinD2,…, MinDn, and MaxD1, MaxD2,…, MaxDn)
After a desired target is included by the enlarged cursor the target selection is achieved by the "quick release" manner (Ramos et al., 2004).

3.2 Zooming Target, Cursor and Background (State 2)

Using direct pointing, the selection speed has an upper limit due to human limitations such that selecting a 10 cm wide object which is within 10 cm of the human user will take less than a second, while a target which is 10 meters away will take at least several seconds to reach. Thus Bubble Radar uses RadarView to bring the targets within arm's reach so that Bubble Cursor can be subsequently easily applied for actual target selection.
Similarly, if the targets are too small and densely packed, it becomes more difficult for the user to visually locate the target. In such cases, enlarging the workspace has the effect of simultaneously increasing *A* and *W* and thus making target acquisition easier. Based on this hypothesis, we decided to enlarge the entire workspace when the target size is smaller than 1.8 mm (about 6 pixels in our experimental setup). (Ren & Moriya's study indicated that 1.80 mm is "the smallest maximum size" (Ren & Moriya, 2000)), or *EW/W* value is less than 2 where *EW* is the effective width. Here, we define *EW/W* as the density of targets, i.e. the amount of void space immediately surrounding a target. The result of pilot studies showed that the selection technique that zooms cursor, target and background at the same time could not show significant advantages above Bubble Cursor when the value of *EW/W* is more than 2. We defined an environment where the *EW/W* ratio was less than or equal to 1.5 as a high density environment, and, when the *EW/W* ratio was greater than 1.5 and less than or equal to 2, we called it a normal density environment. When the *EW/W* value was equal to or greater than 3, this was called a low density environment. High density environments are common in today's applications (e.g., a word processor or a monthly calendar viewer). Fig.2 is an illustrated walkthrough of the technique in State 2.

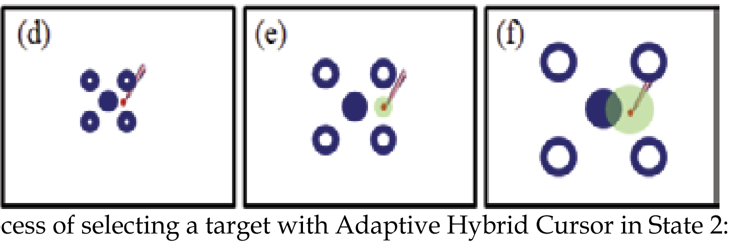


Fig. 2. The process of selecting a target with Adaptive Hybrid Cursor in State 2: Adaptive Hybrid Cursor is able to vary the size of targets, cursor and background simultaneously by pressure when approaching small targets and/or small *EW/W*. (d) the pen-tip lands on the screen; (e) using pressure value to zoom in the targets, the cursor and the background. (f) adjusting pressure and location of the cursor to make the zoomed cursor interact with the desired target. The desired target is selected by quickly lifting the pen-tip.

The maximum zoom ratio is 3 in our current design. The zoom ratio is controlled by the mapped pressure value. At the same time, Adaptive Hybrid Cursor also uses pressure and the "updated" location information of targets to zoom the cursor size according to the principles of Zoom Cursor. When the desired target was interacted by the cursor, the target selection was achieved by the "quick release" motion (Ramos et al., 2004).
The trigger for the enlargement is pen pressure which dynamically adapts the maximum zoom size of the cursor based on the zoomed surroundings, i.e., the cursor should cover no more than one object at a time.

4. Experiment

To evaluate the performance of Adaptive Hybrid Cursor, we conducted a quantitative experiment to compare it with Bubble Cursor and with the traditional technique, the regular cursor (the regular pointing selection in graphical user interfaces) as a baseline. First, Bubble Cursor, which is the current state of the art, has been shown to be the fastest desktop pointing technique. Second, Aliakseyeu et al. (2006) showed that Bubble Radar combined the benefits of Bubble Cursor in a pen-based situation. However, neither Bubble Radar nor Bubble Cursor experiments included very small targets (i.e. less than 1.6 mm). We, therefore, designed the same *EW/W* (1.33, 2, 3) ratios as for Bubble Cursor but with smaller targets (4 pixels). We wondered if Bubble Cursor offered the same advantage in smaller target situations in pen-based environments. Third, Adaptive Hybrid Cursor also employs the effective width of targets just as with Bubble Cursor, targets being allocated effective regions according to a Voronoi diagram.

4.1 Participants

Twelve subjects (11 male and 1 female) all with previous experience using computers were tested for the experiment. The average age was 24.9 years. All subjects used the pen in the right hand. All subjects had normal or a "corrected to normal" vision, with no color blindness.

4.2 Apparatus

The experiment was conducted on a Wacom Cintiq21UX, 43.2x32.4cm interactive LCD tablet display with a resolution of 1600 x 1200 pixels (1 pixel = 0.27 mm), using a wireless pen with a pressure sensitive isometric tip. The pen provides 1024 levels of pressure, and has a binary button on its barrel. The tablet's active area was mapped on the display's visual area in an absolute mode. The experimental software ran on a 3.2GHz P4 PC running Windows XP. The experiment software was implemented in Java 1.5.

4.3 Procedure

Following the protocol (Grossman & Balakrishnan, 2005), we also used a reciprocal pointing task in which subjects were required to select two fixed targets back and forth in succession, but, to simulate a more realistic two dimensional pointing environment, we changed the protocol into a multi-directional reciprocal pointing task which included reciprocal horizontal, vertical and diagonal movements. The targets were drawn as solid circles, and were located at various distances from each other along four directional axes. The goal target, the one intended to be selected, was colored green. When a goal target had been selected, it changed color to red which was an indication that the user now had to select the next goal target. Four red circles were placed around each goal target to control the *EW/W* ratio (Fig. 3).

Subjects were instructed to select the two goal targets alternately. They were told to emphasize both accuracy and speed. When the subject correctly selected the target, he/she heard a beep sound and the targets swapped colors, which was an indication of a new trial. At the start of the each experiment, subjects were given a warm-up block to familiarize themselves with the task and the conditions.

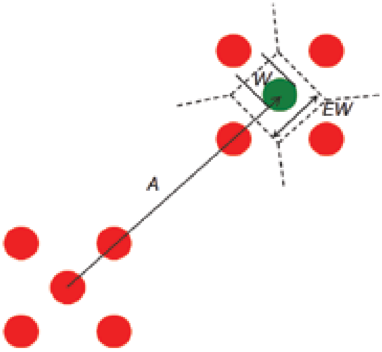


Fig. 3. Experimental setup. The solid red circle that is surrounded by four targets is the start target (as well as one of the two reciprocating goal targets), the green target is the initial goal target. The four circles around each of the start and goal targets are distracters.

4.4 Design

A within-subject design was used. The independent variables were: selection techniques *ST*, amplitude *A* (288, 576, 864 pixels), width *W* (4, 6, 12, 36 pixels), *EW/W* ratios (high = 1.33, normal = 2, low density = 3), and direction *DR* (horizontal, vertical, 2 diagonals). A full crossed design resulted in 432 combinations of *ST, A, W, EW/W*, and *DR*. The order of techniques was counterbalanced using a 3 x 3 Latin-Square.

Each participant performed the entire experiment in one session of approximately 60 minutes at one sitting, including breaks corresponding to changes in selection technique. The session consisted of nine blocks of trials completed for each technique. In each block, subjects completed trial sets for each of the 144 combinations of *A, W, EW/W, DR* appearing in random order. A trial set consisted of 3 effective attempts (4 attempts in total, but the first attempt was the starting point so that it was discarded. Note we had 3 *EW/W* ratios (high = 1.33, normal = 2, low density = 3), as previously defined in Section 3.2, so we could assess the results from different density environments.

In summary, the design of the experiment was as follows:

12 subjects x
3 techniques (Adaptive Hybrid Cursor, Bubble Cursor, Regular Cursor) x
4 target widths (4, 6, 12, 36 pixels) x
3 amplitudes (288, 576, 864 pixels) x
3 *EW/W* (high = 1.33, normal = 2, low density = 3) x
4 directions (horizontal, vertical, 2 diagonals)x
3 effective attempts (4 trials total, but the first trial is discarded due to the same starting point) x

3 blocks
= 46656 total effective selection attempts

After they finished testing each technique, the subjects were asked to fill in a questionnaire which consisted of three questions regarding "selection difficulty", "fatigue", and "overall usability" on 1-to-7 scale (1=lowest preference, and 7 =highest preference). These questions were made by referring to ISO9241-9 (2000)).

4.5 Results

An ANOVA (analysis of variance) with repeated measures was used to analyze performance in terms of selection time, error rate, and subjective preference. Post hoc analysis was performed with Tukey's Honestly Significant Difference (HSD) test.

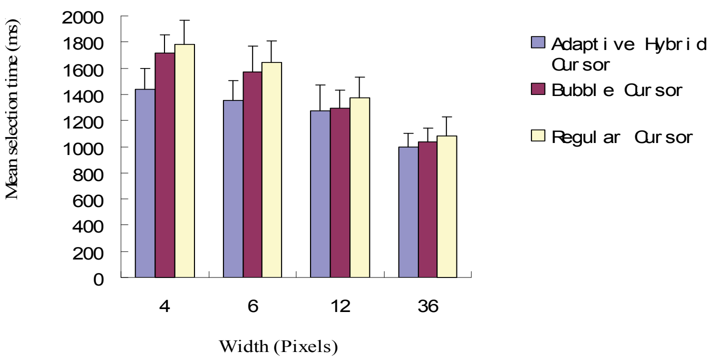


Fig. 4. Mean selection times for different sizes of targets at *EW/W* ratio=1.33.

4.5.1 Selection Time

There was a significant difference in the mean selection times among the three selection techniques, F(2,33)=13.1, p<.0001. The overall mean selection times were 1129 ms for Adaptive Hybrid Cursor, 1177 ms for Bubble Cursor and 1429 ms for Regular Cursor. Tukey HSD tests showed that both Adaptive Hybrid Cursor and Bubble Cursor were significantly faster than Regular Cursor (p<.001). No significant difference was found between Adaptive Hybrid Cursor and Bubble Cursor. Significant interaction was not found between selection technique and block number, F(4,99) = 0.56, p = .69, which indicated the learning improvement did not significantly affect the relative performance of selection techniques.

As shown in Fig. 4, at the *EW/W* ratio value of 1.33 there was a significant difference in selection time between the three selection techniques, F(2,33)=15.1 and 8.9 for the target sizes of 4 and 6 respectively, all p<.001. For target sizes of 4, 6 Tukey HSD tests showed Adaptive Hybrid Cursor was significantly faster than Bubble Cursor and Regular Cursor (p<.01), however, no significant difference was found between Bubble Cursor and Regular Cursor. No significant differences were found between the three selection techniques for the target sizes of 12 and 36.

At the *EW/W* ratio values of 2 and 3, both Adaptive Hybrid Cursor and Bubble Cursor were significantly faster than Regular Cursor, F(2,33)=8.0, 22,9, 8.8 and 19,6 for *EW/W*=2; F(2,33)=24.2, 14.0, 15.2 and 20.1 for *EW/W*=3, at target sizes of 4, 6, 12 and 36, all p<.01. No significant differences were found between Adaptive Hybrid Cursor and Bubble Cursor in both *EW/W* ratios.

The perspective brought by Fitts' law in terms of size and distance effects provided a useful framework for our design. However, it is questionable if it is valid to parameterize our results with a Fitts' law model. Adaptive Hybrid Cursor was more complex than a typical single pointing task in Fitts' law studies because it required the user to perform multiple steps, i.e., enlarge the curser and its contents by pressure, confirm the goal target, and select the goal target. Indeed, we obtained a rather poor fit between the Fitts' law model and the actual data collected, with r2 value at 0.53 for Adaptive Hybrid Cursor, and 0.87, 0.97 for Bubble Cursor, Regular Cursor respectively (we defined *ID* as $log_2(A/EW+1)$ for Adaptive Hybrid Cursor and Bubble Cursor, while for Regular Cursor $log_2(A/W+1)$). The r^2 value for Adaptive Hybrid Cursor was much lower than the values for 0.95 or lower than those found in conventional one-step pointing tasks, e.g. Accot & Zhai (2002); MacKenzie & Buxton (1992). We also looked at the data of State 1 (i.e. Zoom Cursor) described in Section 3.1. We obtained a better fit with r^2 value at 0.87 for Zoom Cursor but still lower than the values for 0.95. This was due to the fact that users had to control the size of the cursor which they do not have to do in conventional one-step pointing. The r^2 value (0.87) for Bubble Cursor was lower than the values for 0.95. This may have been due to the limitations in pen-based systems mentioned in our discussion section.

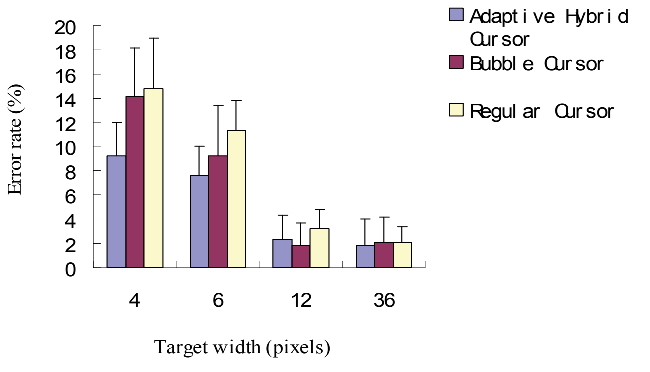


Fig. 5. Mean error rates for different sizes of targets at *EW/W* ratio=1.33.

4.5.2 Error Rate

There was a significant difference in overall mean error rate between the three techniques, F(2,33)=23.4, p<.0001. Tukey HSD tests showed Adaptive Hybrid Cursor was better than both Bubble Cursor and Regular Cursor (p<.05). Bubble Cursor was better than Regular Cursor (p<.01). Overall error rates were 4.2% for Adaptive Hybrid Cursor, 5.4% for Bubble Cursor, and 7.3% for Regular Cursor.

As shown in Fig. 5, at the *EW/W* ratio value of 1.33, there was a significant difference between the three selection techniques for the sizes of 4 and 6, F(2,33)=8.1, 4.2 p<.05. For target size of 4, Tukey HSD tests showed Adaptive Hybrid Cursor was better than both Bubble Cursor than Regular Cursor (p<.05). No significant difference was found between Bubble Cursor and Regular Cursor. For a target size of 6, Tukey HSD tests showed Adaptive Hybrid Cursor was better than Regular Cursor (p<.05). No other significant differences were found among the three techniques. There was no significant difference in error rate between the three selection techniques for the sizes of 12 and 36.

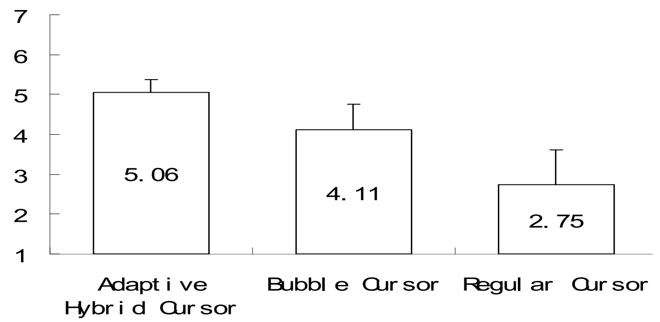


Fig. 6. Subjective ratings for the three techniques (1 = lowest preference, 7 = highest preference).

At the *EW/W* ratio value of 2, there was a significant difference between the three selection techniques for sizes 4 and 6, $F(2,33)=16.2$, 16.6 $p<.01$. For target sizes of 4 and 6, Tukey HSD tests showed both Adaptive Hybrid Cursor and Bubble Cursor were better than Regular Cursor ($p<.01$). No significant difference was found between Adaptive Hybrid Cursor and Bubble Cursor. There was no significant difference in error rate between the three selection techniques for sizes 12 and 36. The results of the *EW/W* ratio value of 3 followed trends similar to those of *EW/W*=2.

4.5.3 Subjective Preference

Fig. 6 shows the subjective ratings for the three techniques. These ratings were based on the average value of the answers given by the subjects to the three questions. Significant main effects were seen between the three selection techniques, $F(2,33)=38.4$ $p<.001$. Tukey HSD tests showed Adaptive Hybrid Cursor was better than Bubble Cursor, and Bubble Cursor was better than Regular Cursor ($p<.01$). Adaptive Hybrid Cursor was the most preferred (mean = 5.06).

5. Discussion

To improve the performance for selecting targets in a dense layout, we designed the Adaptive Hybrid Cursor (including Zoom Cursor), a novel interaction technique for pen-based systems, which enables users to adjust the size of the background, the targets and/or cursor the simultaneously. The Adaptive Hybrid Cursor dynamically adapts the permitted upper boundary of a zoomable selection cursor based on the current index of difficulty of a desired target. As shown in our Experiment, the Adaptive Hybrid Cursor showed advantages over other techniques in performance for small targets in a high density environment. The subjective preferences also showed that the Adaptive Hybrid Cursor was the most preferred technique among the three techniques tested.
Overall, the Adaptive Hybrid Cursor showed significant improvements in a pen-based selection task. It works well with a pen, and in expanding contexts. At the same time, it offers competitive selection performance without losing the background selection capability, and does not expand the context in groups of big targets, in normal and low-density environments. By contrast, many of the other mouse and pen-based interaction techniques have been shown to work well only in low density environments or on isolated targets .

Though Bubble Cursor is comparable to Adaptive Hybrid Cursor in high *EW/W* ratios or groups of larger targets in a high-density environment, it has several limitations compared to our technique, especially in pen-based environments. First, by maximizing utilization of empty screen space, Bubble Cursor trades-off the ability to select an important "target", the background. By contrast, our Adaptive Hybrid Cursor (including Zoom Cursor) allows the user to select the background (by applying lighter pressure). Second, Bubble Cursor lacks the undo function. Our technique provides "natural" cancellation by reversing the pressure value rather than using another mode-switch action like Bubble Radar (Aliakseyeu et al., 2006). Third, Bubble Cursor is not designed for pen-based environments and it does not guarantee continuous, incremental visual feedback of the selection cursor. During the experimental process we found that continuous feedback of Bubble Cursor may not always be available on a pen device (e.g., in tracking mode) because the pen-tip often loses communication with the induction area of the tablet when lifting or landing and feedback suddenly appears or disappears as a consequence. Though continuous feedback is not assured with the Adaptive Hybrid Cursor either, it can control the size of the cursor well by pen-tip pressure. Fourth, though Bubble Cursor allows denser target placement than many previous approaches, its performance advantage largely degrades when a target is closely surrounded by other objects. In theory, when the target's effective width (*EW*) approaches its actual width (*W*), little room can be used to improve the motor space. In fact, it has been shown that as the *EW/W* ratio changes from 3 to 1.33, the advantage of Bubble Cursor degrades (Grossman & Balakrishnan, 2005). In contrast, the Adaptive Hybrid Cursor can enlarge the targets, the background, and the cursor, according to the targets' surroundings. Fifth, neither Bubble Cursor nor Bubble Radar experiments have included very small targets. To further clarify, we also designed the same *EW/W* (1.33, 2, 3) ratios but with a smaller target (4 pixels = 1.08 mm). The experimental results showed that Bubble Cursor suffered from performance limitations in groups of small targets in high density environments.

We varied the essential parameters but we found it necessary to simplify our experimental design in some minor points. First, we set each target in each environment to the same size so that control of the target density parameters could be achieved more easily. Second, we used circular targets so that the distance between start point and destination target was constant in all four directions. Third, in Bubble Cursor's experiment, beside the circles around the target, many black-filled circles were also placed between the starting position and the final target as distracters on the mouse pathway. We omitted intermediate targets (i.e., distracter targets) for the following reasons. In indirect pointing environments, these distracters can significantly impact selection performance, since the subjects' selection pathway can't be avoided by the cursor. However, in a direct pointing pen-based environment, the user simply lifts the pen in the air to move from the starting position to the goal target where an out-of-range state is possible. This hypothesis was confirmed in pilot studies and in our Experiment. In addition, even though the distracters are placed between the start and destination targets, visual load will be similar for each of the techniques. Furthermore, the error rate for Bubble Cursor may increase because if the user selects a distracter he/she cannot perform the "undo" task with Bubble Cursor.

We explored the use of pen pressure for improving the performance of target acquisition tasks in pen-based environments. The experimental results have shown that pen pressure can be used to design more effective selection techniques for pen-based environments. The

Adaptive Hybrid Cursor takes advantage of pressure information. By using pressure, the Adaptive Hybrid Cursor (particularly the Zoom Cursor aspect of the technique) achieves in-place mode switching between background and target selection and requires no additional accessories. This is different from Bubble Radar's approach (Aliakseyeu et al., 2006) which uses an additional button to switch states (Li et al., 2005).
Our study contributes valuable empirical data for applying pressure for target selection techniques which had not been previously addressed in literature. This paper also suggests new ways to further improve target acquisition performance for small targets and high density environments. Future work includes applying a combination of strategies found in (Aliakseyeu et al., 2006; Yin & Ren, 2006) into the Adaptive Hybrid Cursor for large display environments and group selections.

6. Acknowledgments

This study has been partially supported by Grant-in-Aid for Scientific Research (No. 20500118), Microsoft Research Asia Mobile Computing in Education Theme, and Exploratory Software Project of IPA (Information-technology promotion agency in Japan). We are grateful for the work and support of all the members of the Ren Lab in Kochi University of Technology.

7. References

Accot, J. & Zhai, S. (2002). More than dotting the i's - foundations for crossing-based interfaces, *Proceedings of CHI2002*, pp. 73 – 80, ACM Press.

Aliakseyeu, D., Nacenta, M., Subramanian, S. & Gutwin, C. (2006). Bubble radar: efficient pen-based interaction, *Proceedings of. AVI'06*, pp. 19-26, ACM Press.

Barrett, R., Olyha, J., Robert S., & Rutledge, J. (1996). Graphical User Interface Cursor Positioning Device Having a Negative Inertia Transfer Function. Patent # 5,570,111, IBM Corp.

Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B. & Zierlinger, A. (2003). Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch and pen operated systems, *Proceedings of INTERACT '03,* pp. 57-64, IOS Press (IFIP).

Bezerianos, A. & Balakrishnan, R. (2005). The vacuum: facilitating the manipulation of distant objects, *Proceedings of CHI2005,* pp. 361-370, ACM Press.

Blanch, R., Guiard, Y., & Beaudouin-Lafon, M. (2004). Semantic pointing: improving target acquisition with control-display ratio adaptation, *Proceedings of CHI2004,* pp. 519-526, ACM Press.

Buxton, A.S.W. (1990). Three-State Model of Graphical Input, in Human-Computer Interaction, *Proceedings of INTERACT '90,* pp. 449-456, Elsevier Science Publishers.

Cockburn, A. & Brock, P. (2006). Human on-line response to visual and motor target expansion, *Proceedings of Graphics Interface 2006,* pp. 81-87.

Collomb, M., Hascoët, M., Baudisch, P. & Lee, B. (2005). Improving drag-and-drop on wall-size displays, *Proceedings of Graphics Interface 2005,* pp. 8125-8132.

Fitts, P.M. (1954). The information capacity of human motor system in controlling the amplitude of movement, *Journal of Experimental Psychology*, 47, pp. 381-391.

Foley, J.D., Wallace, V. & Chan. P. (1984). The Human Factors of Computer Graphics Interaction Techniques, *IEEE Computer Graphics and Applications*, pp. 13-48.

Grossman, T. & Balakrishnan, R. (2005). The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area, *Proceedings of CHI '05,* pp. 281-290, ACM Press.

Guiard, Y., Blanch, R. & Beaudouin-Lafon, M. (2004). Object pointing: a complement to bitmap pointing in GUIs, *Proceedings of Graphics Interface 2004*, pp. 9-16.

Herot, C.F. & Weinzapfel, G. (1978). One-point touch input of vector information for computer displays, *Proceedings of the 5th annual conference on Computer graphics and interactive techniques,* pp. 210-216, ACM Press.

ISO ISO9241-9 (2000). Ergonomic design for office work with visual display terminals (VDTs)–Part 9: Requirements for non-keyboard input devices. *International Standardization Organization*.

Kabbash, P. & Buxton, W.A.S. (1995). The "prince" technique: Fitts' law and selection using area cursors, *Proceedings of CHI '05,* pp. 273-279, ACM Press.

Li,Y., Hinckley K., Guan, Z. & Landay J. (2005). Experimental analysis of mode switching techniques in pen-based user interfaces, *Proceedings of CHI '05,* pp. 461-470, ACM Press.

MacKenzie, I.S. & Buxton, W.A.S. (1992). Extending Fitts' law to two-dimensional tasks, *Proceedings of CHI '92*, pp. 219-226, ACM Press.

McGuffin, M. & Balakrishnan, R. (2002). Acquisition of expanding targets, *Proceedings of CHI '02*, 57-64, ACM Press.

Nacenta, M.A., Aliakseyeu, D., Subramanian, S. & Gutwin, C.A. (2005). A comparison of techniques for multi-display reaching, *Proceedings of CHI '05*, pp. 371-380, ACM Press.

Parker, K., Mandryk, R., Nunes, M. & Inkpen, K. (2005). TractorBeam Selection Aids: Improving Target Acquisition for Pointing Input on Tabletop Displays, *Proceedings of INTERACT '05*, pp. 80-93, Springer.

Ramos, G., Boulos, M. & Balakrishnan, R. (2004). Pressure widgets, *Proceedings of CHI2004,* pp. 487-494, ACM Press.

Ramos, G. & Balakrishnan, R. (2003). Fluid interaction techniques for the control and annotation of digital video, *Proceedings of UIST''03,* pp. 105-114, ACM Press.

Ramos, G. & Balakrishnan, R. (2005). Zliding: fluid zooming and sliding for high precision parameter manipulation, *Proceedings of UIST''05,* pp. 143-152, ACM Press.

Ren, X. & Moriya, S. (2000). Improving selection performance on pen-based systems: A study of pen-input interaction for selection tasks, *ACM ToCHI*, 7(3), pp. 384-416.

Worden, A., Walker, N., Bharat, K. & Hudson, S. (1997). Making computers easier for older adults to use: area cursors and sticky icons, *Proceedings of CHI '97*, pp. 266-271, ACM Press.

Yin, J. & Ren, X. (2006). The Beam Cursor: A pen-based technique for enhancing target acquisition, *Proceedings of HCI 2006*, pp. 119-134, Springer.

Zhai, S., Conversy, S., Beaudouin-Lafon, M., & Guiard, Y. (2003). Human on-line response to target expansion, *Proceedings of CHI2003*, pp. 177-184, ACM Press.

12

Towards a Conceptual Framework and an Empirical Methodology in Research on Artistic Human-Computer and Human-Robot Interaction

Uwe Seifert, Jin Hyun Kim

1. Introduction: Cognitive musicology, media sciences, and cognitive science

In this paper we give an overview from the point of view of cognitive musicology of our theoretical as well as empirical approach to human-computer interaction, especially human-robot interaction in artistic contexts such as music and new media art.

With our approach we pursue several general purposes. A first goal is to establish the cognitive science of music as a scientific research program on the musical mind as part of a science of mind in the epistemological framework of cognitive science in musicology. Cognitive science of music in the disciplinary field of musicology is termed "cognitive musicology". The second goal of our approach is to extend cognitive musicology to the study of new media art, especially artificial life art and musical robotics, interactive audio programming, and augmented environments. As a third goal we aim at combining research issues from media and cognitive science. Key concepts in this framework are "mediality" and "cognitive artifact". An educational goal of our approach is to facilitate the understanding of "computation" and the use of algorithmic agents in music and the arts for students in the liberal arts or humanities. Therefore, robot programming has been introduced as part of a general education on information technology and cognitive science in some of our courses.

Cognitive musicology is located in the humanities. Musicology is divided into historical musicology, systematic musicology, and ethnomusicology. Traditionally, historical musicology is methodologically associated with hermeneutics and history. Today, ethnomusicology as conceived of as part of cultural or social anthropology is more related to cultural approaches to music than in its beginning. In its beginning ethnomusicology was closely related to biological anthropology and was called "comparative musicology". Systematic musicology, like sociology and psychology, has been strongly influenced by the methodology of the natural sciences, especially the methodology of physics. With the advent of digital computers and cognitive science, that part of systematic musicology that adapted the methodological and epistemological framework of cognitive science was termed "cognitive musicology" (e. g. Seifert 1991a, 1991b 1993, Seifert 2004;

Leman/Schneider 2006). Within cognitive science there are three approaches to a scientific theory of mind: cognitivism, connectionism, and interactionism (Kim/Seifert 2006, Seifert/Kim 2007). Interactionism has

been discussed under different labels such as embodied cognition (Clark 1997), embodied cognitive science (Pfeifer/Bongard 2006, Pfeifer/Scheier 1999), situated cognition (Smith 1999), and distributed cognition (Hollan/Hutchins/Kirsch 2000). In this approach of cognitive science, cognition is viewed as a dynamic interactive relation of a situated agent and its biological as well as socio-cultural environment. Our approach to the musical mind within cognitive musicology adopts ideas from interactionism (Kim/Seifert 2006, Seifert/Kim 2007), which is strongly related to the term "embodied cognitive science of music" (Schmidt 2005, 2007, 2008).

In Germany the "Geisteswissenschaften" or humanities have been methodologically opposed to the natural sciences, assuming a special hermeneutic and historical approach towards understanding the "mind". Today the classical foundations of this approach in German idealism and the concepts of "Geist" (Wilhelm Dilthey) and "Kultur" (Wilhelm Windelband, Heinrich Rickert, Max Weber) as interpretative sciences have been either forgotten or critically questioned. At present the humanities are striving for a new epistemological as well as methodological foundation. Cultural studies or media studies or sciences are under discussion as new research paradigms in the humanities.

Since 2002 our research project in the domain of cognitive musicology has been part of the collaborative research center SFK/FK 427, "Media and cultural communication", addressing questions of a methodological and epistemological foundation by using "medium" and "mediality" from the media science point of view. "Mediality" emphasizes the relevance of external representations and processes mediated by a "medium" which not only serves as a passive means of conveying the message, information or intention, but also participates in shaping. "Mediality" is therefore understood as a functional term. For research on (human) cognition, media science raises the question of how media support and extend the working of the (human) mind. Our research project "Transcriptive Interaction" (2002-2004) focused on interactive music systems in electro-acoustical music performances and musical gestures. It soon became evident that the use of interactive music systems leads to a new understanding of "music" and to the emergence from traditional categories of art (e.g. music, dance, theater, and film) of new media art for which human-computer and human-robot interaction seem to be constitutive. In 2005 we therefore extended our research to new media art in a research project entitled "Artistic Interactivity in Hybrid Networks". One key idea was to use robots as tools for research as well as for teaching computational aspects of new media art and cognitive science. In 2006 we obtained a nearly undocumented prototype of a Khepera-III robot and started the first steps in programming. Because we realized the general difficulty of the task of robot programming, we introduced LEGO Mindstorms at the end of 2006 in our courses. Both systems are now parts of our research project (Schmidt 2006; Schmidt/Seifert 2006; Kim/Seifert 2007). Robotics in the field of musical applications might be termed "musical robotics" (Schmidt/Seifert 2006).

To summarize: In general, our research methodology for the study of the musical mind combines ideas and methods from new directions of media and cognitive science. We claim that human-computer and human-robot interaction in new media art provide the most adequate natural setting in order to explore the musical mind scientifically without reductionism and loss of complexity. In order to see why our claim might be justified, we have to discuss computational modeling approaches to the scientific study of the embodied mind. The role of social interaction, situatedness, affordances and cognitive artifacts for the investigation of the higher mental functions of the human's mind has to be taken into account. Methodologically, the relations between computational modeling, measurement and empirical methods of data acquisition for natural socio-cultural surroundings like human-computer and human-robot interaction in media art and music are to be discussed.

2. Cognitive Artifacts: From Digital Musical Instruments to New Media Art Environments

For scholars from the humanities it is evident that socio-cultural contexts, symbols and other artifacts may function as media and play an important role in studying the mind, its functioning and its historical development. Researchers trained in the natural sciences are often skeptical about these "complex" phenomena and their role for explaining mental functions. They prefer to ground their research on the mind epistemologically and methodologically in biology or physics and experimental research in laboratories. Therefore, let us first address the idea that new media art environments form an extension of the traditional laboratory approach to a more natural (social) setting and that human-computer interaction, and especially human-robot interaction in connection with (embodied) cognitive science are the most interesting approaches for the researchers from humanities and natural sciences to study the functioning of the mind in such surroundings.

For musicologists, musical instruments are tools, media, or artifacts for realizing the sociocultural phenomenon of "music". Music and the production of musical instruments are embedded in a cultural and social context. For natural scientists musical instruments are physical devices for sound production. The production of musical instruments is the construction of a physical device for the production of sound called "music". Usually, the idea of music and a musical instrument is grounded in common sense, general education and accepted social norms. This may be illustrated by most of the interfaces for musical expression presented at the well-known conference "New Instruments for Musical Expression" (NIME). Usually, most of the interfaces presented at NIME conferences can be viewed as an enhancement of the classical idea of a musical instrument and as an extension of craftsmanship to design musical instruments in the realm of computer-based sound generation. A musical instrument is conceived of by most researchers as a physical device optimized for the generation of "music" and "musical expression". Therefore, an interface should enhance "musical expression" for computer-based electro-acoustical sound generation. (e.g. Wanderley/Battier 2000; Miranda/Wanderley 2006). For a better understanding of the current situation, a historical sketch of the developments of design considerations for digital musical instruments will be given.

3. On the history of the design of digital musical instruments and musical robots

"Digital musical instrument" is not identical with "the computer as a musical instrument" (Mathews 1963). The interface designer Axel Mulder prefers the term "virtual musical instrument" instead of "digital musical instrument". But the latter is used as a more general term in the context of computer-aided musical instruments (Miranda/Wanderley 2006). "The computer as a musical instrument" is a term going back to the first invention of the computer as a device to generate sounds by numbers and algorithms. "Digital musical instruments" indicates an extension of the idea of the computer as a musical instrument to computer-aided instruments which not only have the capacity to generate sounds, but also include a gesture interface through which a performer interacts with a computer system by means of her or his bodily actions influencing the mechanisms of sound generation. However, the relation of a gesture interface, a so-called controller, and the sound generator of digital musical instruments – other than physical-acoustic musical instruments – are in principle decoupled from each other, since the process of algorithmic sound generation is not controlled by any physical energy, but purely by information represented as numbers. The artificial coupling of these two units is therefore a core task of the design of digital musical instruments.

Against this background, the strategies of gesture mapping, i.e. mapping from gesture input data into parameters for algorithmic sound synthesis, come to the fore in the design of digital musical instruments. Gesture mapping produces a meaning of a performer's bodily activities which can be interpreted as a musical (instrumental) gesture and an intermedial relationship between bodily gestures and sound structures. In this way, digital technology creating new media (in our case: digital musical instruments) brings about either a reflection on a traditional concept (in our case: music or musical instrument) or a new concept. An established classification of digital musical instruments which up to now have been developed as *augmented musical instruments*, *instrument-like gestural controllers, instrument-inspired gestural controllers, alternate [or alternative] gestural controllers* (Miranda/Wanderley 2006) is related to different strategies of gesture mapping.

The term "*augmented musical instrument*" stands for physical-acoustic musical instruments or electronic musical instruments extended with a computer system. The instrumentalist's significant physical gestures used in playing a musical instrument are detected by sensors which are appropriate for measuring the desired information relating to physical gestures. The sensors' electrical signals are digitalized and used as the input data of a computer system to control algorithmic sound synthesis and processing. *Augmented musical instrument*s are characterized by their extension of control dimensions compared to musical instruments. For instance, a series of string interfaces augmenting physical-acoustic and electronic string instruments belong to this category. Such an interface, e.g. the *MIDI Bow* developed by the composer Jon Rose in cooperation with the Studio for Electro-Instrumental Music (STEIM) since 1986 (see figure 1), allows an instrumentalist to use more or less traditional instrumental techniques such as bowing. Ultrasonic sensors mounted both on the *MIDI Bow* and on the right arm of the violinist allow the computer system to measure the bowing movements of the performer. Sounds are generated both by the physical mechanism of the instrument and by the computer system transforming physical gestures into control parameters for algorithmic sound generation and processing. In this way, the control dimensions of the violin are extended.

Fig. 1. The composer Jon Rose playing his MIDI Bow, courtesy of the Studio for Electro-Instrumental Music (STEIM/Amsterdam)

The difference between *instrument-like gestural controllers* and *instrument-inspired gestural controllers* consists in strategies of gesture mapping which aim at the simulation of physical-acoustic musical instruments on the one hand and at the relations of an instrument-like controller and unexpected musical events on the other. The *SuperPolm*, a virtual violin, developed by the composer and media artist Suguru Goto, can be taken as an example of the latter (see figure 2). This virtual violin does not possess a resonator and is equipped with position-measuring sensors instead of strings. An accelerometer measures the x- and y-axis movements of the violin body to detect bodily posture variations expressively guiding the playing of the *SuperPolm*. The design of this interface allows the instrumentalist, using more or less traditional violin techniques, to explore a new relation between usual instrumental gestures and generated musical results.

Fig. 2. Suguro Goto playing the *SuperPolm*, courtesy of Suguru Goto

A group of interfaces which do not bear any similarity to traditional musical instruments are termed "*alternate [or alternative] gestural controllers*". The function of these controllers can, however, be designed to be similar to that of a traditional musical instrument by gesture mapping, allowing a performer to use musically meaningful gestures to generate comprehensible and reproducible musical events from the intentionally produced gesture segments. Alternatively, they can provide a completely new relationship between physical gestures and sound events, so that the performer has to develop her or his own image of this intermedial relation, although gestural activities may have an effect on some aspects of sound structure. An example of the latter is the data glove interface *Lady's Glove*, invented by the media artist Laetitia Sonami in 1991 and developed in cooperation with STEIM (see figure 3). Different sensors such as Hall-effect sensors, accelerometers and bending sensors measure respectively the distance of the thumb to the rest of the fingers, the speed of hand movements and the bending of fingers. The principles of gesture mapping of this interface vary in each performance of Sonami, so that a certain strategy of gesture mapping serves not only to interface design, but rather as a dramaturgy of media performance.

Fig. 3. Laetitia Sonami's performance with the interface *Lady's Glove*, developed in cooperation with the Studio for Electro-Instrumental Music (STEIM), courtesy of STEIM

The rethinking of the concept of the "musical instrument" has also been an important issue in recent projects on musical robotics modeling artificial bodies involved in music processing. Although projects such as *Heilphon* and *Beatbot* developed by the group *Ensemble Robot* as well as the project *RoboticMusic* developed by the media artist Suguru Goto (see figure 4) were inspired by the idea of traditional musical instruments, they try to go – in words of the current director of *Ensemble Robot* – "beyond traditional boundaries imposed by physical limits of the human body" (Southworth 2006: 17). These projects therefore focus on the development of motors imitating the human motor organs necessary for playing each musical instrument and enabling more capacity needed for instrumental technique.

Fig. 4. *RoboticMusic* by Suguru Goto, courtesy of Suguru Goto

Beyond the simulation of motor capacities for playing a musical instrument, some recent projects of musical robotics explore the aspects of musical interaction necessary for collaborative music-making such as imitation and synchronization. For instance, the percussionist robot *Haile* developed by Gil Weinberg and his team of the Music Technology Group at the Georgia Institute of Technology (see figure 5) has been designed to play in an ensemble with human percussionists based on sequential (decentralized) and simultaneous (centralized) schemata of interaction enabling imitation and synchronization, while a rhythmic motive played by a human percussionist and analyzed by *Haile* serves as a unit of action which can be imitated and synchronically accompanied (Weinberg/Driscoll 2006, 2007).

Fig. 5. Improvisation of *Haile* accompanying human percussionists. Courtesy of Gil Weinberg

This brief sketch of the history of the design of musical instruments indicates a development from purely digital sound-generating systems to systems based on interactions of a performer's body, with a digital system for sound generation. Furthermore, semi-autonomous systems called "musical robots" are designed as "partners" in music making. In general, "body" and "interaction" are becoming relevant for considerations of the design of digital musical instruments. However, a whole media art environment is rarely viewed as a "musical instrument", in the same way as concerning musical robotics a robot is hardly viewed as a real partner in "music" or "art making".

4. New media art environments will change our view of music and art in general and might become "laboratories" for investigating the functioning of the (human) mind

Most researchers involved in the design of digital musical instruments take the sociocultural roles of a craftsman to build musical instruments, of an artist to create art works, of a performer to realize art works, and of a consumer or recipient for enjoying finished works of art. However, these roles seem to change within new media art: The artist will become much more a designer of "scaffolding" for art experiences instead of a creator of art works. She or he will become much more a facilitator for aesthetic experiences than a creator of art works. Artists, e.g. composers in the case of music, in new media art do not create a final work or opus. They will create a framework or an artificial environment to explore or play with sounds and information from other sense modalities.

At the same time, in using these "scaffolds" the recipient or consumer himself/ herself will become a performer or an artist. Traditionally, the social contexts of experiencing music are clubs, concert halls, operas, open-air concerts, and house concerts. They are related to social conventions as well as to musical styles. All this is well known. Our thesis is that with the advent and further development of interactive systems, art will change the roles of the participants in art process, too. Let us explain this idea briefly: There are activities that have their goals outside themselves. The result or product of the activity is of main importance. In other cases the activity itself becomes important, as in playing, music making, thinking, or in aesthetic situations like experiencing or creating art works. The creation of art works has its goals in itself. The main point of our argument is that in media art environments the goal is to get the consumer involved in aesthetic experiences. These experiences might be conceived of as consisting of attention and a pleasant feeling. The main task for the artist designer is accordingly to develop "scaffolds" that enable processes of attention in combination with pleasant feelings.

It seems that in the near future art and entertainment will meet in new media art. The idea of music as an autonomous art and the idea of a musical instrument seem to be changing within new media art.

In new media art mediality and the functioning of cognitive artifacts become obvious, because the user of new media art is obliged to explore the possibilities of an unknown environment and the affordances it supplies. In connection with the exploration of the environment, she or he has to develop habits and concepts to understand and become adapted to the environment. Furthermore, the user is confronted with her or his socio-cultural norms and those embodied in the environment.

Given these circumstances, these new environments provide the opportunity to study social context and cognitive modeling in natural settings. Evidence for the need and relevance of such investigations will be given in the next chapters.

5. From computational modeling intramental processes to modeling intermental processes: Robots as modeling tools in embodied cognitive science and interactionism

From a biological point of view the role of the body and the coupling to its (natural) environment has to be emphasized. This biological aspect might be called embodiment. Jakob von Uexküll's distinction between sign-world (Zeichenwelt) and action-world (Wirkwelt) that constitutes an animal's environing world (Umwelt) is noteworthy: Every biological species, depending on its body and nervous system, has its own action-world. But not only the constitution of an organism is of importance for the investigation of cognitive behavior. An organism's interaction with its environment is supported by the environment's supply of affordances (Gibson 1997). These species-specific affordances have evolved as part of the evolutionary history of an organism's interaction with its environment. In general, one might think of an organism's interaction with its environment more abstractly as an embodied functionality of an embedded agent exploring the affordances of its natural environment. These biological aspects of cognition and an agent's interactions with its natural environment are studied by embodied cognitive science. In embodied cognitive science robots are used as modeling tools to investigate cognitive processes, communicative behavior, and an agent's interaction with its environment (Pfeifer/Bongard 2006). Two aspects are of importance to the study of cognitive processes in the realm of embodied cognitive science: 1) internal computational structures and processes that support the processing of an embodied functionality and an agent's behavior; 2) the embedded agent's interactions with its environment and its communication with members of its own species.

Within embodied cognitive science, robots are used as embedded agents to study both of these aspects of cognitive processing. In analogy to natural systems a robot's sensors are considered as "perceptual" interfaces of a sensory system. Programs are descriptions of an embodied system's internal processing structures for planning, perception, cognition, volition and communication (Arbib 2004, p. 759-761). Actuators of a robot are in analogy to natural systems considered as interfaces of the "action"-systems. Computational modeling of cognition with robots has several advantages: 1) Interactions are taking place within the natural physical world; 2) Implementations of supposed algorithms for interaction and communication in connection with perception and cognition are possible; 3) Measurements and analysis of "internal" sensory data are possible; 4) The view that senses are some kind of measurement instruments can be tested empirically by using sensors in robots.

But the natural environment is not the only factor that should be taken into account in cognitive modeling, since humans are not only embedded in natural environments. They are embedded in social environments, too. Humans interact in a social environment using symbols and other artifacts. Embeddedness in a socio-cultural environment might be called situatedness. "Cognitive artifacts" might be used as a generic term for the social use of artifacts and symbols. As is well known in the humanities and as Lev Vygotsky and Alexander Luria (Luria 2004) pointed out, neuroscientific research focusing solely on brain processes or psychological research focusing solely on individual processes of mental functioning is not sufficient for understanding higher human mental functions: It is important to investigate the influence of intermental functioning, external representations and social contexts in order to understand the intramental functioning of mental and brain processes of cognition in the individual. In other words: "the social interactional dimensions of intermental functioning" and the role of cognitive artifacts should not be neglected in research on the human mind (Wertsch 1999, p. 879). We assume that music is a higher human mental function. So it is research object of cognitive musicology as a science of the musical mind, which is based on a computational approach to music cognition.

Concerning the use of computers, robots or – more generally – algorithmic agents in new media art and music, our approach is based on the assumption – as is true of all other cognitive artifacts – that interaction with such systems has effects on human cognition and behavior. But in comparison with other cognitive artifacts, these artifacts are capable of autonomous actions to some extent, and mimic human cognitive function and behavior to some extent: They are agents or actors, and their behavior in a social setting may serve as a second-order cognitive artifact. In general, cognitive artifacts organize functional skills and are embedded in larger socio-cultural systems. Cognitive artifacts are best considered as "categories of processes that produce cognitive effects by bringing functional skills into coordination with various kinds of structures" (Hutchins 1999, p. 127). Therefore, our goal is to investigate to what extent new media environments and the use of algorithmic agents in music "scaffold", "shape", or "mediate" interactive artistic activities, and how they are related to intermental and intramental cognitive, perceptual, and aesthetic processes.

How the mind operates in a natural and social context is an important question for cognitive science as well as cognitive musicology. Modeling internal processes for social interaction is necessary because of the essential assumption that conscious behavior and cognition as intramental processes are based on such intermental processes. In other words: interaction with an environment and other agents is essential to cognition (Hutchins/Hazlehurst 1995, Hutchins 1999; Luria 2004). Interaction is not only conceived of as interaction within a biological environment but, insofar as humans' mental processes and their activities are involved, as social interaction constituting intermental processes by exploring social affordances supplied by the social environment and cognitive artifacts. We think that the idea of biological and social affordances is also important (Gibson 1977). This can best be studied within social robotics, because a social robot is defined as "an autonomous robot that interacts and communicates with humans or other autonomous physical agents by following social behaviors and rules attached to its role."

Cognitive musicology within the framework of embodied cognitive science of music or interactionism therefore has to take into account interaction, situatedness, embodiedness, cognitive artifacts and social affordances. In our opinion new media art environments as cognitive artifacts offer unknown social affordances for developing aesthetic experiences and entertainment. These social affordances must be studied empirically. But how could these social affordances supplied by cognitive artifacts that shape intermental processing by social interactions be detected and studied?

Environments in new media art can be used as testbeds for the study of more realistic laboratory situations. For example, the Casa Paganini in Genova offers the opportunity to study music-making and measurements in both traditional and new media art performances. In such environments many new ideas might be investigated empirically:

How can robots be studied as partners in the social interaction of music-making or art-making? How do they function as semi-autonomous musical instruments in music making? What kind of social functionality does a situated agent have to embody? How is a specific embodied social functionality related to e. g. embodied perceptual functionalities? What are the important social affordances that facilitate the embodiment of social functionalities? How are social affordances and cognitive artifacts related? How do they influence or shape intramental processing? What are the basic observable units for studying artistic human-robot interaction? To what extent do these units rely on the social environment, the social role of the interacting systems and their tasks? How are they established between humans and robots?

At present no standardized approach to human-robot interaction exists (cf. Fong/Nourbakhsh/Dautenhahn 2003 First, HCI/HRI as a research problem for artistic contexts is not widely recognized. Second, researchers explore using traditional methods from social sciences, psychology or ethology. Third, it is assumed that the methods that are normally applied to the study of human-human interaction could be successfully transferred to the field of human-robot interaction. However, we do not believe that an application of sociological or psychological methods for studying social human-human interaction is possible without changing them considerably. Equally, applications from ethology used to study social animal-animal interaction seem to be inappropriate. In general this seems to be problematic because in the case of human-robot interaction mixed categories of species are under study: humans and robots. At present, not much is known about these special interactions between a natural and an artificial species, how they might take place and how they might adapt to each other. Therefore, we raise the question of how we should or could study human-robot interaction. In our opinion, this can best be undertaken with structured observation as a starting point.

6. Representational measurement theory and structured observation: Observation – some general remarks

Despite the general belief of most researchers educated in the field of natural sciences that observation is epistemologically without presuppositions and results in objective data or hard facts, it has been shown that observation presupposes both perceptual and knowledge structures. First, one has to semantically distinguish process and result, taking into account the fact that observation is the result of observing. Next, one has to bear in mind some conditions underlying observation. A first condition for observation is that the observing system must be capable to distinguish objects or events from a background. So observation presupposes perception. This may be expressed by the following definition of the process of observing: A system s observes a fact X if and only if s perceives an object or event e and subsumes this object or event e under a family F of concepts such as "e has property z", "e is standing in relation x to u" and so on. In order to explain observation one has to explain perception. In general, perception is explained by different psychological approaches to perception in different sense modalities. Nevertheless, from our point of view cognitive science with its computational approach comes into play. Perceptual processes are conceived of as computational processes which can be described by computer programs.
A further condition one has to bear in mind as expressed in the definition of observation is that observation presupposes conceptual structures. From a logical point of view concepts are the rules for the applications of an expression indicating that concept. In axiomatic theories these rules might be definitions or axioms. More naturally, they might be either learned or innate, or even, as in science, consciously chosen for research. A further problem with observation by humans is the application of bio-, socio- and techno-morphic concepts in describing observations, because they might be misleading in pretending to have some explanatory power for an unknown area. Because of these dangers and to avoid "anthropomorphic" pitfalls but without neglecting its conceptual and perceptual presuppositions methodologically, human observation should be related to measurement.

7. Measurement in the psychological and social sciences: Representational measurement theory

Most methods adapted from psychological and sociological research by human-computer and human-robot interaction researchers are based on statistical reasoning and hypothesis-testing. In general, this standard methodology does not support theory-building or theoretical generalization because it is based on the idea of testing singular statements (Bischof 1995, Lehmann 1985, p. IX, Eberlein 1980, p. 527) and its unreflected use in inquiry has been often criticized (e. g. Glymour 2001, pp. 171). Furthermore, it seems to be only the first step in the historical development of psychology as an empirical science (Lehmann 1985, pp. VIII-IX, pp. XV). Four stages in the development of psychological research methodology and theory building can be distinguished (Lehmann 1985, pp. IX-XIII). The first stage, starting in the nineteenth century and now used in the mainstream of psychological research, is the statistical approach of testing singular statements using hypothesis testing. 2) The second stage evolved in the middle of the twentieth century as mathematical psychology, and is concerned mainly with ad-hoc models for fitting experimental data. 3) The third stage is characterized by measurement theory.

Representational measurement theory has been the main approach towards clarifying the concept of measurement in the psychological and social sciences. It started with the work of Patrick Suppes and Dana Scott in 1958 and was developed further in the second half of the twentieth century. Nowadays, it is a standard in psychological and sociological textbooks on statistics, measurement, and mathematical psychology (e. g. Coombs/Dawes/Tversky 1970, p. 4). 4) The fourth stage is based on model and recursion theory. Of course the fourth stage corresponds to the epistemological framework of cognitive science – automata theory. In this framework internal processes are viewed as computational processes, and computer programs may be used in order to describe these processes. Programs, which are conceived of as descriptions of internal processes between perception and action, might therefore substitute the intervening variables or hypothetical constructs of psychological theory-building.

In representational measurement theory, measurement is defined as a homomorphism h from an empirical relative E into a numerical relative N, i.e. (E, N, h). Axioms must be satisfied in the empirical relative to be represented by a numerical relative. If the axioms of the empirical relative are valid in the numerical relative, it is said that the numerical relative represents the empirical relative. Normally this is proven by a representation theorem.

Establishing the existence of a homomorphism for an empirical relative is called the representation problem. The next problem is called the uniqueness problem. It must be shown under which transformations the operations remain valid so that the scale level remains the same. Generally four classes of scales are distinguished: nominal, ordinal, ratio and absolute scales. A third problem is meaningfulness. To what extent do mathematical operations make sense for the domain under study?

8. Representational measurement theory: Structured observation as measurement

It can be shown that structured observation is some kind of measurement in the sense of representational measurement theory (Greve/Wentura 1997). In order to count as measurement in the sense of measurement theory, structured observation must satisfy some logical requirement from measurement theory. For structured observation to be some kind of measurement it must be shown that the observational categories of a coding scheme satisfy the requirements of a nominal scale. In general, these requirements are those for the classification of a certain domain: The classes must be mutually exclusive. All objects of the domain must be classified, and each object is an element belonging only to one class. All classes of the domain must contain at least one object of the domain, i.e. there exists no empty class. In general, an equivalence relation introduces a partition into a given domain of objects and corresponds to a classification. Because structured observation is related to measurement, we think that it is best to start with structured observation in developing a methodology for human-computer and human-robot interaction in new media art environments.

9. Structured observation and observer reliability: Cohen's kappa as a measure of observer agreement

Observational methods may be classified concerning their degree of pre-structuredness and the degree of participation of the observer (Robson 2002). Structured observation is highly structured non-participant observation and has been shown to count as measurement as indicated above (Grewe/Wentura 1997). There are two main steps in developing the method of structured observation (Bakemann/Gottman 1997; Robson 2002).

First, a coding scheme has to be developed. This coding scheme relies on the development of possible categories that might be observed in the domain under study. This first step is conceptually difficult and implies the development of categories or classes for describing the observations. In order for structured observation to count as measurement these categories must satisfy the logical conditions of a nominal scale. These categories must satisfy the requirements of a nominal scale.

Second, observers have to be used as measurement devices. These observers must reliably recognize the observational categories. Therefore, observer training is necessary. Observer calibration ensures that the same results will be obtained from different observers in the same situation, i.e. the observers' judgments are in agreement. To ensure observer reliability, the observers must be "calibrated". A measure for observer agreement is used to calibrate the observers. The reliability of an observer's judgments is tested by comparing her or his judgments with those of another observer observing the same situation. It is measured whether their judgments agree. Different measures of observer agreement or reliability are possible (Wirtz/Caspar 2002). Cohen's kappa coefficient is a well-known measure for observer calibration and for measuring their reliability, in order to ensure the reliability of the data obtained (Bakemann/Gottman 1997, Grewe/Wentura 1997, Robson 2002, Wirtz/Caspar 2002).

The main idea behind Cohen's kappa is correcting for chance agreement percentages of observer judgments by two observers.

Usually the development of the categories of a coding scheme is done in connection with observer training. That is why both steps and observational studies are in general highly time-consuming.

To summarize: The importance of the development of an empirical methodology in accordance with the representational measurement theory of the social and psychological sciences is emphasized. At present, not much is known about this, and no methodology exists – even in human-robot interaction – to address the problem. Therefore, relevant category systems for observational studies of artistic human-robot interaction in the contexts of new media art need to be developed. It is argued that structured observation should be used in empirical research on human-computer interaction, especially on human-robot interaction, because it fulfills the high demands on measurement as required by the representational measurement theory of psychological and sociological methodology.

Observer training and observer calibration using different measures of observer agreement should be used to prepare the data collection. The collection of data should be based on observer protocols from direct observations as well as video recordings and the registration of robotic and human sensor data.

10. Cognitive musicology: Robot studies and structured observation

In order to become acquainted with research on human-robot interaction and robot programming, three exploratory studies using LEGO Mindstorms NXT were carried out. LEGO Mindstorms NXT robots have been used as a tool in order to develop a methodology for research on human-robot interaction in an artistic context.

The first experimental study, a Master thesis by our student Birgitta Burger, was carried out in cooperation with the KTH Stockholm (Burger 2007a, 2007b, Burger/Bresin 2007). In this study a "mainstream" experimental approach using a questionnaire with rating scales was used. The recognition and its enhancement through music of intended "emotions" communicated through a robot's movements using LEGO Mindstorms NXT was investigated. The main result of this study was that there is some difficulty in distinguishing between the movements expressing “joy” and “anger”.

In modifying the experimental set-up of the first study, the second study explored the methodological tool itself: the use of the kappa coefficient for the measurement of observer reliability. This study was carried out in 2007 during the International Summer School in Systematic Musicology (ISSSM-07) at Ghent University. We experimented with observer training in connection with measurements of observer reliability. The usefulness of the well-known kappa coefficient as a measurement for inter-observer agreement was explored in order to use human observers as "measurement instruments" in complex situations such as artistic human-robot interaction. Observers were trained to recognize three classes of movements exhibited by LEGO Mindstorms NXT robots as belonging to the three classes of emotions, “anger”, “joy”, and “sadness”, as used in the first study. In addition, the appearance of one of the robots was changed, in order to test whether a more anthropomorphic appearance of a robot might influence the observers’ judgments. The result of this exploratory study was that the kappa coefficient seems to be an adequate measure to start with, and that anthropomorphic appearance did not influence the identification task.

The third study addressed the problem of finding basic observational units in human-robot interaction using free observation as a heuristic method. This study was carried out during a workshop in connection with the ANIMAX-multimedia theatre at Bonn/Bad Godesberg on LEGO Mindstorms NXT programming for young children. During the teaching of robot programming to the children, video recordings of their behavior and of their interaction with the robots were made. We wanted to study the interaction of the children with robots and their reactions in artistic environments. Our students Julia Wewers and Henrik Niemann analyzed the video material in order to discover some significant behavioral units. The outcome was that the real and interesting interactions between robots and children and the children itself took place only when they were not observed. This indicates that indirect observation should be used in further studies.

In addition to these studies, robot programming is used as a way to introduce students of the liberal arts to computing in media art and cognitive modeling. There are different platforms and curricula for AI courses (Dodds et al. 2006) but only a few ideas outside the engineering domain have been tried in education (e. g. Artbotics (Yang et al. 2007, Martin et al. 2007), the Robot Design Studio (Turbak/Berg 2002) and the Roberta project (Petersen et al. 2007)). We started a course on "Musical Robotics" in 2006. In 2007 and 2008 we integrated robot programming in our courses “Science of Music”, "Embodiment I”, and “Embodiment II”. Our robotics-inspired approach to music research is now part of the curricula for our new Bachelor and Master studies which began in 2007. In 2007 one Master thesis and in 2008 two PhD theses were completed on the topics of robotics, music, and media art since the introduction of robotics, human-robot interaction and embodied cognitive science into musicology. Currently we are transferring the algorithms of the first LEGO Mindstorms NXT project to the more complex Khepera III robot platform.

11. Conclusion

In order to develop a new approach to the scientific study of the musical mind, cognitive musicology has to be complemented by research on human-computer and human-robot interaction. Within the computational approach to mind, interactionism or embodied cognitive science using robots for modeling cognitive and behavioral processes provides an adequate framework for modeling internal processes underlying artistic and aesthetic experiences. The computational framework provided by cognitive science corresponds to the fourth stage in traditional psychological research methodology enabling theory-building and is based on model and recursion theory. The approach of cognitive science to the mind via computational modeling related to psychology may be conceived of as an empirical research strategy resulting from of these theories. This traditional approach of cognitive science focusing mainly on individual internal processing has to be supplemented by a computational approach to the mind taken into account intermental functioning embedded in social environments based on processes of social interaction and the use of cognitive artifacts. Therefore, cognitive musicology has to be supplemented by research on human-computer interaction, especially by research on human-robot interaction, or more generally by research on the interaction of humans with algorithmic agents. For us, new media art environments seem to be the most appropriate place to extend the classical laboratory situation to the study of the relation between intermental and intramental processes in a natural and social environment. We argue that in order to cope with the resulting new research questions, an integrated approach has to be developed. We are trying to develop such an approach using computational modeling of intermental and intramental processing in connection with traditional empirical approaches to data acquisition and analysis from sociology and psychology. Structured observation of human-robot interaction within new media environments seems at present to be the best starting point for prospects of empirical research to achieve an integrated research method, because concerning the accepted methodological standard set by representational measurement theory for psychology and sociology, it can be viewed as measurement. Furthermore, using human observers as measurement devices instead of technical measurement instruments takes into account the complexity of the “stimuli” and “situation” under study.
LEGO Mindstorms robots and structured observation were used in our exploratory studies in order to gain first insights into problems and pitfalls of approaches combining computational modeling from cognitive science and classical empirical research from the

social and psychological sciences in order to study intermental and intramental processing in natural social settings such as artistic human-robot interaction in new media art environments. We hope that our approach might contribute to research on human-computer and human-robot interaction and expect that the development of an integrated methodology might especially contribute to the methodological discussions in the young field of human-robot interaction (Gold et al. 2007).

12. References

Arbib, Michael A. 2004. Programs and Planning. In Gregory, Richard L. (ed.), *The Oxford Companion to the Mind*. 2nd Edition (pp. 759-761). Oxford: Oxford University Press

Bakeman, Roger / Gottman, John M. 1997. Observing Interaction. *An Introduction to Sequential Analysis.* 2nd Edition. Cambridge: Cambridge University Press

Bartneck, Christoph / Okada, Michio 2001. Robotic User Interfaces. *In: Proceedings of the Human and Computer Conference (HC2001),* Aizu, (pp. 130-140)

Bischof, Norbert 1995. *Struktur und Bedeutung*: Eine Einführung in die Systemtheorie für Psychologen zum Selbststudium und für den Gruppenunterricht. Bern: Huber

Burger, Birgitta 2007a. *Musikalischer Ausdruck und Bewegungsgestik*: eine experimental-psychologische Studie zur Kommunikation intendierten Emotionsgehalts durch Roboter. Köln: Universität zu Köln, Magisterarbeit

Burger, Birgitta 2007b. *Communication of Musical Expression from Mobile Robots to Humans:* Recognitions of Music Emotions by Means of Robot Gestures. Stockholm: Master of Science Thesis in Music Acoustics KTH

Burger, Birgitta / Bresin, Roberto 2007. Displaying Expression in Musical Performance by Means of a Mobile Robot. Affective Computing and Intelligent Interaction. *In Proceedings of the 2nd International Conference ACII'07*, Lisbon, Portugal, 12-14 September [Springer LNCS 4738] (pp. 753-754). Berlin: Springer

Clark, Andy 1997. *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press

Coombs, Clyde H. / Dawes, Robin M. / Tversky, Amos 1970. *Mathematical Psychology: An Elementary Introduction.* Englewood Cliffs, NJ: Prentice Hall

Dodds, Zacharia et al. 2006. Components, Curriculum, and Community: Robots and Robotics in Undergraduate *AI Education. AI Magazine*, 27(1), 11-22

Eberlein, G. 1980. Psychologie, Wissenschaftstheoretische Probleme der. In Speck, Josef (ed.), *Handbuch wissenschaftstheoretischer Begriffe*, Bd. 2: G-Q (pp. 523-527). Göttingen: Vandenhoeck & Ruprecht

Fong, Terrence W. / Nourbakhsh, Illah / Dautenhahn, Kerstin 2003. A Survey of Socially Interactive Robots: Concepts, Design, and Applications. *Robots and Autonomous Systems*, 42(3-4), 142-166

Gibson, James J. 1977. *The Theory of Affordances*. In: Shaw, Robert / Bransford, John (eds.), Perceiving, Acting, and Knowing.

Glymour, Clark 2001. *The Mind's Arrow*: Bayes Nets and Graphical Causal Models in Psychology. Cambrigde, MA: MIT Press

Gold, Kevin et al. 2007. *Young Researchers' Views on the Current and Future State of HRI.* HRI'07, March 10–12, 2007, Arlington, Virginia, USA

Greve, Werner / Wentura, Dirk 1997. *Wissenschaftliche Beobachtung* – Eine Einführung. Weinheim: Beltz

Hollan, J. / Hutchins, Edwin / Kirsh, David 2000. *Distributed Cognition*: Toward a New Foundation for Human-Computer Interaction. ACM Transactions on Computer-Human Interaction 7, 174-196

Hutchins, Edwin 1999. Cognitive Artifacts. In Wilson, Robert A. / Keil, Frank C. (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 126-128). Cambridge, MA: MIT Press

Hutchins, Edwin / Hazlehurst, Brian 1995. How to Invent a Lexicon: The Development of Shared Symbols in Interaction. In Gilbert, Nigel / Conte, Rosaria (eds.), *Artificial Societies:* The Computer Simulation of Social Life (pp. 157-189). London: UCL Press

Kim, Jin Hyun 2007. Towards Embodied Musical Machines. In Lischka, Christoph / Sick, Andrea (eds.), *Machines as Agency: Artistic Perspectives* (pp. 18-35). Bielefeld: transcript

Kim, Jin Hyun / Seifert, Uwe 2006. Embodiment: The Body in Algorithmic Sound Generation. Contemporary Music Review: Special Issue "*The Body Technology – Instrument Technology Paradigm*" 25(1-2), (pp. 139-149)

Kim, Jin Hyun / Seifert, Uwe 2007. Embodiment and Agency. Towards an Aesthetics of Interactive Performativity. *In Proceedings of the 4th Sound and Music Computing Conference (SMC'07),* Lefkada, (pp. 230-237)

Leman, Marc / Schneider, Albrecht 1997. Origin and Nature of Cognitive and Systematic Musicology. In Leman, Marc (ed.), *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology* (pp. 13-29). New York: Springer

Lehmann, Günter 1985. *Modell- und rekursionstheoretische Grundlagen psychologischer Theoriebildung*. Berlin: Springer

Luria, Alexander R. 2004. Luria on Reductionism. In Gregory, Richard L. (ed.), *The Oxford Companion to the Mind* (pp. 537-539). Oxford: Oxford University Press

Martin, Fred et al. 2007. Artbotics: *Challenges and Opportunities for Multi-Disciplinary, Community-Based Learning in Computer Science, Robotics, and Art.* 2007 Workshop on Research on Robots for Education at the Robotics Science and Systems Conference, June 30, 2007

Mathews, Max V. 1963. *The Digital Computer as a Musical Instrument. Science* 142(3592), 553-557

Miranda, Eduardo Reck / Wanderley, Marcelo M. 2006. New Interfaces for Musical Expression: *Control and Interaction Beyond the Keyboard*. Middleton: A-R Publications

Petersen, Ulrike et al. 2007. Roberta – Abschlussbericht St. Augustin: *Fraunhofer Institute AIS Autonomous Intelligent Systems*. PDF

Pfeifer, Rolf / Scheier, Christian 1999. *Understanding Intelligence*. Cambridge, MA: MIT Press

Resnick, M., Berg, R., / Eisenberg, M. 2000. Beyond Black Boxes: Bringing Transparency and Aesthetics Back to Scientific Investigation. *The Journal of the Learning Sciences*, 9(1), 17-35

Robson, Colin 2002. Real World Research – *A Resource for Social Scientists and Practioner-Researchers.* 2nd edition. Oxford: Blackwell

Schmidt, Lüder 2005. Towards an "Embodied Cognitive Science of Music": Incorporating Mobile Autonomous Robots Into Musical Interaction. *In Proceedings of the 2nd Asia Pacific Society for the Cognitive Science of Music APSCOM05*, Seoul, (pp. 148-154).

Schmidt, Lüder 2007. Embodied Cognitive Science as Research Paradigm for Music Research. In Lischka, Christoph / Sick, Andrea (eds.), *Machines as Agency: Artistic Perspectives* (pp. 48-65). Bochum: transcript

Schmidt, Lüder 2008. *Embodied Cognitive Science of Music*. Cologne: PhD-Thesis, Institute of Musicology, University of Cologne

Schmidt, Lüder / Seifert, Uwe 2006. Musikalische Robotik: Körperlichkeit und Musik. *Neue Zeitschrift für Musik*, 4, 44-45

Smith, Brian Cantwell 1999. Situatedness/embeddedness. In Wilson, Robert A. / Keil, Frank C. (eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 769-770). Cambridge, MA: MIT Press

Seifert, Uwe 2004. Kognitive Musikwissenschaft: naturalisierte Erforschung musikalischen Geistes. *Journal of Science and Practice of Music*, 18, 177-214

Seifert, Uwe 1991a. Cognitive Science as a New Research Program for Musicology. *Interface*, 21(3-4), 219-238

Seifert, Uwe 1991b. Scientific Musicology as Cognitive Science. Computing in Musicology. *A Directory of Research*, 7, 83-85

Seifert, Uwe 1993. Time Constraints and Cognitive Modelling: Preliminary Remarks on the Methodology of Music Research in Cognitive Musicology. Laaksamo, Jouko / Louhivuori, Jukka (eds.), *Proceedings of the First International Conference on Cognitive Musicology*, 26-29 August 1993, University of Jyväskylä, Finland (pp. 288-299). Jyväskylä: University of Jyväskylä

Seifert, Uwe / Kim, Jin Hyun 2007. Entelechy and Embodiment in (Artistic) Human-Computer Interaction. In Jacko, Julie A. (ed.), *Human-Computer Interaction*, Part I, HCII 2007, LNCS 4550 (pp. 929-938). Berlin/Heidelberg: Springer

Southworth, Nöelle 2006. Music and Robots: Designing Robotic Musical Instruments and Writing Music for Robots in Conjunction with Human Performers, *Master Thesis at the Brown University*

Turbak, Franklin / Berg, Robert 2002. Robotic Design Studio: Exploring the Big Ideas of Engineering In a Liberal Arts Environment. *Journal of Science Education and Technology,* 11(3), 237-253

Wanderley, Marcelo / Battier, Marc (eds.) 2000. *Trends in Gestural Control of Music*. CD-ROM; Paris: IRCAM - Centre Pompidou

Weinberg, Gil / Driscoll, Scott 2006. Toward Robotic Musicianship. *Computer Music Journal* 30(4), 28-45

Weinberg, Gil / Driscoll, Scott 2007. The Robotic Percussionist – Bringing Interactive Computer Music into the Physical World. In Lischka, Christoph / Sick, Andrea (eds.), *Machines as Agency* (pp. 66-82). Bielefeld: transcript

Wertsch, James L. 1999. Vygotsky, Lev Semenovich. In Wilson, Robert A. / Keil, Frank C. (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 878-879). Cambridge, MA: MIT Press

Wirtz, Markus / Caspar, Franz 2002. *Beurteilerübereinstimmung und Beurteilerreliablität – Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriesystemen und Ratingskalen*. Göttingen: Hogrefe

Yanco, H. A. et al. 2007. Artbotics: Combining Art and Robotics to Broaden Participation in Computing." *Proceedings of the AAAI Spring Symposium on Robots and Robot Venues: Resources for AI Education*, Stanford, CA, March

13

Novel Multimodal Interaction for Industrial Design

Bahar Sener, Owain Pedgley
Middle East Technical University
Turkey

1. Introduction

This chapter sets out the need to change the human-computer interaction (HCI) provided in digital tools for carrying out professional industrial design practice. Three-dimensional computer-aided design (3D CAD) has become an indispensable tool for industrial design, being the primary means for modelling and communicating product design proposals. However, a recurring complaint among industrial designers is that 3D CAD is too rooted in engineering design, and is directed towards neither their own creative practices for defining the form of a product (i.e. the activity of 'form creation') nor their underlying need for sketching (Shillito et al., 2003; Hummels, 2000).

The general concern in the literature is that the creatively intense early phase of industrial design, where the form of a product is in a conceptual and 'fluid' state, is very poorly supported. Presently only two systems are marketed as supporting conceptual form creation for industrial design: AliasStudio™ (Autodesk, 2008) and the FreeForm® virtual clay modelling system (SensAble Technologies Inc., 2008). The former makes good use of imported sketch elevation drawings and can be connected to a tablet PC to allow direct freehand drawing. The latter utilises haptic technology (force and kinaesthetic feedback) to harness designers' sense of touch, which is prominent during sketch modelling with workshop media such as Styrofoam® and clay.

It was against this backdrop of dissatisfaction that an empirical research programme was undertaken to identify and address the shortfalls of current 3D CAD systems used by industrial designers. The research programme had the aim of bridging gaps between current 3D CAD packages and envisioned systems specifically devised for industrial design practice. It addressed the research question: in what ways can digital design tools be enhanced or superseded to fit better to industrial designers' needs for conceptual form creation? Of concern was a need to examine in documentary detail what industrial designers liked and disliked about form creation in a variety of modelling media, and then to propose concepts for ways in which computers can – and could – provide improved support. It is worthwhile noting that conceptual design receives little attention in HCI literature, with the balance of research firmly on technical developments and system-specific evaluations. This is most surprising, given that matters of detail become largely academic if underlying or established concepts are found to be inadequate.

The primary motivation for the work was that until improved digital tools are realised, industrial designers will be resigned to adapt to 3D CAD essentially built for other professions (Sener et al., 2002; Hanna & Barber, 2001). This does not seem to be a reasonable or sensible situation. Only through examination at a fundamental operational and conceptual level can the nature of inadequacies with 3D CAD for industrial design be revealed.

2. Enaction, Cognitive Development Theories and CAD

Before introducing the account of the empirical research, it is pertinent to provide a brief historical overview of the interplay of design modelling, human cognitive development theories and the evolution of CAD, up to and including state-of-the-art solutions. Historically, computer support for industrial design commenced with command-line 2D drafting and evolved from the 1980s into multimedia-driven 3D solid and surface modelling systems. It can be said that this transition has been from relatively crude to relatively sophisticated support. This comes as no surprise, since only through technological breakthroughs, which necessarily take time to develop and implement, does the opportunity arise for computer interaction to become more advanced and more sensitive to the needs of users. The idea of sophistication in CAD is an issue that demands closer inspection, and may be usefully illuminated by examining the human cognitive development theories of Piaget (1971) and Bruner (1966).

Piaget's theory involves three stages of cognitive transformation, commencing with sensori-motor representations (from simple reflexes to progressively controlled actions, for achieving effects in the world), through visual manipulations (drawing upon simple single representations to complex multiple representations), to fully matured formal operations (involving cognitive manipulation of complex symbolic systems). Each stage focuses on a new approach rather than an advancement of the preceding approach. Transition through each stage provides gradual clarity and depth of understanding making representations increasingly open to conscious and reflective manipulation (O'Malley and Fraser, 2004). Bruner's theory of intellectual development also involves three stages (enactive, iconic, and symbolic 'modes' or 'mentalities'), signifying transitions from implicit, tacit or sensori-motor representations to gradually more explicit representations (Figure 1).

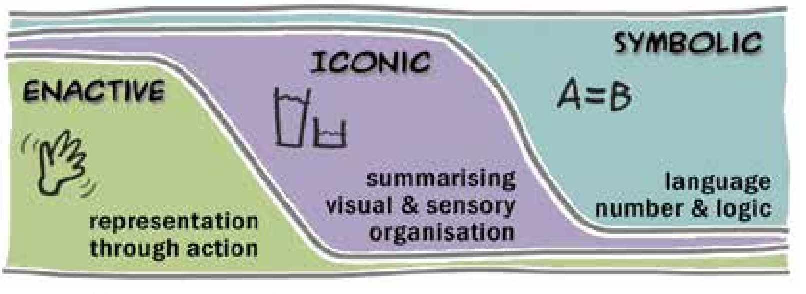


Fig. 1. Transition of learning according to Bruner

Where Piaget's and Bruner's theories differ is that stage transitions are independent of domain or subject area for Piaget, whereas for Bruner each new domain of learning commences with the first stage of intellectual development, and is irrespective of age. The characteristics of Bruner's three cognitive development stages are described in Table 1.

Enactive Mode	Knowledge is generated and demonstrated through domain activities: through learning by doing and the exercising of, and reaction to, motor responses, especially in skilled physical activities requiring dexterity and subtle exercising of, and reactions to, tacit motor responses (e.g. typing, driving a car, dancing, playing a musical instrument, crafting objects).
Iconic Mode	Knowledge about the domain generated in the enactive mode is organised and structured. Knowledge is represented and communicated primarily in the form of images.
Symbolic Mode	Rules are abstracted from the structure and inter-relations of knowledge generated in the iconic mode. Knowledge is represented and communicated as words, mathematical symbols and other notation.

Table 1. Bruner's three modes of cognitive development

Kay, the visionary who created the object oriented software language *Smalltalk*, has offered the insightful slogan 'DOING with IMAGES makes SYMBOLS', as a theoretical underpinning for HCI (Kay, 1987; 1996). The slogan directly relates HCI to Piaget's 'stage model' and Bruner's 'mentalities model' (Figure 2). Kay's slogan implies (as did Bruner) that to be compliant with cognitive development theory, the design of HCI should commence with, and be grounded in, 'DOING with IMAGES', and only then be carried into the more abstract 'makes SYMBOLS'. In reality, the opposite progression has been the case, owing to the technological and conceptual difficulties of creating computer systems that operate in an enactive mode (Verplank, 2003).

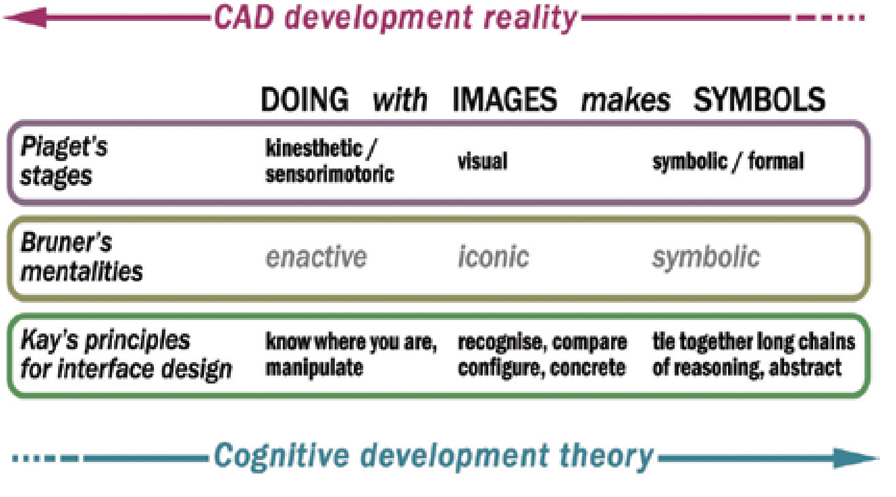


Fig. 2. Human cognitive development versus history of CAD development

So what relevance does the pursuit of enactive interfaces have for computer support for industrial design? The answer is that enactive modes of operation are at the core of industrial designers' designing-and-making and pen-and-paper sketching activities (Sener, 2007). Thus support for enaction can be considered a natural, immediate and intuitive means of HCI for industrial design. Currently only one commercial 3D CAD system is built around an enactive interaction: the aforementioned FreeForm® system. As with other enactive interfaces, FreeForm® delivers multimodal HCI through dual attention to the software and the physical devices that allow access to, and manipulation of, virtual objects (O'Malley and Fraser, 2004; Sener and Pedgley, 2005).

Whilst FreeForm® is a commercially successful product, most enactive interfaces for professional applications are at a pre-commercialisation stage (European Enactive Network of Excellence, 2008). For example, software and hardware developers within research organisations are currently showing how enactive interfaces that integrate a combination of visual technologies (e.g. virtual reality, immersion, holography) and spatial/dynamic technologies (e.g. haptic interaction, tactile interaction, kinaesthetic interaction, gestural sketching) can have application in design disciplines (Bordegoni and Cugini, 2006).

Overall, the implementation of enactive interfaces within 3D CAD is still at a very early stage, largely as a result of technical limitations and the previously mentioned legacy in which HCI for 3D CAD has developed from a starting point of text-based instructions. FreeForm®, for example, still relies on considerable CAD-like command interaction (i.e. menus, keyboard input, mouse actions) to accompany its haptic capabilities.

3. Research Methodology

The kinds of structure and interaction that digital industrial design tools will possess in the future is not a subject frequently visited in the literature. Furthermore, reports of empirical research into industrial designers' form creation activities are also relatively sparse. As a general observation, previous studies have failed to examine in detail the comparative strengths and weaknesses of the various modelling media designers use, instead concentrating in the main on modelling technique and good practice – often for just one modelling medium. This situation is not so surprising, since to generate empirical evidence that exposes designers' general form creation needs, rather than their perceived needs or scattered and generalised anecdotal evidence, requires considerable effort. Generation of this evidence base across modelling media was deemed a vital first step for ensuring the credibility of any new HCI concepts for 3D CAD.

The research programme therefore commenced with a substantial documentary study of industrial designers' form creation activities with two established media (Styrofoam®, conventional CAD) and one state-of-the-art medium (FreeForm®). The study comprised a series of 40 design and modelling experiment sessions, each lasting approximately two hours, conducted with a total of 16 UK-based participants. The participants were split into two groups of 8 participants each. Group 1 spanned employed, freelance, university staff and postgraduate industrial designers. Group 1 participants were involved in experiments covering all three media (24 sessions in total). Group 2 comprised solely industrial design undergraduates, who were involved in experiments covering only Styrofoam® and conventional CAD (16 sessions in total). A variety of 3D CAD systems were used (3DStudio Max, AutoCAD, I-DEAS, Lightwave 3D, Mechanical Desktop, Pro/Engineer, Rhinoceros, SolidWorks) – these being the participants' preferred and familiar systems.

Each session involved a participant creating the form for a small-sized household item (perfume container or salt and pepper shaker) or consumer electrical product (computer mouse or computer speakers) using just one modelling medium. To limit any order effect, the participants worked on different products for each modelling medium, thereby reducing the likelihood of transferring experiences from one session to another. In addition, each participant completed a mood adjective checklist before and after each session (Mackay et al., 1978). The evaluation of the checklist revealed no experiment conditioning and no order effects between the various combinations of sessions and modelling media. The participants were free to use pen and paper sketching during the sessions.

Real-time data were generated through video recording and researcher observation, and were accompanied by questionnaires completed by the participants during the sessions, which gauged the participants' first-hand experiences of modelling. A review interview was held at the end of each session to provide an opportunity to clarify any element of the generated data, and to allow participants an opportunity to volunteer further insights into their modelling. As far as is known, the experiments comprised the first major comparative study of 3D sketch modelling in industrial design.

A full account of the data collection, code-based processing and analysis would be too lengthy for inclusion here, so readers are referred to its primary documentation (Sener, 2004). The data originating from the questionnaires and review interviews comprised approximately 800 individual statements on the strengths and weaknesses of modelling with Styrofoam®, conventional CAD and FreeForm®. A hybrid strengths – weaknesses – opportunities – threats (SWOT) analysis, based on guidelines by Ulrich & Eppinger (1995), was followed to translate the strength and weakness statements into a set of customer need statements for improved digital industrial design tools. Briefly, this involved collating and consolidating the collective strengths across the three modelling media and redressing (i.e. reversing the expressed negativity) of the collective weaknesses. Figure 3 contains the definitive set of customer need statements. The processing procedure determined a priority position for each customer need statement, so that Figure 3 presents the statements in priority order from the especially important at the head to the moderately important at the foot. The terms 'quick', 'easy' and 'good' were merged during the data processing because participants used them interchangeably.

Figure 3 can be regarded as an explicit guide to desirable specifications for digital industrial design tools and their associated HCI. Four key themes can be identified.

Bulk/sketch form creation. The highest priority customer need statement – quick/easy/good basic form creation – refers to sketching of product form in a proportionally correct and simplified manner free of constraints and dimensions.

Control of form creation. Seven customer need statements pointed to controlled form creation: constrainable tools, constrained form creation, precise, quick/easy/good detailing, quick/easy/good attribute control, and quick/easy/good uniform surface finish/texture.

Ease of form creation. Six customer need statements pointed to minimal effort and removing obstacles in creating form: user-friendly interfaces, useful variety of modelling tools, high proficiency with minimal practice, form creation guidance, form construction aids, and comfortable input devices.

Life-like form creation. Five customer need statements pointed to replicating the multimodal sensory experience of creating physical models: life-like model appearance, life-like tool/material contact, model interaction with hands, haptic feedback, and tools analogous to workshop tools.

#	Customer need
1	quick/easy/good basic form creation
2	model interaction with hands
3	precise
4	quick/easy/good uniform surface
5	haptic feedback
6	quick/easy/good form detailing
7	quick/easy/good form attribute control
8	unlimited model viewpoints
9	quick/easy/good curvy form creation
10	quick/easy/good shape/form modification
11	life-like model appearance
12	life-like tool/material contact
13	reversible
14	user-friendly interface
15	form creation guidance
16	useful variety of modelling tools
17	form construction aids
18	comfortable input devices
19	high proficiency with minimal practice
20	quick/easy/good form visualisation
21	constrainable tools
22	multiple/interchangeable input devices
23	reasonable workload on left & right hands
24	suitable for product design
25	tools analogous to workshop tools
26	quick/easy/good model export
27	manufacturable output
28	size/proportion feedback
29	quick/easy/good model spatial arrangement
30	constrained form creation

Fig. 3. Prioritised customer needs for improved digital industrial design tools

The first two of these themes are clearly in tension. Thus, the results showed that the absence of convincing digital sketching provision, and in particular a lack of marriage between sketch form creation and constrained form creation, in both 2D and 3D modelling environments, is a major issue to be redressed in the design of 3D CAD systems for industrial designers. The combination remains elusive in currently available systems.

4. HCI Design

4.1 Ideation

As is normal, the customer need statements fall short of providing *ideas* for tangible HCI solutions: they point to issues to be resolved, but not how they can be resolved. It is only through creative input – designing – that customer needs can be acted upon and design ideas proposed. It was therefore essential to integrate a design project into the research programme so that the customer need statements could be translated into envisaged new 3D CAD systems. This was achieved during a period of practice-led design research (Arts and Humanities Research Council, 2007; Pedgley & Wormald, 2007).

The designing was undertaken by the first author, a trained industrial designer, who drew upon a variety of specific experiences and sources of information to assist in ideation. The major inputs were: (i) the set of 30 customer needs in Figure 3, (ii) the prior art reviews concerning the history of 3D CAD development, and (iii) a four month industrial placement at Procter & Gamble Technical Centres UK. During the placement, the author was employed to design, model, and prototype new consumer goods using the FreeForm® system. As a practising designer in Procter & Gamble, she had significant professional authority for her investigations of how 3D CAD was perceived and used by other designers. Her involvement in the day-to-day business of the company's NPD programmes provided evidence of the uses of 3D CAD that otherwise would have been impossible to obtain. It also allowed for an effectiveness evaluation of 3D CAD driven by haptic interfaces, within a commercial context.

Other inspiration to aid the designing came from wider reading in the crossovers between communication technologies and contemporary product design, as well as personal experiences of 3D sketch modelling in a variety of media. Specific sources included Philips' *Vision of the Future* (Philips Design, 1996) and several Hollywood sci-fi movies including *Minority Report*, *The Matrix Trilogy*, and *The World Is Not Enough*.

4.2 User Participation

The purpose of involving target end users in the design project was to share ideas, create synergies and generally enrich the design activity and outcomes. Six participants were recruited from amongst the pool of eight within group 1 of the design and modelling experiments. All of the participants therefore possessed a heightened awareness of the research aims and had first-hand experience of state-of-the-art modelling through their FreeForm® sessions.

The first author assumed the role of facilitator and note taker during two three-hour sessions with users. The first session focused on generating individual ideas for new kinds of form creation tools and environments. It was explicitly stated that acceptable ideas could be either incremental improvements to existing technologies or 'future-gazing' solutions. The set of customer need statements was provided as a stimulus. The second session employed scenario building (Hasdogan, 1997) to elicit ideas on how individual ideas could be combined.

The participants communicated their proposals through A2 sketch sheets and verbal reports. Following the sessions, the proposals were examined for common features, which were then visually or verbally grouped under keyword headings. The proposals showed a general desire for designers to work within more dedicated, customisable surroundings enhanced by digital technologies.

4.3 Concept Development

Concept development continued as a solo effort by the first author. Two strands identified from the user participation sessions were adopted to aid the process: 'workspace concepts' (broadly referring to digitally enhanced environments for industrial designers to work in) and 'form creation concepts' (broadly referring to new HCI for industrial designers to digitally define product form). The separation is acknowledged to be forced, particularly since with technologies such as immersive virtual reality (VR), as far as users are concerned the 'workspace' becomes almost imperceptible and the 'form creation' dominates. The separation was nonetheless helpful in directing the emerging design ideas and communicating the final concepts.

A balance was sought between concepts that were immediately realisable (and that suggested incremental improvements to, and combinations of, existing technologies), and those that would require technology to advance. Special attention was paid towards developing concepts that offered plausible new routes for digitally sketching product form, especially through multimodal interactions. In all cases, the concepts were required to be coherent in the sense that they combined individual ideas into a convincing system. It was an explicit objective to satisfy the highest number of customer need statements as practically possible, although readers will appreciate that not all customer needs could be satisfied within a single concept without that concept becoming too incoherent. Matrices were used to check the compliance of each concept against the 30 customer need statements.

A portfolio containing eleven individual concepts was prepared: four workspace concepts and seven form creation concepts. A name was assigned to each concept and a text description of the essence of the concept was written. The concepts were prepared as presentation-quality colour illustrations in a purposefully 'cartoon style'. The style choice was important: it was chosen to promote flexibility in interpretation, rather than finality in specification that would accompany a photorealistic rendering or physical mock-up.

5. Concept Portfolio

5.1 Workspace Concepts

WC1 '*Desktop Computing*' enhances the sensory information experienced by designers within a contemporary desk environment, utilising multiple and interchangeable input devices connected to large flat screen displays, including haptic devices and stereovision glasses.

The idea behind WC2 '*Immersive Room*' is to set an immersive theme and mood within a dedicated collaborative workroom, in a similar way to desktop themes and wallpaper in Microsoft Windows or Mac OS. The environment can be instantaneously switched from project to project, with full-scale projections of, for example, CAD software, moodboards, competitor products and products in use. The workspace is used in conjunction with wireless tablet PCs and optional haptic gloves and stereovision glasses. Designers are free to sit or stand.

The technology embodied in WC3 '*Intelligent Screens*' allows designers to move away from desktop cubicles towards open-plan environments that facilitate collaboration and sharing of information. Touch screens and finger-based haptic devices allowing two-handed interaction are prominent. Programmable finger thimbles are used to perform various functions with the touch screens. To aid collaboration, design updates can be wirelessly streamed between designers, and the screens can operate in either single-sided mode (opaque) or double-sided mode (transparent) to further strengthen collaboration. Designers are free to sit or stand and can communicate via an audio link.

The idea behind WC4 '*Advanced Wireless Virtual Reality*' is a fully programmable VR system based on non-invasive technology that is wireless, miniaturised and lightweight. It represents a technological progression of current VR and haptic applications within a dedicated workroom. Programmable VR software is customised to designers' preferences and is coupled to multiple and interchangeable peripherals (e.g. motion-trackers, wireless haptic fingertip sensors, stereovision headset, foot control sensors, voice command receivers). The concept allows two-handed interaction, full-scale model evaluation, and collaborative working through shared VR information. Figure 4 shows WC1 to WC4.

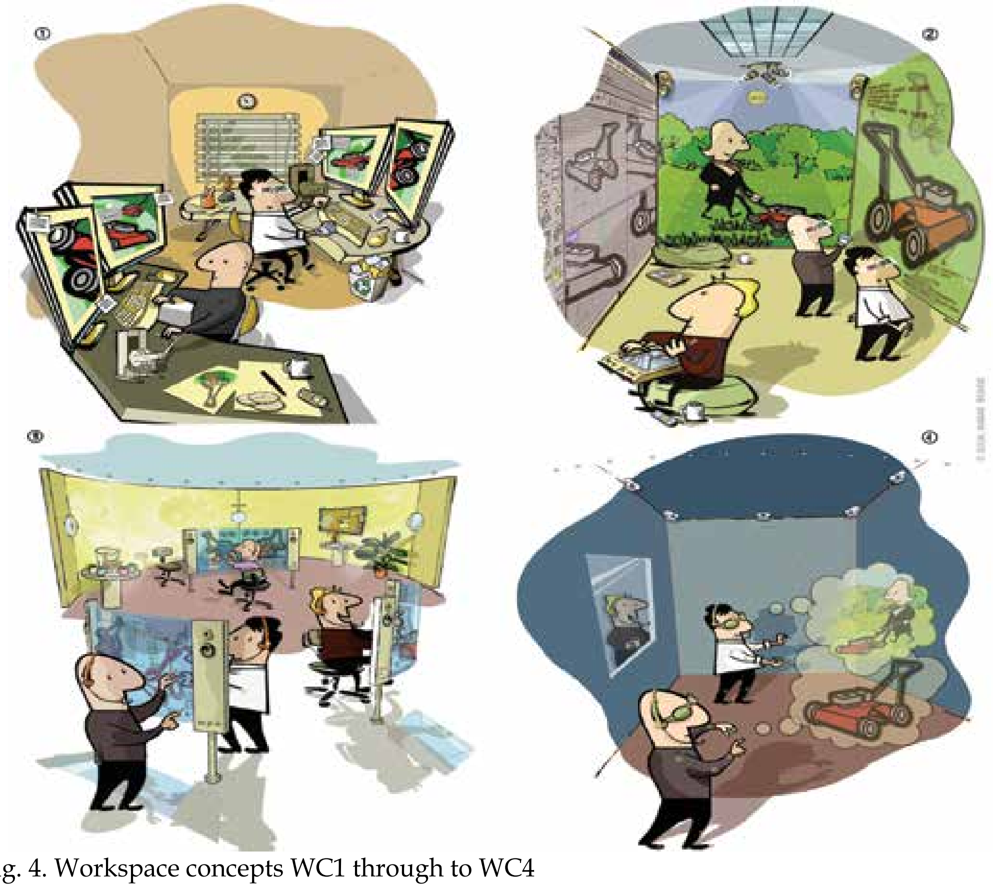


Fig. 4. Workspace concepts WC1 through to WC4

5.2 Form Creation Concepts

Two-handed haptic feedback is the essential feature of FC1 '*Dual Co-Located Haptic Devices*', which combines elements of the FreeForm® system with bespoke hardware manufactured by Reachin Technologies AB (2008). The latter comprises a monitor that displays a CAD model stereoscopically and 'in mirror image' so that when viewed on a reflective screen the model appears convincingly in 3D and in correct orientation. Thus, the on-screen cursor (modelling tool) controlled by the FreeForm® input device (Phantom®) is co-located (hand-eye coordinated) with the physical nib position of the Phantom®. This concept specifically addresses the absence of two-handed control and realistic movement within present haptic systems. One hand is intended to grasp the model (using a haptic glove), whilst the other is intended to shape the material (using the Phantom®). The palette of modelling tools could be for any virtual material, although Styrofoam® and clay are most suited to industrial design. Tools ranging from delicate hand tools to large machine tools would be recreated digitally.

The intention behind FC2 '*Smart Material*' is to make form creation with digital tools as interactive and spatially unconstrained as the manual shaping of workshop materials. It relies on the use of malleable material impregnated with particles that can be continuously position-detected in 3D space, allowing a digital equivalent of designing-and-making.

FC3 '*Haptic Holographic Representation*' uses a form of non-physical rapid prototyping, allowing visual, haptic and 'walk around' evaluation of an emerging product form projected holographically from a pod. It is an entirely waste-free and instantaneous system, independent of modelling software, and is intended to assist form modifications in real-time and promote collaborative product evaluation between remote sites. Optional stereovision glasses and haptic finger thimbles are used to allow enhanced multimodal evaluation.

The premise behind FC4 '*Sequential Scanning*' is that organic and texturised forms are easy to create in non-digital media (e.g. Styrofoam®, clay). The concept builds upon this and includes intelligent reverse engineering software to automatically create high-quality editable surface models (i.e. constructed from splines, arcs, circles, lines etc.) from point cloud scan data of pre-modelled forms.

FC5 '*Squidgy Sponge*' is a highly interactive wireless input device that can be manipulated and deformed in 3D, with the resulting deformations mapped onto selected areas of a digital model in real-time. The device can be twisted, indented, squeezed, tapered, stretched, squashed, folded etc. The device can also be deformed by pressing a physical object into it.

FC6 '*Verbal/Gestural Input*' extends the application of gestural sketching (Hummels, 2000), in which the movement of one's hands, arms or head becomes a tool for sketching, and in so doing overcomes spatial and functional limitations of 2D (planar) movement associated with pen and paper sketching. At its heart is personal expression, allowing designers to 'act out' and 'talk through' their ideas for product form. The system uses motion trackers and microphones to capture input data, whilst stereovision glasses may be optionally worn.

The familiarity of paper-based sketching is harnessed in FC7 '*Automated 2D-to-3D Translation*' and augmented by intelligent software to create 'clean' model geometry and a correspondingly high-quality surface model. The software shows in real-time how a product sketched in 2D elevations on a tablet PC will appear as a 3D form. This concept takes influences from sketch mapping (Tovey, 2002) and proven methods of 2D-to-3D translation (Igarashi & Hughes, 2003) and represents an attempt to harness and surpass functional and qualitative aspects of paper-based sketching. It is intended to relieve designers of the relatively mundane task of 3D CAD geometry construction. Figure 5 shows FC1 to FC7.

Fig. 5. Form creation concepts FC1 through to FC7

6. Evaluation Method

The evaluation of the concepts was carried out through a questionnaire distributed to ten participants: the same eight participants of group 1 in the design and modelling experiments, and two additional staff industrial designers at Loughborough University. A 100% return rate was achieved. The participants were chosen because collectively they represented an 'elite group' of especially well informed designers, having had significant prior involvement in the research. They had also demonstrated proficiency in Styrofoam® and 3D CAD and had practical experience with FreeForm®. The continued involvement of the same participants was viewed positively and was expected to lead to particularly critical evaluation of the concepts.

A questionnaire was chosen over individual interviews to allow the participants to pace themselves during their evaluations and to create a standardised set of data (Jorgensen, 1989). The overall aim was to identify the most favoured and least favoured concepts, and to identify the features and characters of those concepts that led to their particularly supportive or unsupportive evaluation. Each concept was requested to be evaluated individually against three principal criteria: *enjoyment*, *inspiration* and *assistance*. When combined, these three criteria were intended to create a good assessment of the overall desirability of the concepts as measured by long-term use (enjoyment), stimulation for design ideas (inspiration) and utilitarian benefits (assistance). Figure 6 shows the questionnaire template and the use of Likert scale statements to elicit participants' reactions.

WORKSPACE CONCEPT 1 : DESKTOP COMPUTING

"I would find this WORKSPACE **enjoyable** to work in."			
Strongly Agree ☐	Tend to Agree ☐	Tend to Disagree ☐	Strongly Disagree ☐
"This WORKSPACE would be **inspirational** to my designing."			
Strongly Agree ☐	Tend to Agree ☐	Tend to Disagree ☐	Strongly Disagree ☐
"This WORKSPACE would **assist** my designing."			
Strongly Agree ☐	Tend to Agree ☐	Tend to Disagree ☐	Strongly Disagree ☐
"I would **prefer** this WORKSPACE to my current workspace."			
Strongly Agree ☐	Tend to Agree ☐	Tend to Disagree ☐	Strongly Disagree ☐
Summarise your reasons for agreeing:		Summarise your reasons for disagreeing:	

Fig. 6. Questionnaire template

The Likert scale deliberately lacked a neutral response to encourage the participants to express an opinion. The fourth statement, concerning preferential use, was added to directly assess participants' acceptance of change and overall impressions. This 'preference data' would allow a +ve/-ve correlation to be established against the researcher-constructed 'overall data' (comprising a summation of enjoyment, inspiration and assistance data) and would therefore act as a methodological test. A short summary of the participants' reasons for agreeing or disagreeing with the questionnaire statements was also requested.

A briefing session was held prior to delivery of the questionnaire, to remind the participants of the purposes of the work and the specific aims of the concept evaluation. Written instructions on how to complete the questionnaire were provided. The concepts were presented within a ring-bound portfolio, containing the concept illustrations and text descriptions. A time limit of two hours was set to view the portfolio and complete the questionnaire.

6.1 Data Analysis Procedure

The data were analysed by assigning numerical scores to each of the Likert scale grades, so that a quantitative measure of success for each concept could be calculated (Brace, 2004). The data were scored as follows: strongly agree (+2), tend to agree (+1), tend to disagree (-1) and strongly disagree (-2). The score range per criterion was ±20 (±2 maximum/minimum score, 10 participants). The overall score range per concept was ±60 (±20 per criterion, 3 criteria). To aid comparisons and discussion, all data were converted to percentage of score range, creating the following categories.

$x \geq +50\%$ = participants strongly agreed
$+50\% > x > 0\%$ = participants tended to agree
$-50\% > x > 0\%$ = participants tended to disagree
$x \geq -50\%$ = participants strongly disagreed

The participants' comments regarding their agreement or disagreement with the questionnaire statements were logged verbatim. Keywords were extracted from the comments to develop a deeper understanding of the successes and failures of each concept.

7. Results

Figures 7, 8 and 9 show the results of the individual evaluations for enjoyment, inspiration and assistance. Figure 10 presents the results of the overall evaluation, as a summation of the individual evaluations, whilst Figure 11 presents the results of the preference evaluation.

The first general observation to note is that the concepts scored very highly for enjoyment (mean = +61%) and assistance (mean = +48%), and reasonably well for inspiration (mean = +32%). Negativity towards any of the concepts was extremely isolated, occurring in only 2 out of a possible 33 evaluations (11 concepts, 3 individual evaluations): WC1 (-20% for inspiration) and FC6 (-5% for preference).

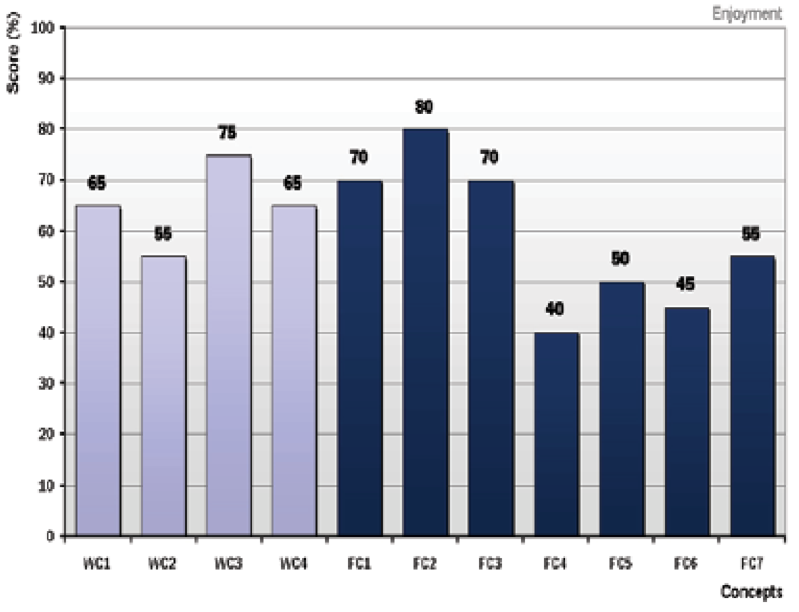


Fig. 7. Results – enjoyment evaluation

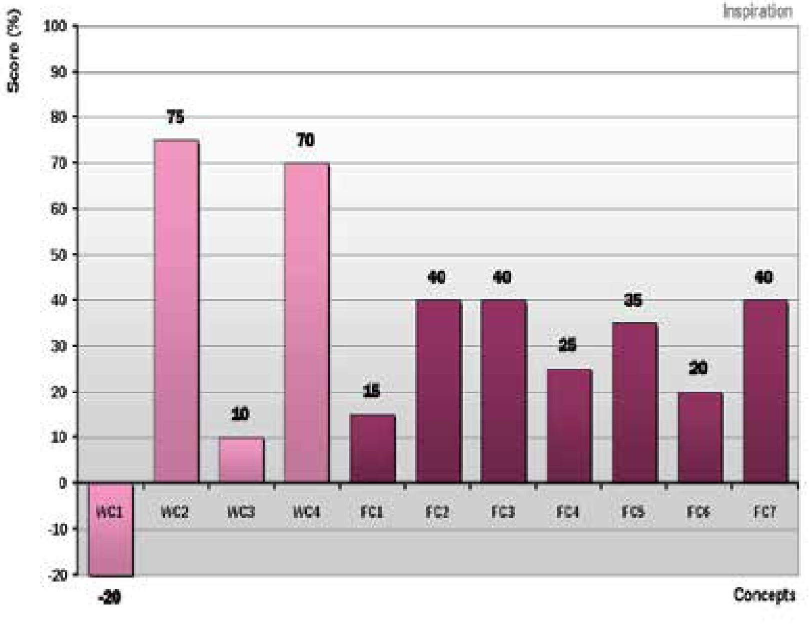


Fig. 8. Results – inspiration evaluation

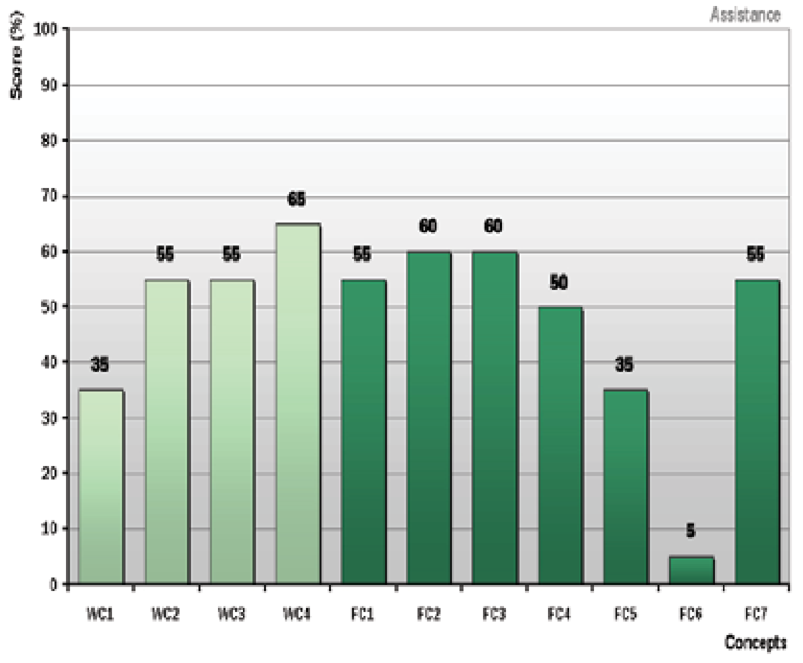


Fig. 9. Results – assistance evaluation

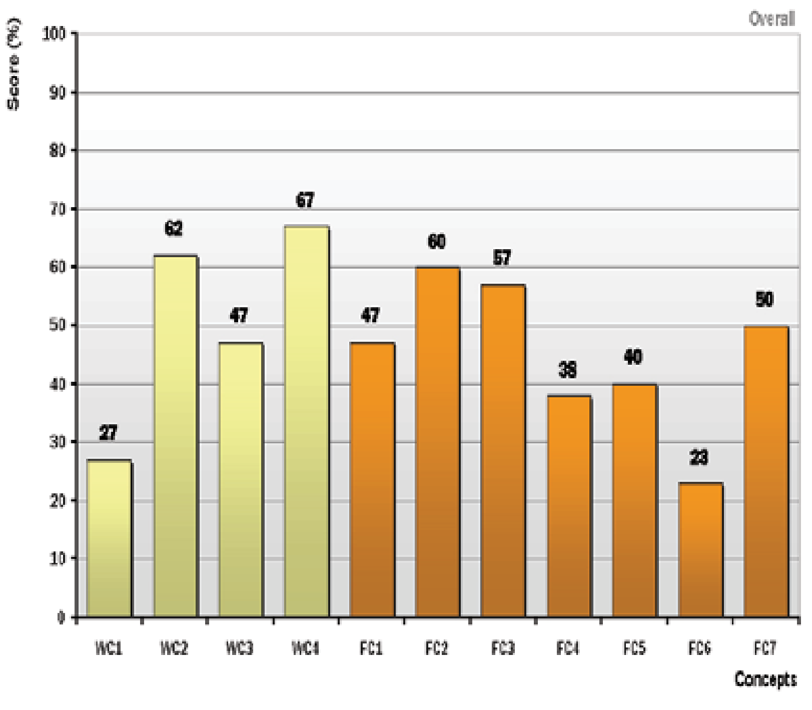


Fig. 10. Results – overall evaluation

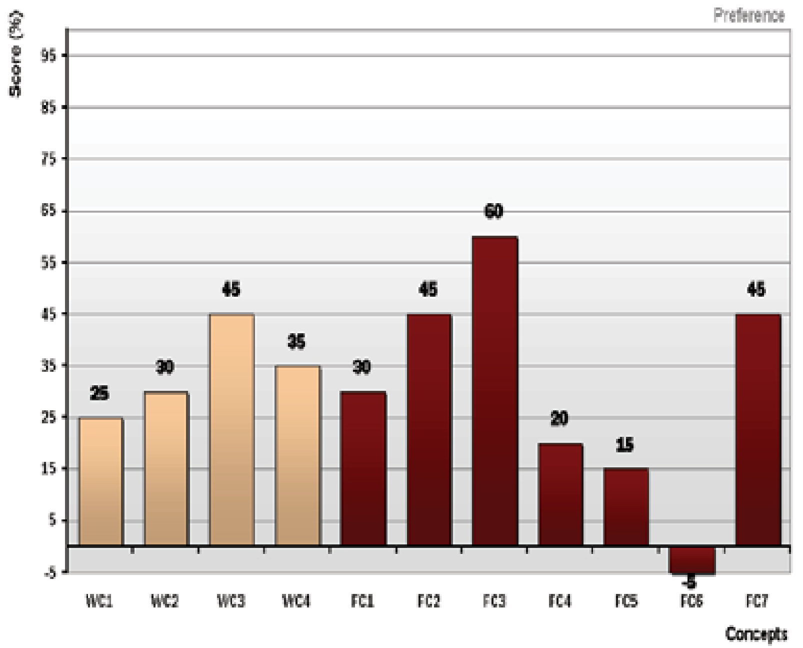


Fig. 11. Results – preference evaluation

7.1 Enjoyment Evaluation

Participants *strongly agreed* that all the concepts would be enjoyable to use, except FC4 (+40%) and FC6 (+45%), which they only *tended to agree* would be enjoyable to use. The participants' comments revealed their relative lack of enthusiasm towards FC4 was because the concept was seen as too procedural and slower than alternative methods of digital form creation. For FC6, the participants raised concerns over modelling accuracy, difficulties in describing complex forms and intricate details, usability, and anxiety about talking aloud and making gestures.

7.2 Inspiration Evaluation

In general, the participants *tended to agree* that the presented concepts would be inspirational to their design practice. However, WC2 (+75%) and WC4 (+70%) were found to be *strongly* inspirational whereas participants *tended to disagree* that WC1 (-20%) was inspirational. Many aspects of WC2 were praised by the participants: the immersive approach, project interchangeability, concentration and variety of information, high levels of communication, opportunities for teamwork and the ability to visualise ideas full-scale. Similarly high praise was given to WC4, with participants keen on its advanced interactive visualisation, its versatility, the ability to visualise ideas full-scale, the general scope of sensory information that it affords, and its facility for upstream virtual product testing. The participants' negativity towards WC1 was shown in comments referring to the normality of a modern-day office, poor interaction between people, few external stimuli and a confined workspace.

7.3 Assistance Evaluation

In general, participants *strongly agreed* that the concepts would be assistive to their design practice. However, three concepts fell within the *tended to agree* category: WC1 (+35%), FC5 (+35%) and FC6 (+5%). The relative lack of enthusiasm for WC1 and FC6 was accounted for in sections 5.1 and 5.2. Comments on FC5 showed the participants to be concerned about accuracy, control, realisation of form details, difficulties in achieving organic forms, and its limitation as a purely deformation-making tool.

7.4 Overall Evaluation

Five concepts received overall scores ≥ +50%: WC2 (+62%), WC4 (+67%), FC2 (+60%), FC3 (+57%), and FC7 (+50%). These five concepts represent the participants' most favoured potential uses of digital technologies for product form creation. The scores for the remaining six concepts ranged from +23% to +47%, indicating that participants possessed overall support for all eleven concepts in the portfolio, with none of the concepts having overall rejection.

7.5 Preference Evaluation

Participants' direct preference data (Figure 11) provided an opportunity for comparison and corroboration with the researcher-derived summed overall score combining enjoyment, inspiration and assistance (Figure 10). The results showed that the *rank order* of the participants' direct preference scores correlated well with that of the researcher-derived overall scores, although some differences existed in the score values, which will be examined shortly. On the whole, the combination of enjoyment, inspiration and assistance criteria successfully indicated designers' willingness to change from current digital modelling systems to new ones. Their adoption as evaluation criteria was therefore methodologically vindicated.

The participants *tended to agree* that the concepts were preferable to their present systems (mean = +31%). Exceptions to this were: FC3, which was considered *strongly* preferable (+60%), and FC6, which was *not* considered preferable (-5%). The participants expressed a strong preference for FC3 because of its full-scale visualisation capabilities, the ability to walk around a projected product and view it from unlimited viewpoints, its 3D sensory feedback, and the attractiveness of appending it to existing CAD systems. The negativity towards FC6 echoed the comments reported previously.

With regard to the workspace concepts, WC1 was rated the least popular under both evaluations and received consistent scores (+27% overall, +25% preference). WC3 was ranked differently under the evaluations (third for overall, first for preference), although it received a consistent score of +47% and +45% respectively. Some inconsistencies were shown for WC2 and WC4. Although both of these concepts received similar rankings under both evaluations (WC2 either second or third; WC4 either first or second), the scores under the two evaluations differed (WC2 +62% overall, +30% preference; WC4 +67% overall, +35% preference). The participants were therefore considerably less enthusiastic about adopting WC2 or WC4 in preference to their current systems. This may be because WC2 and WC4 are technologically quite advanced from current systems and generate some scepticism over their likely success of implementation, despite acknowledged conceptual benefits.

Cross-comparisons were also made for the form creation concepts. FC1, FC4, FC5 and FC6 were ranked as the lowest four concepts under both evaluations, with FC1 consistently fourth least popular and FC6 consistently the most least popular. The rank order of FC4 and FC5 swapped between the two evaluations. A comparison of the scores received for these four concepts revealed that the preference evaluation was consistently less favourable than the overall evaluation, indicating that despite acknowledging individual benefits within these lowest ranked concepts, the participants were not convinced that overall they would be preferable to their current systems.

In contrast, FC2, FC3 and FC7 were the three highest ranked concepts under both evaluations. FC7 was consistently ranked third, whilst the rank order of FC2 and FC3 swapped between the two evaluations. The scores between the two evaluations of FC2, FC3 and FC7 were reasonably close (FC2 +60% overall, +45% preference; FC3 +57% overall, +60% preference; FC7 +50% overall, +45% preference), showing that the participants considered these concepts to be strong, whether assessed as a whole or analysed against individual criteria.

8. Technological Implications

Without doubt, the quality of haptic feedback offered by enactive HCI will need to dramatically improve if digital modelling experiences are to become convincing reproductions of designing-and-making and pen-and-paper sketching performed in the physical world. The most valuable technological advances will be those that make it possible to grasp models, to have two-handed control of modelling tools, and that provide a facility to rub one's fingertips and palms across model surfaces to evaluate and adjust for ergonomics, aesthetics and other matters of fitness of form. Haptic devices that are less invasive (e.g. smaller, less heavy) and that have multipoint sensors (e.g. on fingers, thumbs and palms) will be necessary to create more authentic modelling experiences.

Furthermore, any new system should be based on surface modelling technology (e.g. NURBS: non-uniform rational b-spline surfaces), rather than polygon mesh models, to maximise usefulness in downstream manufacturing and analysis applications.

9. Conclusions

The chapter has made a case for industrial designers to be served with specialised 3D CAD systems. The thrust of the argument is that a conceptual shift in HCI must take place if industrial design is to be supported by digital tools that properly satisfy industrial designers' needs for sketching and developing product forms. The research demonstrated how current 3D CAD systems fail to fully support these needs, and that recent technological developments in HCI for 3D CAD do not yet offer a fully satisfactory resolution.

Eleven concepts for 3D CAD specialised for industrial design were generated. The concepts were subjected to a first-stage evaluation by expert users, whose assessment was based on various criteria attributable to the HCI inherent in the concepts. Five concepts were evaluated as especially desirable to users: WC2 '*Immersive Room*', WC4 '*Advanced Wireless Virtual Reality*', FC2 '*Smart Material*', FC3 '*Haptic Holographic Representation*' and FC7

'*Automated 2D-to-3D Translation*'. Overall, users were found to favour HCI providing naturalistic, spontaneous and expressive tools for sketch form creation, specifically away from the paradigm of conventional desktop CAD.

For workspaces, users showed most enthusiasm towards dedicated and customisable workrooms, where an immersive environment can be set and switched seamlessly from project to project. For form creation tools, users showed most enthusiasm towards what may be termed 'virtual workshops' (digital emulations of existent skills in modelling with physical materials) and 'intelligent environments' (supplementing cognitive modelling skills – mental imaging – through assistive digital visualisation, specifically away from the command-led interactions of conventional 3D CAD).

The results of the research justify initiatives for developing prototype and pre-commercial systems for new digital industrial design tools, and for creating R&D collaborations between specialist HCI and industrial design communities. The next stage for this work is to cooperate with experts in human and computer sciences to develop the favoured concepts to a prototype stage, so that a second-stage evaluation may be performed with a larger and more general group of industrial designers. It will be important to use multimedia techniques and mock-ups to communicate the essence of the concepts in a manner that is more advanced than 2D illustrations and text descriptions. The findings of the second-stage evaluation will be valuable for finalising directions for new commercial systems.

10. Acknowledgements

This chapter builds upon an article originally published in *Design Studies*, entitled 'User evaluation of HCI concepts for defining product form'.

11. References

Arts and Humanities Research Council (2007). *Research Review: Practice Led Research in Art, Design and Architecture.* Available at: http://www.ahrc.ac.uk/apply/research/sfi/ahrcsi/research_reviews.asp (accessed 10 May 2008)

Autodesk (2008). Available at: http://www.autodesk.com (accessed 10 May 2008)

Bordegoni, M. & Cugini, U. (2006) 'Haptic modelling in the conceptual phases of product design', *Virtual Reality*, Vol. 9, pp. 192–202.

Brace, I. (2004). *Questionnaire Design,* Kogan Page Ltd., London

Bruner, J. (1966). *Towards a theory of instruction*, Harvard, Cambridge

European Enactive Network of Excellence (2008). Available at: http://www.enactivenetwork.org (accessed 10 May 2008)

Hanna, R. & Barber, T. (2001). An inquiry into computers in design: attitudes before–attitudes after *Design Studies,* Vol. 22, No. 3, pp. 255-281

Hasdogan, G. (1997). Scenario building as part of the design process, in McGrory, P (ed) *The Challenge of Complexity* Helsinki, University of Arts and Design, pp. 134-141

Hummels, C. (2000). *Gestural Design Tools: Prototypes, Experiments and Scenarios,* University of Technology Publications, Delft

Igarashi, T. & Hughes, J. (2003). Clothing manipulation, *Proceedings of ACM SIGGRAPH 2003*, UIST and I3D reprise session, San Diego, p. 67

Jorgensen, D. (1989). *Participant Observation – A Methodology for Human Studies,* Sage Publications, Newbury Park

Kay, A. (1987). *Doing with images makes symbols: communicating with computers*, Apple Computer Inc., Cupertino

Kay, A. (1996). 'The early history of SmallTalk' in *Programming languages*, eds T. Bergin and R. Gibson, ACM Press, New York

Mackay, C.; Cox, T.; Burrows, G. & Lazzerini, T. (1978). An inventory for the measurement of self-reported stress and arousal *Journal of Social and Clinical Psychology,* Vol. 17, pp. 283-284

O'Malley, C. & Fraser, D. (2004). *Literature review in learning with tangible technologies*, a report for NESTA Futurelab, Available at: http://www.futurelab.org.uk/resources/publications_reports_articles/literature_reviews, (accessed 10 May 2008)

Pedgley, O. & Wormald, P. (2007). Integration of Design Projects within a PhD. *Design Issues*. Vol. 23, No. 3, pp. 70-85

Philips Design (1996). *Vision of the Future*, V&K Publishing, Blaricum

Piaget, J. (1971). *The psychology of intelligence*, Routledge and Kegan, Boston

Reachin Technologies A.B. (2008). Available at: http://www.reachin.se (accessed 10 May 2008)

Sener, B. (2004). *Enhancing the Form Creation Capabilities of Digital Industrial Design Tools,* PhD thesis, Loughborough University, UK

Sener, B. (2007). Rethinking digital industrial design: a mandate for virtual workshops and intelligent environments, *Digital Creativity*, Vol.18, No.4, pp. 193-206

Sener, B. & Pedgley, O. (2005). 'Next-generation digital industrial design tools driven by enactive interfaces', *Proceedings of Enactive-05 Second International Conference on Enactive Interfaces*, Enactive Network, Genoa, Italy [CD-ROM]

Sener, B.; Wormald, P. & Campbell, I. (2002). Evaluating a haptic modelling system with industrial designers, in Wall, S, Riedel, B, Crossan, A and McGee, M (eds) *Proceedings of the 2002 EuroHaptics International Conference* Edinburgh, University of Edinburgh, pp 165-170

SensAble Technologies Inc. (2008). Available at: http://www.sensable.com (accessed 10 May 2008)

Shillito, A.; Scali, S. & Wright, M. (2003). Haptics: for a more experiential quality in a computer interface, *Proceedings of the 5th European Academy of Design Conference,* 28-30 April 2003, Barcelona, available at: http://www.ub.es/5ead/PDF/9/ShillitoSW.pdf (accessed 10 May 2008)

Tovey, M (2002). Concept design CAD for the automotive industry, *Journal of Engineering Design,* Vol. 13, pp. 5-18

Ulrich, K. & Eppinger, S. (1995). *Product Design and Development*, McGraw-Hill, New York

Verplank, B. (2003). *Interaction design sketchbook*, Available at: http://ccrma.stanford.edu/courses/250a/lectures/IDSketchbok.pdf, (accessed 10 May 2008)

14

Augmented Reality E-Commerce: How the Technology Benefits People's Lives

Yuzhu Lu
Wolfram Research, Inc.
yuzhu@wolfram.com

Shana Smith
National Taiwan University, Department of Mechanical Engineering
ssmith@ntu.edu.tw

1. Introduction

In general, technology can benefit people's lives. For example, during the past 20 years, with the deveolpment of computer and Internet technology, e-commerce and online shopping have rapidly progressed, due to the convenience that they provide consumers. E-commerce websites, such as Amazon.com, Dell.com, and eBay.com, have become an integral part of many shoppers' lives.

However, according to most shoppers' experiences, e-commerce and online shopping are still not able to fully replace onsite shopping, especially for products like clothing, shoes, jewelry, and furniture. For many such products, onsite shopping has many distinct advantages over online shopping. One of the main advantages of onsite shopping is that it usually provides more direct interaction with the actual product. In contrast, conventional online shopping websites often cannot provide enough information about a product for the customer to make an informed decision before checkout. Onsite shoppers frequently engage in some sort of interaction with their potential purchase to discover the scent, texture, appearance, and/or sound characteristics of a product before buying it. Such experience is often impossible with current online purchases.

However, technology is progressing. In particular, Augmented Reality (AR), an emerging Human-Computer Interaction technology, which aims to mix or overlap computer-generated 2D or 3D virtual objects and other feedback with real world scenes, shows great potential for enhancing e-commerce systems. Unlike VR, which replaces the physical world, AR enhances physical reality by integrating virtual objects into the physical world. The virtual object becomes, in a sense, an equal part of the natural environment.

This chapter presents a new type of e-commerce system, AR e-commerce, which visually brings virtual products into real physical environments for user interaction. The new approach gives customers a chance to "try" a product at home or in another use environment. The chapter presents development of a prototype AR e-commerce system and a user study of the developed prototype. Experiment results and data both validate the new

AR e-commerce system and provide suggestions for improvement. Overall results of the study show that the AR e-commerce system can help customers make better purchasing decisions.

2. Background

2.1 VR in E-commerce

Virtual reality (VR) is a computer-simulated environment that allows users to manipulate 3D virtual models online. Recently, researchers have been using VR in e-commerce to provide consumers with a new type of shopping experience by interacting with virtual product models. Hughes et al (2002) presented an adaptive navigation support system for using a virtual environment for online shopping. Sanna et al. (2002) presented a VR e-commerce system based on VRML. They used QuickTime 3D to generate 360-degree image-based immersive backgrounds and an animated virtual human to help online shoppers navigate through their e-commerce environment. Bhatt (2004) analyzed the interactivity, immersion, and connectivity of several major VR-ecommerce websites, such as amazon.com, ebay.com, and schwab.com. Daugherty et al. (2005) conducted five experiments to study the usability of VR for e-commerce. Their results showed that users acquired more information about products when using a VR-based e-commerce system than when using traditional website tools. Fomenko (2006) developed a tool for creating online VR shops, which also gave domain experts more control during the website development process. With Fomenko's tool, developers can use high-level concepts to model and semi-automatically generate a complete VR shop.

2.2 Moving from VR to AR

Although prior studies show that VR can enhance e-commerce, by providing more product information through enhanced human-computer interaction, current VR methods for e-commerce still only provide scaled virtual product models displayed on traditional computer screens. New, more advanced, methods are needed to provide consumers with more realistic product models, with respect to size, customer experience, and user interaction.

AR is a technology which can mix or overlap computer-generated virtual objects with real-world scenes or objects. Unlike VR, which experientially replaces the physical world, AR enhances physical reality by integrating virtual objects into a physical scene. Generated virtual objects become, in a sense, an equal part of the natural environment.

In recent years, much research has focused on developing AR applications, which could be generally classified into two types, based upon the different devices used: optical see-through AR and video see-through AR. Optical see-though AR uses a semi-transparent screen onto which computer generated objects can be projected; users, can simultaneously view the computer generated images and see through the screen to view the natural background environment and, thus, see an integrated AR scene. Video see-through AR uses cameras to capture the live scene as a video stream. For each viewed image frame, a captured video image frame is processed and computer generated virtual objects are added. One advantage of video see-through AR is that the mixed scene can then be displayed on different devices. With video see-through AR, markers and computer vision methods are often used for tracking. Between the two prominent AR methods, video-based AR has attracted the most attention from researchers.

Although AR methods and applications have progressed significantly over recent years, there has been little research conducted related to using AR to enhance e-commerce. In 2001, Azuma et al. reviewed new advances in AR which, after 1997, included display devices and methods, indoor and outdoor tracking, model rendering, and interaction technologies. At that time, they identified several problems that still needed to be addressed, such as occlusion, broader sensing, advanced rendering, and user perception issues. In addition, in 2005, Swan et al's survey showed that, although there were an increasing number of AR applications, research which considered usability was only a small part (less than 8%) of the total, and most of the usability studies were neither formal nor systematic.

Among the limited number of prior related studies, Zhu et al. (2006) proposed AR in-store shopping assistant devices, which provided personalized advertising and dynamic contextualization. Their study was aimed at using AR technology to enhance in-store shopping. Zhang et al. (2000) proposed and developed a prototype direct marketing system that used AR technology. Salespeople could use the system to show the main features of a product by manually holding a plate with specially designed markers. With their marker-based system, they could mix a 3D virtual product with a real scene, videotape the resulting scene, and then send the video tape to interested customers by email. However, their method of using AR in e-commerce did not make full use of the advantages of AR. With their method, online shoppers had no direct interaction with either physical objects or virtual product models. With only video recordings of AR scenes, customers still might not know whether products are suitable for them in their real physical environments.

Two industry companies: metaio and bitmanagement (http://www.ar-live.de/main.php)(2007), are also trying to cooperate and extend e-commerce systems with AR technology. Users are asked to upload a photo of the personal environment with markers. The mixed scene can then be visualized through an online tool. With their application, online users can visually see how a model fits in their personal environment. However the static-picture approach greatly limits uses' direct interaction with virtual product models in a natural way, and their flexibility to try the virtual product in their environment.

In this study, a new AR e-commerce system was developed using, video see-through AR technology, since the devices needed for this type of AR system is more available to online consumers. Video see-through AR technology is also more flexible because the mixed AR scene can be displayed on different devices, rather than with a special optical see-through device only. The system integrates a full-sized virtual product model into an online shopper's physical environment and provides the customer methods for "realistically" interacting with the virtual product. With this system, online shoppers can directly and freely interact with the product model in their environment and in a more nature way. For example, they can physically move around in their environment to see how the product fits in their space from different viewpoints, and they can also move markers around in their environment to move the virtual products to different locations in their environment. This paper presents both the design of the AR e-commerce assistant system and related usability studies. Several key issues related to using AR to enhance e-commerce are also discussed and analyzed.

3. System and User Interface Design

In this study, an AR e-commerce assistant system was designed to provide consumers with more realistic product experiences and interactions. With the developed AR e-commerce assistant, online consumers can bring a product into their physical environment and even try out and visualize the product in their physical environment while shopping from their computers.

3.1 Structure

Like traditional e-commerce systems, our AR e-commerce system uses the Internet as the primary user interaction platform. However, with our AR e-commerce system, a video camera is needed to capture the consumer's physical environment and then integrate it with virtual objects in real time.

The system was developed as an Active X plug-in for an e-commerce web page. Online users can use web page navigation to search for and view pictures and product related information, just as they would with a traditional e-commerce website. However, online shoppers can also use the plug-in to bring virtual products into their physical environment and then interact with the products to determine if the products are suitable.

The client-server plug-in was made using the MFC and OpenGL libraries. The plug-in works between clients and an e-commerce assistant server through an Internet Explorer interface, so that online consumers can easily log onto the Internet, using different hardware, like a computer, cell phone, or Personal Digital Assistant (PDA), to access the server as shown in Figure 1. In this system, an extra video camera is needed, so that consumers can bring product models into their home, auto, outdoor, or other scenes. ARToolkit (Kato and Billinghurst, 1999) was used for tracking, and Open VRML was used for rendering models. The complete structure of the system is shown in Figure 2.

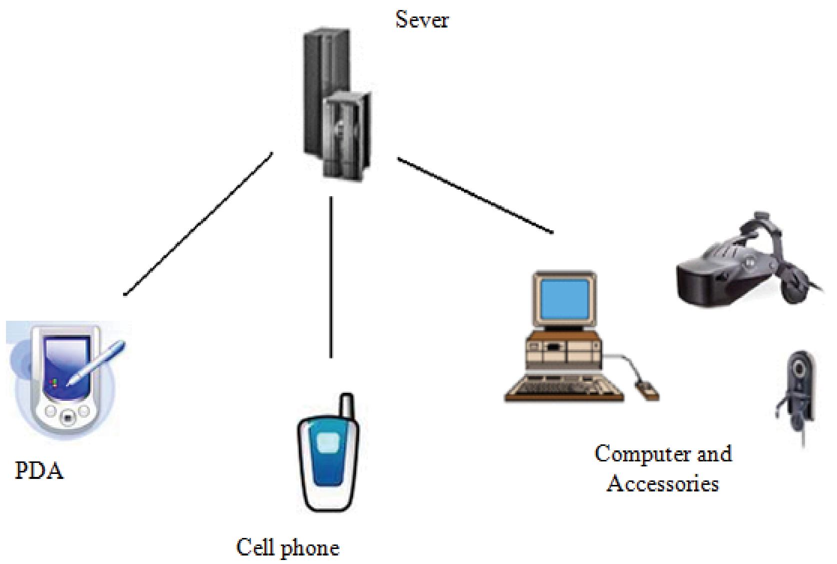


Fig. 1. AR e-commerce assistant system working model

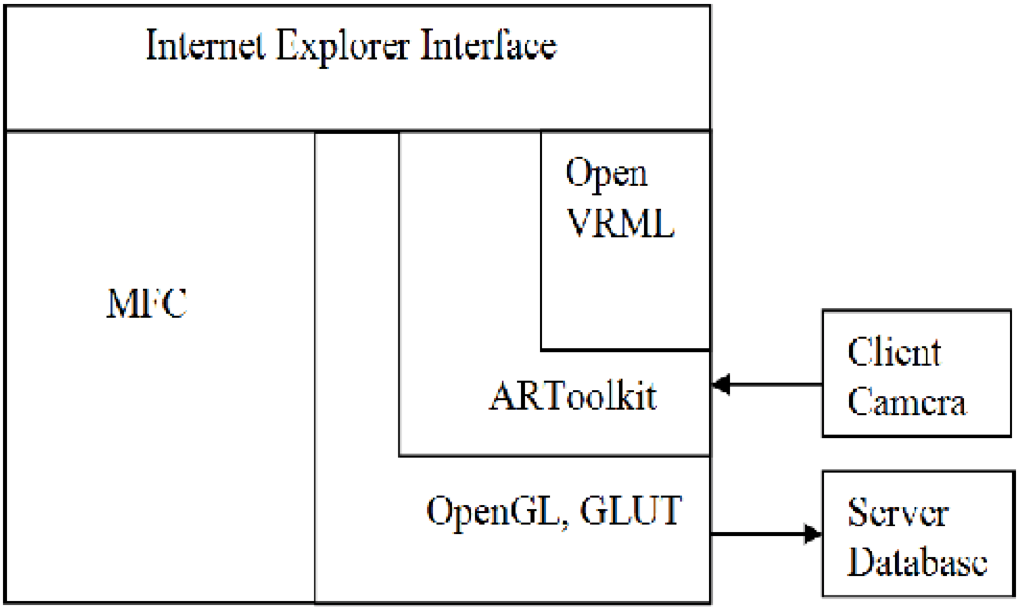


Fig. 2. The structure of the AR e-commerce assistant system

3.2 Interfaces

Primary users of the system are expected to be common computer users, with minimal computer experience. As a result, the user interface of the system was made as simple and user-friendly as possible. In the study, according to our analysis, we determined that consumer shopping typically includes three main tasks:

1. Searching for products.
2. Interacting with products.
3. Acquiring product information.

As a result, the user interface was designed to facilitate the three primary shopping tasks. The three tasks were combined into a two-level menu system within the AR window as shown in Figure 3. A 2D menu system was used, since it is still the most intuitive interaction method for computer users, due to their previous computer experience. Through the menu, users can access the full interaction capability which was designed for AR e-commerce. Shortcut keys are also available to simplify and accelerate interactions between the user and the AR scene.

To provide convenient product searching, a product search interface is provided in the AR window, as shown in Figure 4. As a result, users do not need to exit the AR application every time they want to find another product at web page level and then reopen another AR application for comparing products. Several capabilities were also developed to make product searching efficient, such as searching by keywords, sorting by properties, image viewing, listing operations, and displaying prices. With the tool, users can recursively search for and switch product models in an AR display, to compare products, and thus gain enough direct information to make purchasing decisions. Within the system, for tracking purposes, different markers are used which correspond to different types of products. As a result, online shoppers can also combine different types of products together when shopping. For example, a shopper can combine a table with different chairs or sofas to check the appearance of different combinations in their home.

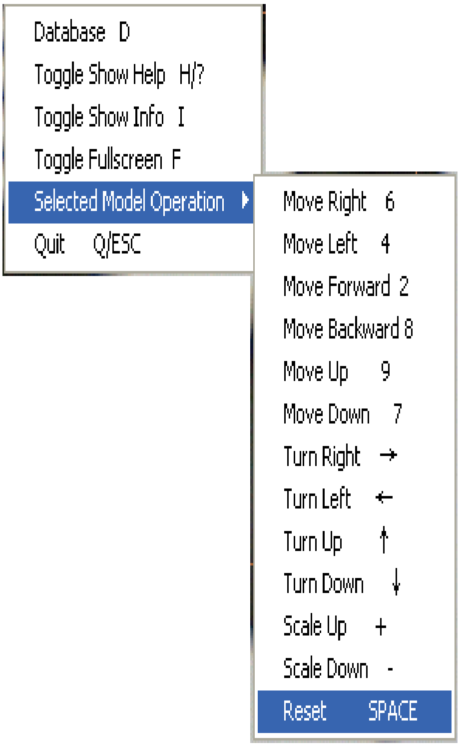


Fig. 3. User interface menu system

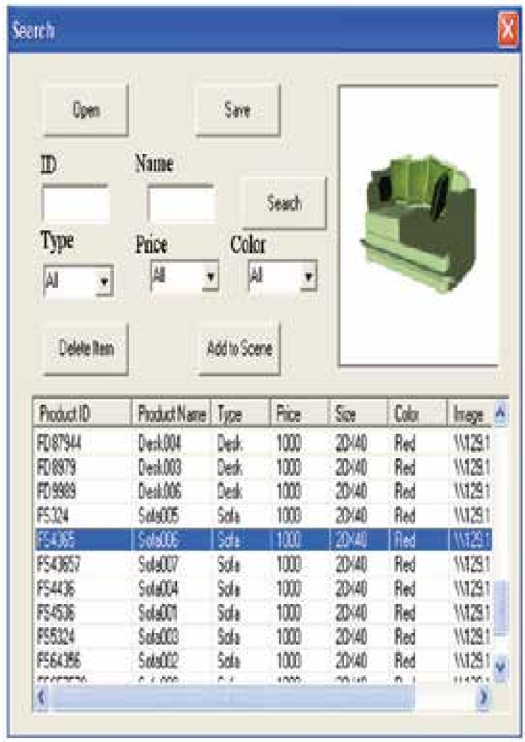
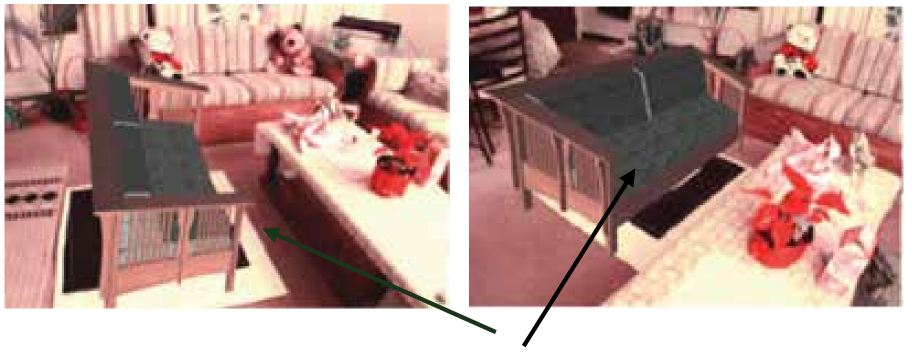


Fig. 4. Product search interface

With models selected from a well-built and normalized database of product loaded to the AR scene, the products can be visualized in actual size within the live background environment which is captured by the local video camera. Users can also pick one of virtual products and manipulate it, for example, move or rotate the model, and view specific information about the selected product, such as name, price, size, and color, to help them make their purchasing decision.

With AR e-commerce, users can have special interactions, which are not available with other applications. Users can walk around their environment, with their laptop, PDA, or cell phone and camera, to see how a product fits in their environment from different viewpoints, as shown in Figure 5. Users can also interact with the AR scene by moving or rotating markers used for tracking.

As mentioned above, the ARToolkit library is used for marker-based tracking in real scenes (Kato and Billinghurst, 1999). Large markers are used for large virtual objects, such as furniture, as shown in Figure 6. Using large markers makes recognition and registration easier and more reliable. With large markers, online consumers can bring virtual furniture or other large virtual products into their homes, and view them from greater distances. Other techniques would cause more instability, since marker tracking is based on computer vision technology. Product models also need to be normalized with respect to marker size so that users see product models in actual size to help them make better buying decisions.



Virtual sofa from different angles

Fig. 5. A virtual model in a real scene

Fig. 6. A Big marker was used

4. Usability Study

A usability study was conducted to compare the developed AR-enhanced e-commerce system with a traditional e-commerce system and a VR-enhanced e-commerce system. To avoid web page design bias, all three web pages were designed using the same design template, which included a word description of the product and a visualization of the product, as shown in Figures 7.-9. The word description parts of the three e-commerce web pages were the same. The only difference among the three types of e-commerce systems was in the visualization component.

For visualization, traditional e-commerce web pages typically use several static 2D pictures of a product, from different perspectives, as shown in Figure 7. With a traditional e-commerce web page, users can visually examine the static 2D product pictures before they buy the product. They can also usually interactively switch between the images. The traditional method is the most commonly used e-commerce approach generally used today.

VR-enhanced e-commerce web pages typically use JAVA applets for visualization. The JAVA applets dynamically download 3D product models in real-time and provide different manipulation capabilities (translate, rotate, zoom) to users, as shown in Figure 8. With VR-enhanced e-commerce web pages, users can easily control and select viewpoints for looking at virtual product models. There might be different types of VR e-commerce web pages. However, this type of design is more representative, since similar types of designs have been used in prior user studies of VR e-commerce (Daugherty 2005) in commercial websites, such as Compaq.com and Dell.com.

Our AR-enhanced e-commerce web page uses ActiveX controls for visualization, as described earlier. System users can visually bring products into their actual physical environments, as shown in Figure 9. With the developed AR-enhanced system, users can hold a laptop, which has a camera, and move around their environment to see how a virtual product model looks, corresponding to the traditional translation, rotation, and zoom interactions in VR e-commerce, and pick operations in traditional e-commerce, and then decide if they want to buy the product. They can also move markers to position the virtual products at different locations to help them make their buying decisions. Figure 10. shows an example of our AR e-commerce system running on a laptop. To control different interaction bias with VR e-commerce and traditional e-commerce, during the user study, participants were not asked to use the developed AR e-commerce menu system.

Fig. 7. Traditional e-commerce with three static 2D images

Fig. 8. VR e-commerce with interactive 3D model

Fig. 9. AR e-commerce interface

(a) AR (b) AR scene on computer

Fig. 10. AR application running on a laptop computer

4.1. Experiment Design

Based on a pilot user study for home furniture products (Lu and Smith, 2006), a formal user study was designed and conducted to test the usability of the developed AR e-commerce system. In the full study, different types of e-commerce web pages were designed for office products (wall hangings and decorative plants) to avoid product-based bias, as shown in Figure 11.

The experiment was designed as within-subjects for types of e-commerce, so that each subject accessed all three e-commerce systems because subjects inevitably differ from one another. In between-subject designs, these differences among subjects are uncontrolled and are treated as error. In within-subject designs, the same subjects are tested in each condition. Therefore, differences among subjects can be measured and separated from error (Howell 2007). Removing variance due to differences between subjects from the error variance greatly increases the power of significance tests. Therefore, within-subjects designs are almost always more powerful than between-subject designs. Since power is such an important consideration in the design of experiments, this study was designed as a within-

subjects experiment to compare user's subjective satisfaction level of using the three different types of e-commerce systems. As a result, by design, different participants' rating standards should not affect the comparisons.
Tests were carried out with six volunteer participants in each of the four office environments. In total, twenty-four participants were tested in the experiment. At the beginning of the experiment, participants were trained to use the three types of e-commerce systems. During the experiment, real-time help concerning how to use the systems was also provided. In the test, participants were asked to use the three types of e-commerce systems to buy different office products for the different environments, without considering budget. Users were asked to select wall hangings and decorative plants and then compare the three types of e-commerce systems. During the experiment, the process was recorded and observed. After the experiment, participants were asked to fill out a questionnaire and to give their evaluations of usability. Four main variables (overall evaluation, information provided, ease of use, and confidence level in the final decision) were measured for each type of e-commerce system for each participant.

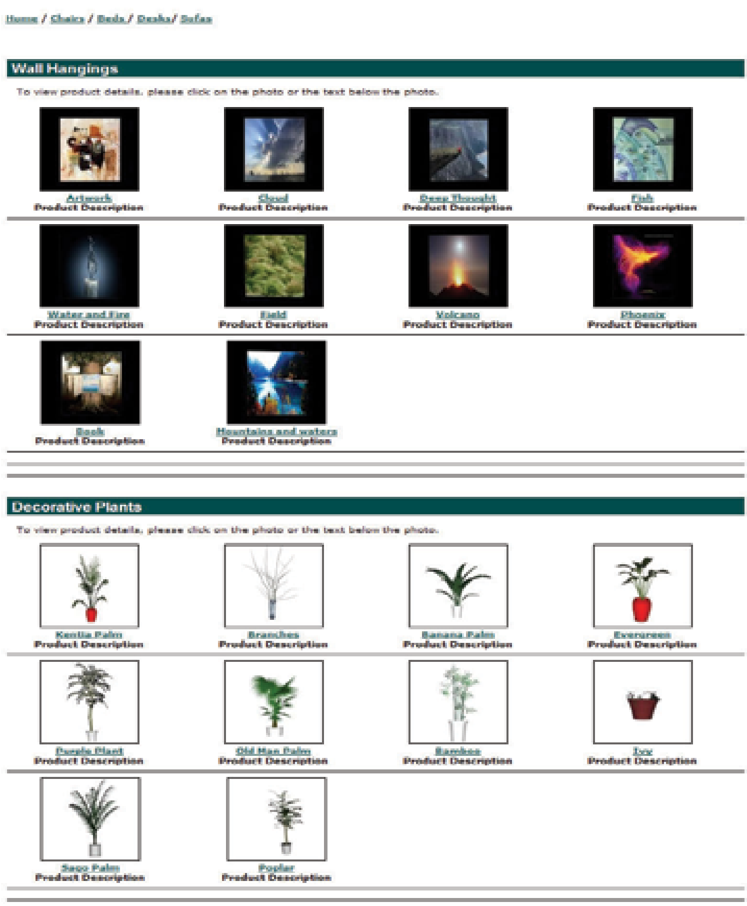


Fig. 11. Office products

In the study, the independent variables were the three different types of e-commerce systems, four different environments (an open space office, a cubicle, a single-user single-room office, and a multi-user single-room shared office). Within each environment, presentation of the e-commerce systems was systematically varied to control the "carryover" effects of a within-subjects design. Since we assigned 6 subjects to each environment, we were able to test all possible presentation orders of the three e-commerce systems (3 choose 1 * 2 choose 1 * 1 choose 1) = 6 different testing orders: (T, VR, AR), (T, AR, VR), (VR, T, AR), (VR, AR, T), (AR, T, VR), and (AR, VR, T). The dependent variables in the research question were four main variables: overall evaluation, information provided, ease of use, and confidence level in the final decision.

To test whether the usability results were affected by experience order, the six user study participants in each of the four environments were randomly assigned to one of the six orders. Evaluations of the four main variables were also compared for the different orders.

The formal study addressed the following hypotheses:

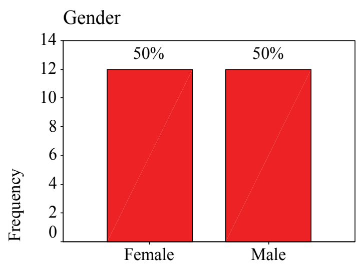
- *Hypothesis 1*: The overall evaluation and satisfaction level of using the AR e-commerce system is higher than using the other two e-commerce systems.
- *Hypothesis 2*: The AR e-commerce system provides more visualization information to online shoppers than the other two e-commerce systems.
- *Hypothesis 3*: The ease of use rating for the AR e-commerce system is lower than the other two e-commerce systems.
- *Hypothesis 4*: Users of the AR e-commerce system have a higher confidence level in their final decision than users of the other two e-commerce systems.
- *Hypothesis 5*: User performance in the different e-commerce systems is not affected by locations.

To test the 5 hypotheses, different ratings given by the participants, after using the three types of e-commerce systems, were compared.

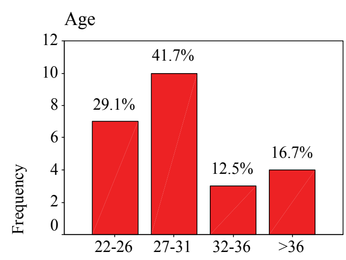
4.2. Experiment Participants

All participants for the study were individuals from Iowa State University who responded to an invitation email. They represented students, staff, and faculty. Figure 12. shows the composition of subjects for the study.

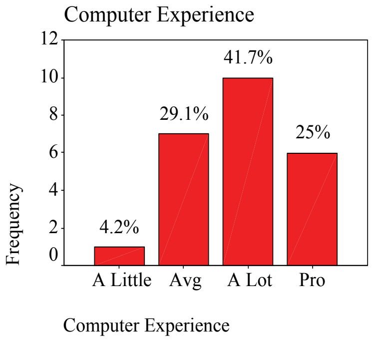
Figure 12. shows that the gender of participants was equally distributed. Since most of the participants were students, the age distribution of participants was skewed toward lower age groups, and computer experience level was skewed toward high levels ("A little" mean little computer experience while "Pro" means professional computer experience), which might have caused some sample bias.



Gender



Age



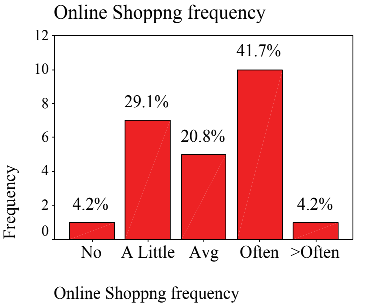


Fig. 12. Participants' self description

4.3 Results

4.3.1 Overall Evaluation

The first research question in the questionnaire was designed to capture overall feelings about the three different types of e-commerce systems, without being affected or guided by later questions. The participants' overall evaluations are listed in Table 1, by locations and by experience orders, which were also separately tested using Factorial ANOVA.

LOCATION	SUBJECTS	PARTICIPANT RATING		
		T	VR	AR
Open space office (1)	1	2	4	5
	2	1	5	5
	3	1	3	4
	4	2	3	5
	5	1	3	4
	6	2	4	5
	Mean/Std. Dev	1.5/0.548	3.667/0.816	4.667/0.516
Cubicle office (2)	7	3	5	4
	8	2	5	4
	9	2	3	4
	10	1	5	5
	11	3	4	5
	12	1	3	5
	Mean/Std. Dev	2/0.894	4.167/0.983	4.5/0.548
Single-user single-room office (3)	13	3	5	5
	14	1	3	5
	15	3	3	5
	16	5	4	4
	17	1	3	5
	18	1	3	5
	Mean/Std. Dev	2.333/1.633	3.5/0.837	4.833/0.408

Multi-user single-room shared office (4)	19	3	4	5
	20	3	4	5
	21	2	3	4
	22	1	2	4
	23	3	4	4
	24	5	5	4
	Mean/Std. Dev	2.833/1.329	3.667/1.033	4.333/0.516
Mean		2.167	3.75	4.583
Std. Dev.		1.204	0.897	0.504

Table 1. Overall evaluation (1=lowest 5=highest)

As shown in Table 1., the mean overall evaluation for traditional e-commerce was 2.167, the mean overall evaluation for VR enhanced e-commerce was 3.75, and the mean overall evaluation for AR enhanced e-commerce was 4.583. As shown in the between-subjects effects and within-subjects effects analysis of Table 2., the p-value for the effect of the type of e-commerce system is very small (<0.05), which indicates that there is a statistically significant difference in mean overall evaluations between the three types of e-commerce systems. In contrast, the p-values for the effect of location is 0.7913, which indicates that there is no statistically significant difference in mean overall evaluations for different locations.

Figures 13., clearly shows that the main effect for different types of e-commerce systems is obvious and that the overall evaluation for the AR e-commerce system is higher than the ratings for the traditional and VR e-commerce systems. The p-value for interaction between types and locations is 0.1407, which indicates that there are no statistically significant interaction effects for types and locations. Thus, interaction effects, and location effects were neglected in the refined analysis model shown in Table 3.

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F	SIG.
Location	3	1.2222	.4074	.3476	.7913
Error	20	23.4444	1.1722		
Type	2	72.3333	36.1667	55.1695**	.000**
Location*Type	6	6.7778	1.1296	1.7232	.1407
Error	40	26.2222	.6556		

**p<0.05

Table 2. Tests of Between-Subjects Effects and Within-Subjects Effects (Dependent Variable: Overall evaluation)

	N	Subset		
Type		1	2	3
Traditional	24	2.1667		
VR	24		3.7500	
AR	24			4.5833
SIG.		1.000	1.000	1.000

Table 3. Homogeneous Subsets Tukey HSD (Dependent Variable: Overall evaluation)

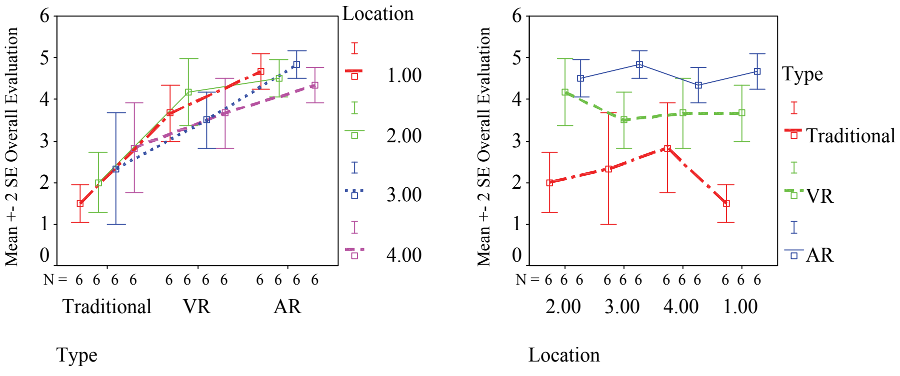


Fig. 13. Interaction between type and location for Overall evaluation

To determine differences in overall evaluations for the three types of e-commerce systems, multiple mean comparisons (Tukey HSD) was used. The analysis results in Table 3. show that each pair of mean overall evaluations for the three types is significantly different.

In comparing the three e-commerce systems, the AR enhanced e-commerce was rated highest by users, which indicates that users preferred the AR enhanced e-commerce system more than the other two for office decoration. Therefore, research hypothesis 1 is accepted. Based on the strength and weakness of AR e-commerce compared to the other two types of e-commerce, customers still preferred AR e-commerce. One of the participants stated, "It is a very high potential method, especially for products like furniture." From the statistical analysis of survey results, there is also no significant evidence that location has any effect on users' overall evaluations. Therefore, the AR e-commerce approach appears to be generally useful in various environments.

4.3.2 Visualized Information Provided

In the questionnaire, users were asked to rate how much information they gained from the three different types of e-commerce systems. Participants' ratings for information provided are listed in Table 4., by locations and by experience orders, which were also tested separately using Factorial ANOVA.

From Table 4., the mean rating for information provided by the traditional e-commerce system was 1.958, the mean information provided by the VR-enhanced e-commerce system was 3.542, and the mean rating for information provided by the AR-enhanced e-commerce system was 4.542. As shown in the between-subjects effects and within-subjects effects analysis of Table 5., the p-value for the effect of type of e-commerce system is very small (<0.05), which indicates that there is a statistically significant difference in mean information provided between the three types of e-commerce systems. However, the p-value for the effect of location is 0.9555, which indicates that there is no statistically significant difference in mean information provided for different locations and different experience orders.

Figure 14. clearly shows that the information users gained from the AR e-commerce system was more than the information they gained from the traditional and VR e-commerce systems. The p-value for the interaction between type and location is 0.9677, which indicates that there was no statistically significant interaction effect between type and location. Thus, the location effect, and interaction effects on information provided were neglected in the refined analysis model shown in Table 6.

To determine the differences between the information users gained for the three types of e-commerce systems, Tukey HSD was used, without considering location or order. With an experiment-wise error rate of 0.05, Table 6. shows that the differences in information provided between the AR e-commerce system and both the traditional e-commerce and VR enhanced e-commerce system are both statistically significant. So the research hypothesis 2 is accepted. Participants also mentioned, in their feedback, that the AR e-commerce system provides the capability to see how products fit in the physical space, so that they can gain more visualization information: "It is very vivid, as if you put a real product into the place where you want. You can efficiently evaluate product information, such as color and size, and determine whether it can match with the scene very well."; "It can provide people an interesting experience and help people gain more information and a much more correct judgment." In addition, statistical analysis of survey results showed that there is no significant evidence that location has an effect on information provided.

LOCATION	SUBJECTS	PARTICIPANT RATING		
		T	VR	AR
Open space office (1)	1	3	3	3
	2	1	3	5
	3	1	3	5
	4	3	4	5
	5	1	4	4.5
	6	3	4	5
	Mean/Std. Dev	2/1.095	3.5/0.548	4.583/0.801
Cubicle office (2)	7	3	4	4.5
	8	2	4	4
	9	2	2	5
	10	1	3	4
	11	3	5	4
	12	1	2	5
	Mean/Std. Dev	2/0.894	3.333/1.211	4.417/0.492
Single-user single-room office (3)	13	3	5	5
	14	1	3	5
	15	1	4	4
	16	4	5	4
	17	1	3	5
	18	1	3	5
	Mean/Std. Dev	1.833/1.329	3.833/0.983	4.667/0.516
Multi-user single-room shared office (4)	19	3	4	5
	20	2	4	5
	21	1	3	4
	22	1	3	4
	23	2	3	4
	24	3	4	5

	Mean/Std. Dev	2/0.894	3.5/0.548	4.5/0.548
Mean		1.958	3.542	4.542
Std. Dev.		1.000	0.833	0.569

Table 4. Information provided (1=lowest 5=highest)

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F	SIG.
Location	3	.3472	.1157	.1062	.9555
Error	20	21.8056	1.0903		
Type	2	81.4444	40.7222	69.4787**	.000**
Location*Type	6	.7778	.1296	.2212	.9677
Error	40	23.4444	.5861		

**p<0.05

Table 5. Tests of Between-Subjects Effects and Within-Subjects Effects (Dependent Variable: Information Provided)

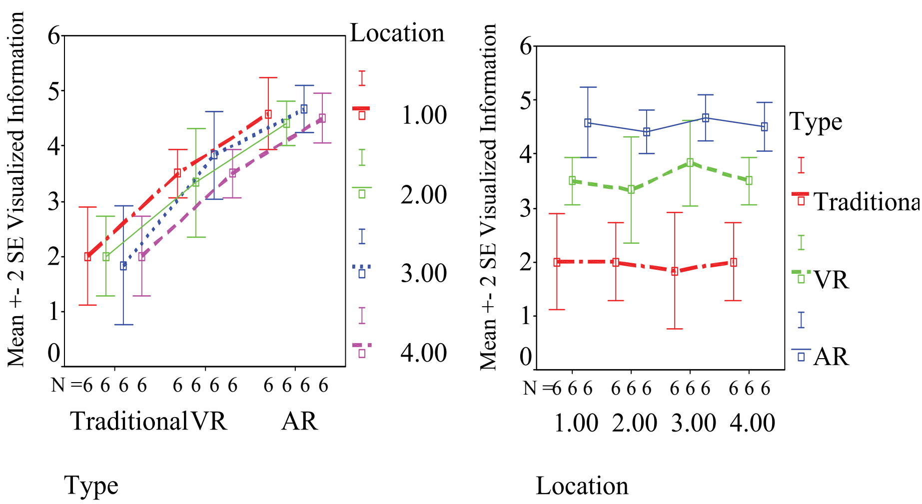


Fig. 14. Interaction between type and location for Information Provided

	N	Subset		
Type		1	2	3
Traditional	24	1.9583		
VR	24		3.5417	
AR	24			4.5417
SIG.		1.000	1.000	1.000

Table 6. Homogeneous Subsets Tukey HSD (Dependent Variable: Information Provided)

4.3.3 Ease of Use

Participants' ratings concerning ease of use for the three different types of e-commerce systems are listed in Table 7., by location and by experience order, which were also tested separately using Factorial ANOVA.

LOCATION	SUBJECTS	PARTICIPANT RATING		
		T	VR	AR
Open space office (1)	1	5	4	2
	2	5	1	5
	3	5	4	3
	4	2	3	5
	5	5	4.5	4.5
	6	2	5	4
	Mean/Std. Dev	4/1.549	3.583/1.429	3.917/1.201
Cubicle office (2)	7	5	5	3
	8	4	4	4
	9	5	3	2
	10	5	5	4
	11	4	5	3
	12	5	4	3
	Mean/Std. Dev	4.667/0.516	4.333/0.816	3.167/0.753
Single-user single-room office (3)	13	5	4	4
	14	5	5	5
	15	5	4	5
	16	5	4	3
	17	5	4	3
	18	5	4	2
	Mean/Std. Dev	5/0	4.167/0.408	3.667/1.211
Multi-user single-room shared office (4)	19	5	5	3
	20	5	3	5
	21	4	4	3
	22	5	3	2
	23	4	4	3
	24	5	5	3
	Mean/Std. Dev	4.667/0.516	4/0.894	3.167/0.983
Mean		4.583	4.021	3.479
Std. Dev.		0.881	0.938	1.037

Table 7. Ease of use (1=lowest 5=highest)

The mean ease of use for the traditional e-commerce system was 4.583, the mean ease of use for the VR enhanced e-commerce system was 4.021, and the mean ease of use for the AR enhanced e-commerce system was 3.479. As shown in the between-subjects effects and within-subjects effects analysis of Table 8., the p-value for the effect of type of e-commerce system is 0.0027 (<0.05), which indicates that there is a statistically significant difference in mean ease of use between the three types of e-commerce systems. In contrast, the p-value for the effect of location is 0.4033, which indicates that there is no statistically significant difference in mean ease of use for different locations.

Figure 15. shows the main effect of different types of e-commerce systems. Ease of use for the AR e-commerce system is much lower than ease of use for the traditional and for the VR e-commerce systems. The p-value for the interaction effect between type and location is 0.5186, which indicates that there are also no statistically significant interaction effects for type and location or type. Thus, the interaction effects, for ease of use were neglected in the refined analysis model shown in Table 9.

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F	SIG.
Location	3	1.9444	.6481	1.0234	.4033
Error	20	12.6777	.6333		
Type	2	14.6319	7.3160	6.8721**	.0027**
Location*Type	6	5.6181	.9363	.8795	.5186
Error	40	42.5833	1.0646		

**p<0.05

Table 8. Tests of Between-Subjects Effects and Within-Subjects Effects (Dependent Variable: Easiness to Use)

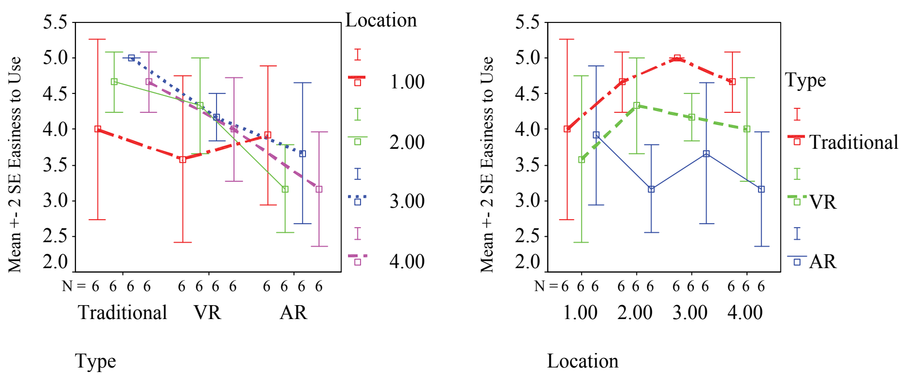


Fig. 15. Interaction between type and location for Easiness to use

	N	Subset	
Type		1	2
AR	24	3.4792	
VR	24	4.0208	4.0208
Traditional	24		4.5833
SIG.		.128	.110

Table 9. Homogeneous Subsets Tukey HSD

To determine the differences between ease of use for the three types of e-commerce systems, Tukey HSD was used, without considering location or order. With an experiment-wise error rate of 0.05, Table 9. shows that the difference in ease of use between the traditional e-commerce system and the VR enhanced e-commerce system is not statistically significant. The difference between the VR enhanced e-commerce system and the AR enhance e-commerce system is also not statistically significant. However, ease of use for the traditional e-commerce system is significantly better than ease of use for the AR enhanced e-commerce system.

So the research hypothesis that ease of use for the AR e-commerce system is lower than for the traditional e-commerce systems is accepted. Participants mentioned in their feedback that the AR e-commerce system needs more high-end hardware equipment, and that it is inconvenient to use: "It is not very convenient to hold the laptop with your hands all the time." There are two explanations for the finding. The first is that AR e-commerce uses more devices and needs more computer skills. The second is that users were still not familiar with AR and AR system interactions. Meanwhile, there is also no significant evidence that location has significant effects on ease of use.

4.3.4 User Confidence Level for Decision

The final main dependent variable measured in the questionnaire was the user's confidence level in their decision (buy or not buy). Participants' ratings are listed in Table 10., by location and by experience order, which were also tested using Factorial ANOVA.

The mean user confidence level for the Traditional e-commerce system was 2.25, the mean user confidence level for the VR enhanced e-commerce system was 3.542, and the mean user confidence for the AR enhanced e-commerce system was 4.646. As shown in the between-subjects effects and within-subjects effects analysis of Table 11., the p-value for the effect of type of e-commerce system is very small (<0.05), which indicates that there is a statistically significant difference in user confidence level between the three types of e-commerce systems. However, the p-value of the effect of location is 0.1184, which indicates that there is no statistically significant difference in user confidence level for different locations.

Figures 16. clearly shows the main effect for different types. User confidence level for the AR e-commerce is much higher than user confidence level for either the traditional or the VR e-commerce systems. The p-value for the interaction effect of type and location is 0.3923, which indicates that there is no statistically significant interaction effect for type and location, Thus, location effect, and interaction effects on user confidence level were neglected in the refined analysis model as shown in Table 12.

To determine the differences in user confidence level for the three types of e-commerce systems, Tukey HSD was used, without considering location or order. With an experiment-wise error rate of 0.05, Table 12. shows that the difference in user confidence level between the AR e-commerce system and both the traditional e-commerce system and the VR enhanced e-commerce system was statistically significant.

The results show that users had a higher confidence level in their shopping decisions when using the AR enhanced e-commerce system, rather than the other two e-commerce systems, for purchasing office decoration products. Therefore, research hypothesis 4 is accepted. Participant comments included: "AR e-commerce makes shopping more visually intuitive."; "The user naturally sees what will happen before actually buying."; "It gives you a real-time experience in your own environment so that you can instantly tell whether or not the product is a good fit." Meanwhile there was also no significant evidence that location had an effect on user confidence level.

LOCATION	SUBJECTS	PARTICIPANT RATING		
		T	VR	AR
Open space office (1)	1	2	2	4
	2	1	3	5
	3	1	3	4
	4	2	4	5
	5	1	4	4.5
	6	2	3	5
	Mean/Std. Dev	1.5/0.548	3.167/0.752	4.583/0.491
Cubical office (2)	7	2	5	4
	8	2	4	5
	9	3	4	5
	10	2	4	3
	11	3	4	5
	12	1	2	5
	Mean/Std. Dev	2.167/0.752	3.833/0.983	4.5/0.837
Single-user single-room office (3)	13	3	4	5
	14	3	5	5
	15	4	3	5
	16	4	4	3
	17	1	3	5
	18	2	3	5
	Mean/Std. Dev	2.833/1.169	3.667/0.816	4.667/0.816
Multi-user single-room shared office (4)	19	3	4	5
	20	2	3	5
	21	3	4	5
	22	2	4	5
	23	2	3	4
	24	3	3	5
	Mean/Std. Dev	2.5/0.548	3.5/0.548	4.833/0.408
Mean		2.25	3.542	4.646
Std. Dev.		0.897	0.779	0.634

Table 10. User confidence level for decision (1=lowest 5=highest)

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F	SIG.
Location	3	4.2049	1.4016	2.2107	.1184
Error	20	12.6806	.6340		
Type	2	69.0208	34.5104	64.6229**	.0000**
Location*Type	6	3.4514	.5752	1.0772	.3923
Error	40	21.3611	.5340		

**p<0.05

Table 11. Tests of Between-Subjects Effects and Within-Subjects Effects (Dependent Variable: User Confidence Level for Decision)

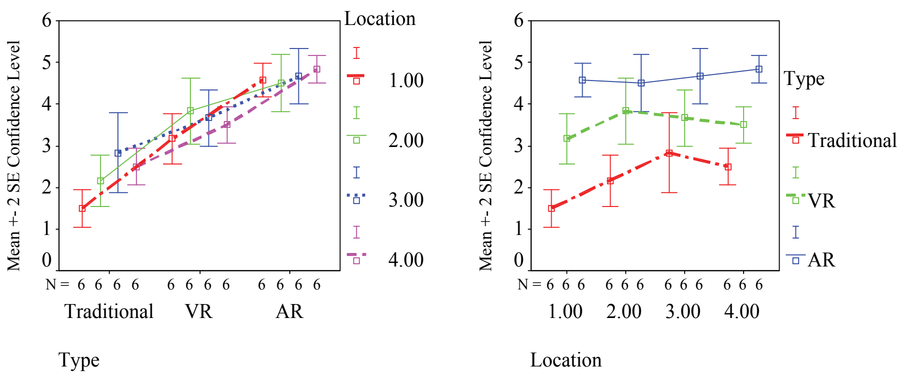


Fig. 16. Interaction between type and location for Confidence Level for Decision

	N	Subset		
Type		1	2	3
Traditional	24	2.2500		
VR	24		3.5417	
AR	24			4.6458
SIG.		1.000	1.000	1.000

Table 12. Homogeneous Subsets Tukey HSD (Dependent Variable: User Confidence Level in Decision)

4.4 Observations and Users' Comments

4.4.1 "As Is" View

95.8% of participants mentioned, in their feedback, that the AR e-commerce system provides the capability to see how products fit in the physical space. Users' comments included: "It is visually intuitive."; "The user naturally sees what will happen before actually buying."; "It gives you a real-time experience in your own environment so that you can instantly tell whether or not the product is a good fit."; "It presents products in a real scale relative to the environment, and is able to show views from several perspectives."; "AR makes shopping more confident.”; “It is cool and helpful for making the decision.”; “It is very vivid, as if you put a real product into the place where you want. You can efficiently evaluate product information, such as color and size, and determine whether it can match with the scene very well.”; “It can provide people an interesting experience and help people gain more information and a much more correct judgment.”

4.4.2 Ease of Use

87.5% of participants mentioned in their feedback that the AR e-commerce system needs more high-end hardware equipment, and that it is inconvenient to use. Users’ comments included: “You have to have a laptop or mobile device.”; “It is not very convenient to hold the laptop with your hands all the time.”; “It is constrained to a marker.”; “It is limited to certain viewing areas.”; “If the designer could use a small device (like a cell phone) to replace the laptop, it would be more convenient for customers.”; “It is slower for the user and more complicated.”; “If it was more user friendly and more easy to use, it would be widely used.”; “Not as convenient as VR and traditional e-commerce.”
However, 12.5% of participants believed that the AR e-commerce system was convenient to use. Users’ comments included: “It is very easy.”; “There is not much I have to learn to dive right in.”; “It is friendly and looks real.”; “It is easy to manipulate. It is a more natural interactive method than mouse interaction.”; “It is more convenient, and otherwise, it is difficult to shop at onsite stores that are far away.”

4.4.3 Unstable

29.2% of participants mentioned in their feedback that the AR e-commerce system is unstable: “The images on the screen are not stable, and sometimes disappear due to problems with light intensity.”; “If people could easily change the position of the target, without considering light problems, it would be better.”; “The smoothness of motion tracking needs to be improved.”; “There are limited spots where you can see the product.”; “Sometimes I cannot see the virtual image.”

4.4.4 Real Modeling and Rendering

25% of participants said that the virtual objects in the AR e-commerce display were not very real: “If it looked more realistic, it would be better.”; “If the models looked the same as the real objects, it would be better.”; “The model should be designed more accurately.”; “It needs some easy way to directly transfer real things into 3D virtual models.”; “It needs accurate illumination.”; “It would be great if I could feel the texture of a product”.

4.4.5 Internet Speed

25% of participants felt that the Internet wireless connection speed used was not fast enough for AR e-commerce. They considered the process of downloading models to be slow. However, they believed this problem would be solved with further development of technology. One user said: “While I thought that the quality of the graphics of the product would be an issue, I found that the AR system provided me with an excellent sensation of the product. The lack of a very high graphical representation of the product did not bother me at all.”

5. Discussion and Conclusions

Traditional e-commerce systems have reached a limitation that needs to be overcome, because they do not provide enough direct information for online shoppers, especially when they are shopping for products like furniture, clothing, shoes, jewelry, and other decorative products. In this study, we developed an AR e-commerce system and studied the effectiveness of AR for enhancing e-commerce.

A formal usability study was designed and conducted. Usability experiment results verified that the developed AR e-commerce system could be used to provide more direct product information to online shoppers and thereby help them make better purchasing decisions. Additionally, in the study, users preferred the AR e-commerce system more than traditional e-commerce and VR e-commerce systems.

Although the AR e-commerce system provides more information and interaction capability than the other e-commerce systems, it is also evident that some limitations still exist in the proposed approach. According to the study participants, the major limitation of using the AR e-commerce system is that it is currently not as easy to use as the traditional or VR e-commerce systems. The AR e-commerce system's interaction method still needs to be improved, to make it more convenient for users. For example, the system could offer online shoppers differet modes for using the system. Such as uploading static pictures with markers, or uploading pre-made videos so that users do not need to carry a laptop computer around for viewing each product. The application could also be implemented on PDAs and cell phones, which are available to most consumers and which are also light and easy to carry.

The rendering methods used also need to be improved to help integrate virtual models into real scenes more seamlessly. For example, more texture mapping could be used to improve virtual product realism. Real time occlusion could also be implemented to help consumers' depth perception and visualization of virtual products placed in their environments. The computer vision algorithm used in the AR system needs to be improved, to make the marker tracking more stable, even in a poor lighting condition. New and better algorithms should also be studied and developed for partial marker tracking so that users do not need to worry about the virtual product disappearing because the marker is partially occluded. In addition, the system should be updated to use the latest high-speed wireless Internet technology, when available, since current wireless Internet technology is currently still not fast enough to transfer high-resolution product models in real-time.

6. References

ARToolkit document, Retrieved March 31, 2007 from http://www.hitl.washington.edu/artoolkit/documentation/

Azuma, R.; Baillot, Y.; Behringer, R.; Feiner, S.; Julier, S. & MacIntyre, B. (2001). Recent Advances in Augmented Reality, *IEEE Comp. Graph. & App* 21 (6), 34-47.

Bhatt, G. (2004). Bringing virtual reality for commercial Web sites, *International Journal of Human-Computer Studies* 60 (1), 1-15.

Cho, Y.; Im, I.; Hiltz, R. & Fjermestad, J. (2002). An Analysis of Online Customer Complaints:

Implications for Web Complaint Management, *35th Hawaii International Conference on System Sciences (HISS '02)* 7, 176.

Daugherty, T.; Li, H. & Biocca, F. (2005). Experiential commerce: A summary of research investigating the impact of virtual experience on consumer learning, *Society of Consumer Psychology: Online Advertising*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fomenko, V. (2006). Generating Virtual Reality Shops for E-commerce, Ph.D. Dissertation, Vrije Universiteit Brussel.

Hughes, S.; Brusilovsky, P. & Lewis, M. (2002). Adaptive navigation support in 3D e-commerce activities. *AH'2002*, Malaga, Spain, 132-139.

Kato, H.; & Billinghurst, M. (1999). Marker Tracking and HMD Calibration for a Video based Augmented Reality Conferencing System, *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, San Francisco, CA, 85-94

Lefebvre, L.A. & Lefebvre, E. (2002). E-commerce and Virtual Enterprises: Issues and Challenges for Transition Economies, *Technovation* 22 (5), 313-323.

Lu, Y. & Smith, S. (2006). Augmented Reality E-Commerce Assistant System: Designing While Shopping, *Proceedings of IDETC/CIE'06*, Philadelphia, PA, paper number DETC2006-99401.

Sanna, A.; Zunino, C. & Lamberti, F. (2002). HAVS: a human animated VRML-based virtual shop for e-commerce, *SCI'02 Proc.*, vol. XII, 24-29.

Standing, C. (2000). Internet Commerce Development, *Artech House Computing Libraty*, Hardcover.

Swan II, J.E. & Gabbard, J.L. (2005). Survey of User-Based Experimentation in Augmented Reality, *Proceedings of 1st International Conference on Virtual Reality*, Las Vegas, Nevada.

Zhang, X.; Navab, N. & Liou S.P. (2000). E-Commerce Direct Marketing using Augmented Reality. *IEEE International Conference on Multimedia and Expo (I),* 88-91.

Zhu, W.; Owen, C.B.; Li, H. & Lee, J.H. (2006). Design of the PromoPad: an Automated Augmented Reality Shopping Assistant. *12th Americas Conference on Information Systems*, Acapulco, Mexico.

Howell, D.C. (2007). *Statistical Methods for Psychology*. 6th Edition. Thomson Wadsworth.

15

Multi-Dimensional Force Sensor Design for Haptic Human-Computer Interaction

Aiguo Song
Department of Instrument Science and Engineering, Southeast University
P.R.China

1. Introduction

Haptic human-computer interaction (HapHCI) is interaction between a human and a computer with realistic sense of touch. Haptic interaction between human and computer involves solving challenging problems in mechanical design, sensor, actuator, computer graphics, physical-based modelling and rendering algorithm, human capabilities, and other areas. With the increasing applications of HapHCI in virtual reality, teleoperation, rehabilitation, tele-surgery, entertainment, etc, the importance of the sense of touch for human-computer interaction has been widely acknowledged [Gabriel 2006]. For example, in virtual surgery training system, the surgeon controls the surgical tools and characterizes virtual tissues as normal or abnormal through the sense of touch provided by the HapHCI device. Another example is HapHCI based rehabilitation system for post-stroke patient exercise. During active rehabilitation exercise process, it requires accurate damping force control, and during passive rehabilitation exercise process, relatively accurate traction force is necessary.

HapHCI technique usually consists of three fundamental parts: force/tactile measuring, haptic modelling, and haptic display device. Haptic modelling as well as haptic display hardware has been discussed a lot and exploited for ten years, particularly in the area of virtual reality. However, so far, little attention has been paid to the design of multi-dimensional force sensor for HapHCI, and the existing commercial six degree-of-freedom (DOF) force sensors are designed mainly for industrial robot control, which are too expensive and often over designed for HapHCI in axis and in bandwidth. As an important component in the HapHCI system, multi-dimensional force sensor not only measures the human hand force/torque acted on the interactive hardware device, such as hand-controller, master-manipulator, joystick etc, as a command input the computer, but also provides force/torque information for close-loop control of precise haptic display.

A number of multi-dimensional force sensors have been developed during the past decades, which are intended for use at the end effector of a robot to monitor assembly or machine force. Most of them are six axes force/torque sensors [Watson, Drake, 1975] [Lord Corporation, 1985] [Nakamura et al., 1987] [Kaneko, Nishihara, 1993] [Kim, 2001], which measure three axes forces Fx, Fy, Fz, and three axes torques Mx, My, Mz. And some of them

are three axes force sensors, such as RIXEN EFS-202 [Emplus Corporation, 1991], three-axis gripper force sensor [Kim, 2007], which only measure three axes forces Fx, Fy, Fz.
Although there exist several different types of multi-dimensional force sensors, Lorenze et al pointed out that most of the conventional force sensors are not suitable for use in HapHCI systems, which often over designed for measuring the interaction force between human and machine, and Lorenze et al presented a new type of force sensor for HapHCI, which only measures x and y force components [Lorenze et al, 1999].
This chapter focuses on the multi-dimensional force sensor design for HapHCI. We firstly discuss the role played by the force/torque sensor in the HapHCI systems, and build the dynamics model of the force/torque sensor in the HapHCI. Then we give the general principles of force/torque sensor design for HapHCI. According to the proposed design principles, a novel 4 DOF force/torque sensor for HapHCI is developed, which is designed to measure three axes forces Fx, Fy, Fz, and one axis torque Mz by ignoring the other two axes torques. In this chapter, the mechanical structure of the 4 DOF force/torque sensor is presented, and the strain of the elastic body is analyzed in theory and by FEM analysis software ANSYS, respectively. At last, the calibration results of the 4 DOF force/torque sensor are given. The FEM analysis and calibration results show the new force/torque sensor has low cross sensitivity without decoupling matrix calculation. This 4 DOF force/torque sensor is easier to fabricate with lower cost than the existing commercial force/torque sensors. It is well suitable for use in HapHCI systems.

2. Dynamics model of the force sensor in the HapHCI

A typical haptic human-computer interaction system is show in Figure 1. The human operator holds the human-computer interaction device (e.g. hand controller, master manipulator, Phantom hand, etc.) to control the avatar (e.g. virtual hand, virtual probe, etc.) touch with the virtual objects in virtual environment. The force sensor is usually installed between human hand and the human-computer interaction device, which is used to measure the interactive force between them. The position sensor on the human-computer interaction device acquires three dimensional space position of the human hand as command input to the virtual environment. The computer calculates the interactive force between avatar and virtual objects by using the haptic model and feeds the touch force back to the human hand through the human-computer interaction device.
In general, the HapHCI system in Figure 1 can be represented by the block diagram of Figure 2.

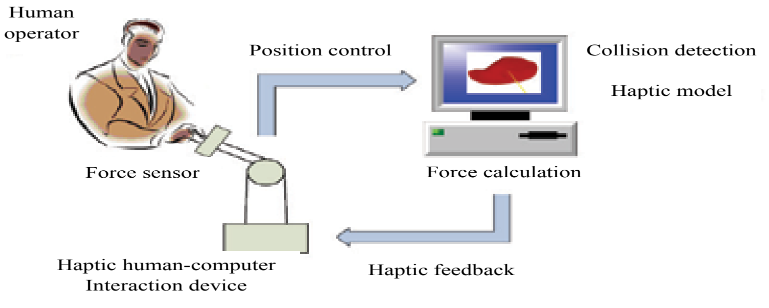


Fig. 1. A typical haptic human-computer interaction system

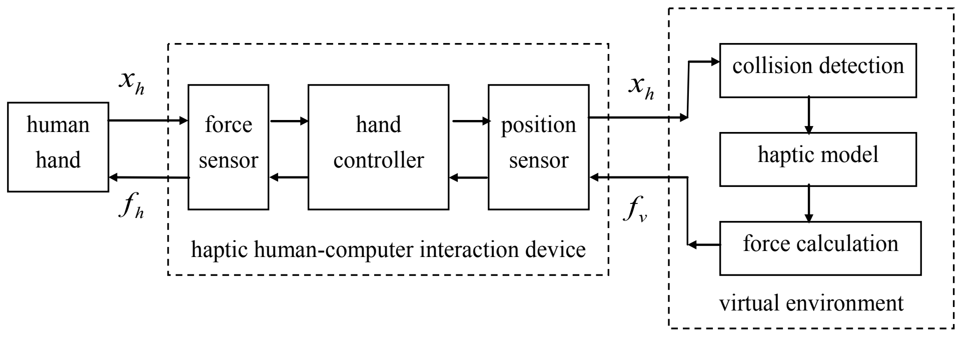


Fig. 2. The block diagram of haptic human-computer interaction system

The dynamics of interaction between human hand and HapHCI system is given as follows.

$$\begin{aligned} f_h - f_v = {} & (M_{fsensor} + M_m + M_{psensor})\ddot{x}_h + (B_{fsensor} + B_m + B_{psensor})\dot{x}_h \\ & + (K_{fsensor} + K_m + K_{psensor})x_h \end{aligned} \tag{1}$$

where x and f denote position and force, respectively; M, B and K denote mass, damp and spring coefficient, respectively; subscript 'fsensor', 'm' and 'psensor' denote the force sensor, hand controller and the position sensor, respectively; fh is interactive force between human hand and HapHCI device, and fv is calculation force by computer based on haptic model. Considering the position sensor is usually relatively very small in size and very light in weight, the equation (1) can be rewritten as

$$f_h - f_v = (M_{fsensor}\ddot{x}_h + B_{fsensor}\dot{x}_h + K_{fsensor}x_h) + (M_m\ddot{x}_h + B_m\dot{x}_h + K_m x_h) \tag{2}$$

For an ideal virtual environment, its dynamics can be expressed as

$$f_v = M_v\ddot{x}_v + B_v\dot{x}_v + K_v x_v \tag{3}$$

$$x_v = x_h \tag{4}$$

According to analogy between mechanical and electrical systems [Anderson, Spong, 1989], equation (3) (4) can be rewritten by using Laplace transformation as follows

$$F_h - F_v = Z_{fsensor}V_h + Z_m V_h \tag{5}$$

$$F_v = Z_v V_v \tag{6}$$

$$V_v = V_h \tag{7}$$

Where, $Z_{fsensor} = M_{fsensor}s + B_{fsensor} + K_{fsensor}/s$ is mechanical impedance of force sensor, $Z_m = M_m s + B_m + K_m/s$ is mechanical impedance of HapHCI device, and $Z_v = M_v s + B_v + K_v/s$ is mechanical impedance of virtual enviroment. $F(s)$ and $V(s)$ are Laplace transforms of $f(t)$, $\dot{x}(t)$, respectively.

Thus, we can represent the HapHCI system as a circuit, shown in Figure 3.

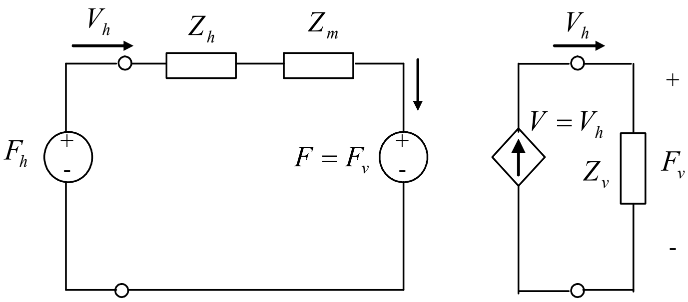


Fig. 3. Circuit representation of HapHCI system

Ideally, human operator can feel as if he is directly touching the virtual environment by maneuvering the HapHCI device, that is to say, the virtual avatar motion $\dot{x}(t)$ equals to the human hand motion $\dot{x}(t)$, and the force of human hand acted on the HapHCI device f_h equals to the virtual interactive force between avatar and virtual object f_v. Lawrence defined the transparency notion for a teleoperation system [Lawrence, 1993]. Here, we extend the transparency notion to evaluate HapHCI system. If a HapHCI system is completely transparent only when it satisfies the condition

$$Z_{feel} \overset{def}{=} \frac{F_h}{V_h} \equiv Z_v \overset{def}{=} \frac{F_v}{V_v} \tag{8}$$

where Z_{feel} is the virtual impedance felt by the human operator.

From equtions (5) (6) (7) (8), we have

$$Z_{feel} = Z_{fsensor} + Z_m + Z_v \tag{9}$$

So it is obvious, reducing the mechanical impedances of force sensor and HapHCI device $Z_{fsensor}$, Z_m can increase the transparency of the HapHCI system, especially when $Z_{fsensor}$, Z_m both equal to zero, the HapHCI system is completely transparent.

Because the damp of force sensors is near zero and stiffness is relatively very high, the mechanical impedance of force sensor is mainly determined by its mass.
Therefore, from the viewpoint of transparency, one of the important requirements of multi-dimensional force sensor design is mass minimization.

3. Principles of force/torque sensor design for HapHCI

Human-computer interaction requires different properties of a force sensor than typical robot applications such as machining and assembly. These differences have substantial impact on how a force sensor can be designed.

3.1 Fewer degrees of freedom required

Owing to the difficulty of mechanical design and motor control for HapHCI device with 6 DOF force feedback, most of the existing HapHCI devices are designed with 3 DOF force feedback, sometimes one torque feedback in addition, although they may be able to move in six directions including 3 DOF translations and 3 DOF rotation.
In spite of six axes force/torque information may be required for some cases in HapHCI systems, the 4 axes force/torque signals, that is three axes forces Fx, Fy, Fz, and one axis torque Mz, are key components of the six axes force/torques, because the torques Mx, My are easy to calculate from the measured forces Fx, Fy and their contact points [Nagarajan et al, 2003]. That is to say the four axes force/torque signals Fx, Fy, Fz, and Mz are sufficient for force sensor design.
Commercial multi-axis force/torque sensors typically measure all six axes forces and torques. In the existing commercial 6 DOF force/torque sensors, there are at least 32 necessary strain gauges stuck to the cross elastic beam, as shown in Figure 4. Owing to difficulty of accurately sticking so many strain gauges to the cross beam, the 6 DOF force/torque sensors usually are very expensive, which restrict their application in HapHCI systems. Another problem of the existing 6 DOF force/torque sensors is coupled interference or noise among six axes, which causes the calibration become much complicate and difficult.

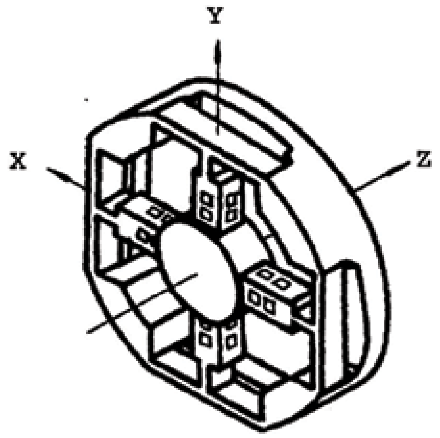


Fig. 4. Mechanical structure of 6 DOF force/torque sensor

3.2 Sensitivity and stiffness requirement

The multi-dimensional force sensor for industrial robot needs very wide bandwidth (more than 1000 Hz bandwidth is often required), which causes the conflict between sensitivity and stiffness during the force sensor design. However, this problem is not faced at all when designing force sensor for HapHCI system. The first reason is the interactive force between human hand and HapHCI device often changes at lower frequency mainly owing to the softness of human hand. The second reason is human is relatively insensitive to small force change and small displacement. So the sensitivity and stiffness of the multi-dimensional force sensor for HapHCI can be lowered a lot (just over 100 Hz bandwidth is needed), which will greatly reduce its expense of fabrication.

3.3 Size and weight requirement

The size and weight of the force/torque sensor for HapHCI is very important. Section 2 has concluded the mass minimization is necessary for force sensor design for HapHCI systems, that means less weight and small size is required. Another reason is that if its diameter is larger or its thickness (length) is longer, it can produce larger inertial force when human hand pulls or pushes the HapHCI devise with a speed, which will reduce both the precision of force measurement and human sense of touch. Furthermore, big size of force sensor will cause it is not easy to install on the existing HapHCI devices (e.g. hand controllers, master manipulators, Phantoms, etc.).

4. A new mechanical structure of the force/torque sensor

We have developed a novel mechanical structure for 6 DOF wrist force/torque sensor before [Huang et al, 1993]. By improving this mechanical structure, we design a new mechanical structure for 4 DOF force/torque sensor for HapHCI [Song et al, 2007], as illustrated in Figure 5.
The elastic body of 4 DOF force/torque sensor consists of center support of the elastic body, cross elastic beam, compliant beams and the base of the elastic body. Where, the cross elastic beam is composed of four symmetric horizontal beams. And four vertical compliant beams connect the four corresponding horizontal beams to the base, respectively.
The whole elastic body is designed to be monolithic and symmetric. Thus, the mechanical structure of the 4 DOF force/torque sensor is light and simple.

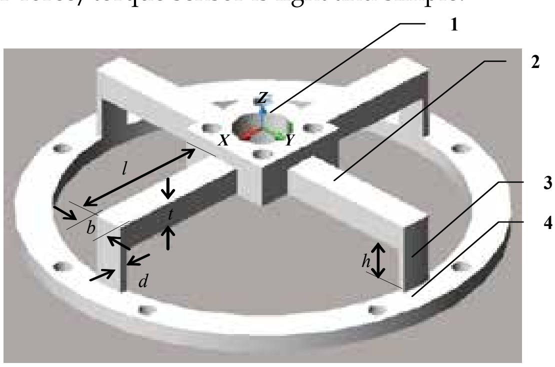


Fig. 5 The mechanical structure for novel force/torque sensor. (1) center support of the elastic body, (2) cross elastic beam, (3) compliance beam, (4) base of the elastic body.

Where, l, b, t are length, width, thickness of the horizontal beam, respectively. And h, d are height, thickness of the vertical compliant beam, respectively. Usually, b=t, $d \le \frac{1}{3}b$.

5. Strain analysis in theory

It can be assumed before analysis that:
(a) The stiffness of the elastic body designed is strong enough for force and moment to be applied. The deformation of the cross elastic beam is within the elastic region for the maximum force and moment applied on it.
(b) The strain gauges are glued correctly, symmetrically and stably.
(c) Every line of the component force passes through the center of the elastic body.
Figure 6 shows the skeleton drawing of the 4 DOF force/torque sensor. When a single force in X direction Fx is applied to the elastic body through its center, the two horizontal beams in X direction OA and OC are float owing to the two vertical beams AA´ and CC´ act as compliant beams, while the other two horizontal beams in Y direction OB and OD become a freely supported beam and produce bending deformation owing to the two vertical beams BB´ and DD´ act as rigid beams.
As is the case for a single force in Y direction Fy, when it is applied to the elastic body through its center, the beams OA and OC become a freely supported beam and produce bending deformation.
When a single force in Z direction Fz is applied to the elastic body through its center, the two horizontal beams OA, OC and two horizontal beams OB, OD become two freely supported beams and produce identical bending deformation.
When a single torque in Z direction Mz is applied to the elastic body through its center, the four horizontal beams OA, OB, OC, OD produce identical bending deformation.

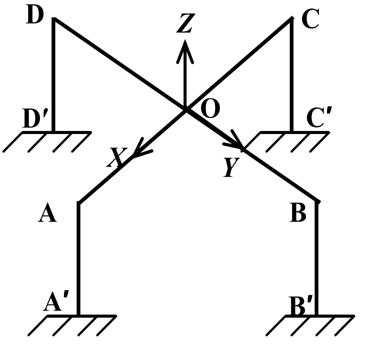


Fig. 6. The skeleton drawing of the sensor

For the novel 4 DOF force/torque sensor, only 16 strain gauges is sufficient for measuring three axes forces and one axis torque, which is twice less than that of 6 DOF force/torque sensor. So it is much easier to stick the strain gauges on the cross elastic beam accurately. Figure 7 depicts skeleton drawing of the distribution of 16 strain gauges on the cross beam.

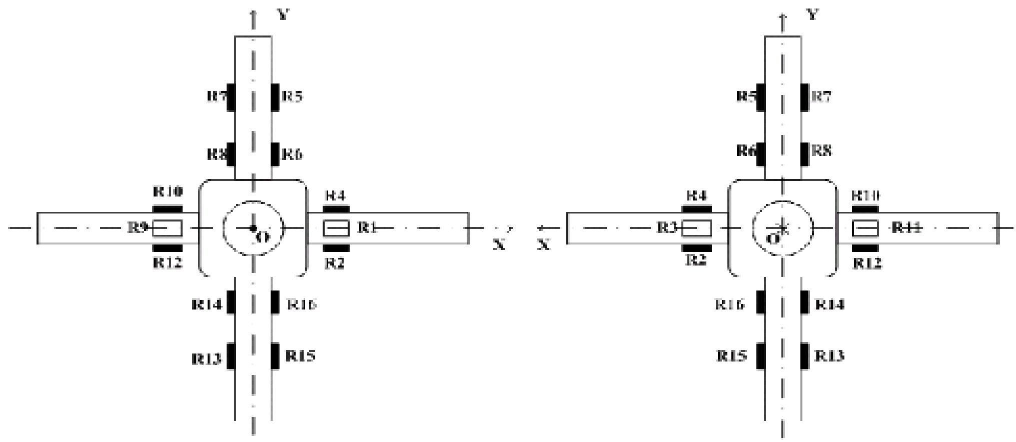


Fig. 7. The distribution of 16 strain gauges on the cross beam

Assuming the strain outputs obtained from the 16 strain gauges R_1, R_2, …,R_{16} are s1, s2, … ,s16, respectively, then we analyze the relationship between the 16 strain outputs and each of the six axes force/torques by using the theory of Mechanism of material.

The 16 strain gauges are divided into four groups and hard wired into four full Wheatstone bridge circuits to measure the four axes force/torques, respectively, as shown in Figure 8.

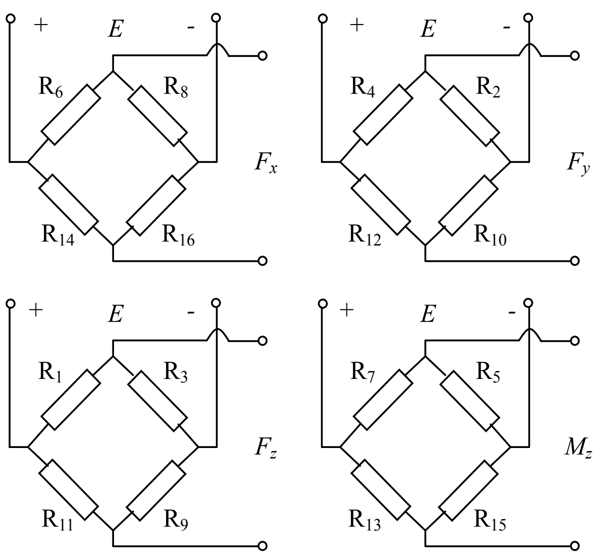


Fig. 8. Four Wheatstone bridge circuits for four axes force/torques measurement

Where, E is voltage of the power supply.

In Reference [Huang, 1993], we have proved an important case of strain gauge output, if a strain gauge is glued at the neutral axis of a beam, when the beam is under bending moment in its flank, the output of the strain gauge is unchanged, as shown in Figure 9.

Here, it is easy to prove another important case of strain gauge output. When a beam is under a torque around its center axis, the output of the strain gauge on its side will increase as a result of the enlargement of gauge length, as shown in Figure 10.

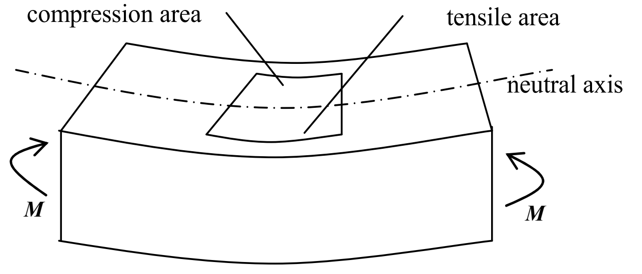


Fig. 9. The beam is under bending moment in its flank

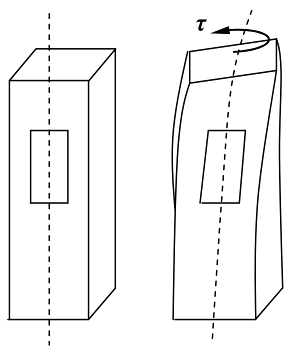


Fig. 10. The beam is under torque moment

From the theory of Mechanics of material, the measured force vector can be easily determined as

$$U_F = \begin{bmatrix} U_{Fx} \\ U_{Fy} \\ U_{Fz} \\ U_{Mz} \end{bmatrix} = \begin{bmatrix} K_1(s_6 + s_{16} - s_8 - s_{14}) \\ K_2(s_4 + s_{10} - s_2 - s_{12}) \\ K_3(s_1 + s_9 - s_3 - s_{11}) \\ K_4(s_7 + s_{15} - s_5 - s_{13}) \end{bmatrix} \tag{10}$$

here, K_1, K_2, K_3, K_4 are coefficients of the U_{Fx}, U_{Fy}, U_{Fz}, U_{Mz}, respectively, which are determined when the 4 DOF force/torque sensor is designed.

When single one of the six axes force/torques is applied to the sensor, it is not difficult to deduce the relationship between 16 gauge outputs and each of the six axes force/torques from the theory of Mechanics of material. The results are seen in table 1. Here, "+", "-" denote the increment and reduction of gauge output, respectively, and "0" means fixedness.

	Applied force/torques					
	F_x	F_y	F_z	M_z	M_x	M_y
s_1	**0**	**0**	+	**0**	+	**-**
s_2	**0**	**-**	**0**	**-**	+	**0**
s_3	**0**	**0**	**-**	**0**	+	+
s_4	**0**	+	**0**	+	+	**0**
s_5	+	**0**	**0**	**-**	**0**	+
s_6	+	**0**	**0**	**-**	**0**	+
s_7	**-**	**0**	**0**	+	**0**	+
s_8	**-**	**0**	**0**	+	**0**	+
s_9	**0**	**0**	+	**0**	+	+
s_{10}	**0**	+	**0**	**-**	+	**0**
s_{11}	**0**	**0**	**-**	**0**	+	**-**
s_{12}	**0**	**-**	**0**	+	+	**0**
s_{13}	**-**	**0**	**0**	**-**	**0**	+
s_{14}	**-**	**0**	**0**	**-**	**0**	+
s_{15}	+	**0**	**0**	+	**0**	+
s_{16}	+	**0**	**0**	+	**0**	+

Table 1. The gauge output changes under each applied force/torque

Substituting the data in table 1 into equation (1) yields the outputs of the sensor under six axis force/torques, shown in Table 2. Table 2 indicates that in theory there is no any coupled interference among six axis force/torques in the sensor, which implies the novel elastic body is mechanically decoupled.

	Applied force/torques					
	F_x	F_y	F_z	M_z	M_x	M_y
U_{Fx}	$4K_1s_6$	0	0	0	0	0
U_{Fy}	0	$4K_2s_4$	0	0	0	0
U_{Fz}	0	0	$4K_3s_1$	0	0	0
U_{Mz}	0	0	0	$4K_4s_7$	0	0

Table 2. Outputs of the sensor

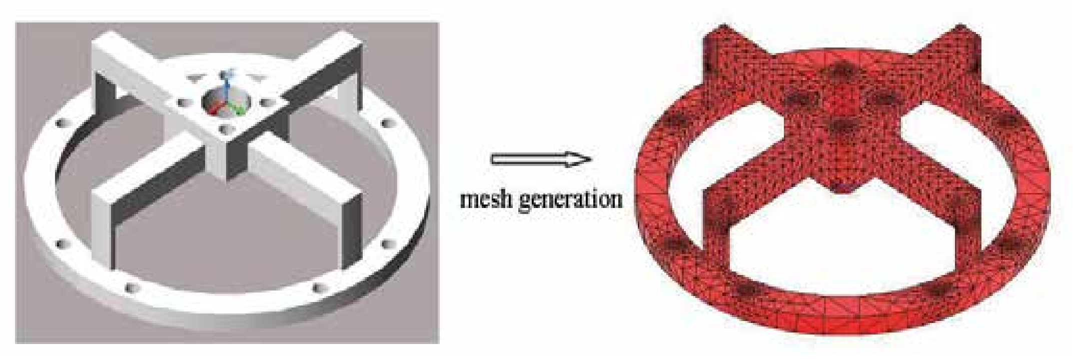
6. Coupled interference analysis by using Finite Element Method

Finite Element Analysis Method (FEM) as the name implies can be used for exact analysis of the elasticity problems. We use the commercial FEM software called ANSYS, produced by ANSYS Corporation, USA, to analyze the coupled interference of the new 4 DOF force/torque sensor.

6.1 Finite element model of the elastic body

The discretization of the domain into sub-regions is the first of a series of steps that must be performed for FEM. The subdivision is usually called mesh generation, and a finite number

of sub-domains are called elements. The discretization of the body involves the decision as to the element number, size and shape of sub-regions used to model the real body.
We discrete the elastic body of the 4 DOF force/torque sensor into sub-regions by using ANASYS software. Here, the element type is set as SOLID95 high-precision element available in the ANSYS, which is much suitable for analysis of bending and twisting of the elastic beam. And the Smart-Size function of the ANSYS is used for mesh generation control. Figure 11 shows a FEM model of the elastic body of the sensor with 49720 element nodes and 28231 elements after mesh generation.



(a) (b)
Fig. 11. Discretization of the elastic body into sub-regions. (a) elastic body of the 4 DOF force/torque sensor, (b) finite element model of the elastic body

The material of the elastic body is aluminium with the parameters as follows:
Young's modulus is 72×10^9 Pa, Poisson ratio is 0.33, and density is 2.78×10^3 kg/m^3.
The size of the elastic body is shown in Table 3.

	Cross elastic beam	Compliant beam	Center support
length (mm)	l=21	h=7	14
width (mm)	b=4.5	b=4.5	14
thickness (mm)	t=4.5	d=1.3	9.5

Table 3. Size of the elastic body

6.2 Strain analysis under six axes force/torques
(1) bound condition set
The elastic body is fixed on the shell of the force/torque sensor through eight bolts on the base, so the connection between them can be regarded as rigid connection. Therefore the total degree of freedom of the base of the elastic body can be set as zero.
(2) applied force/torques
Each single one of the six axis force/torques is applied to the elastic body through its center, respectively. When a single force or torque is applied to the elastic body, the overall

deformation of the elastic body is easy to calculate by using the ANSYS software. What we care about is the strain outputs at the 16 points on the cross beam, to which the 16 strain gauges are stuck, shown in Figure 7. In section 5, we have assumed that s_1, s_2, ……, s_{16} are strain outputs of 16 strain gauges, respectively. The strain of the tensile surface of the beam is defined as positive strain, and the strain of the compressed surface is defined as negative strain.

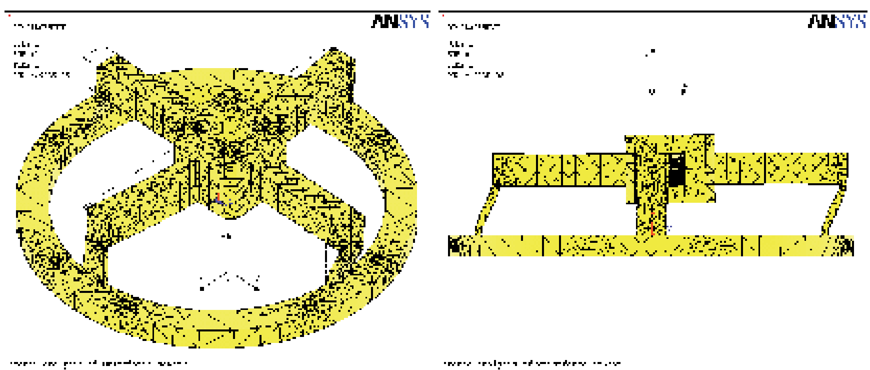
The measurement range of the analyzed 4 DOF force/torque sensor is designed as F_x =±20N, F_y =±20N, F_z =±20N, M_z =±20×4.5 $N.mm$, respectively.

Because the structure of elastic body is symmetric, the strain circumstance under the single force F_x, is similar to that of F_y, and the strain circumstance under the single torque M_x is similar to that of M_y. For simplification of analysis, we only analyze the strain outputs under each one of the force/torques F_y, F_z, M_z, M_x, respectively.

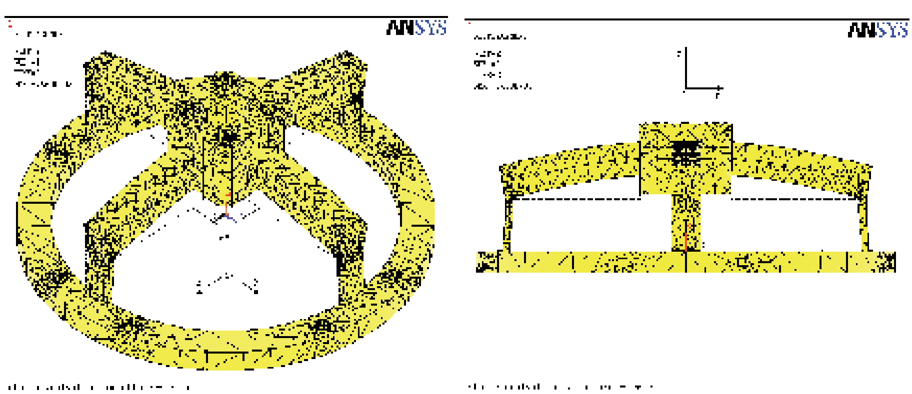
For the convenience of FEM analysis, the applied force/torques to elastic body are chosen as the maximum 20N or 20×4.5 N.mm.

(3) analysis results of FEM

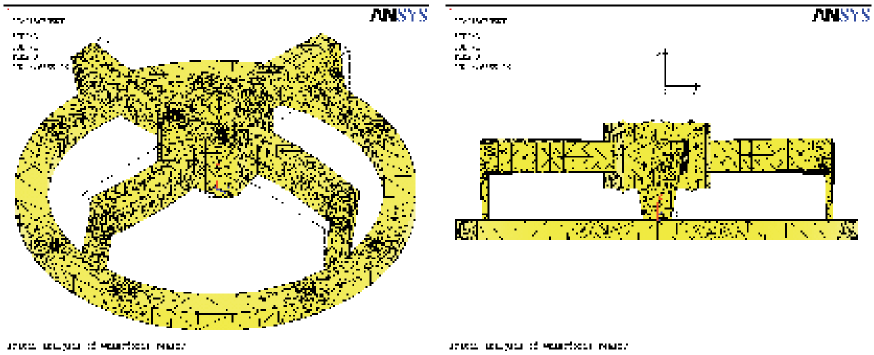
We apply single force F_y=20N, F_z=20N, and single torque M_z=20×4.5$N.mm$, M_x=20×4.5$N.mm$ on the sensor, respectively. The deformations of the elastic body under each single force/torque calculated by the FEM software are shown in Figure 12, and the strain outputs are seen in the Table 4.



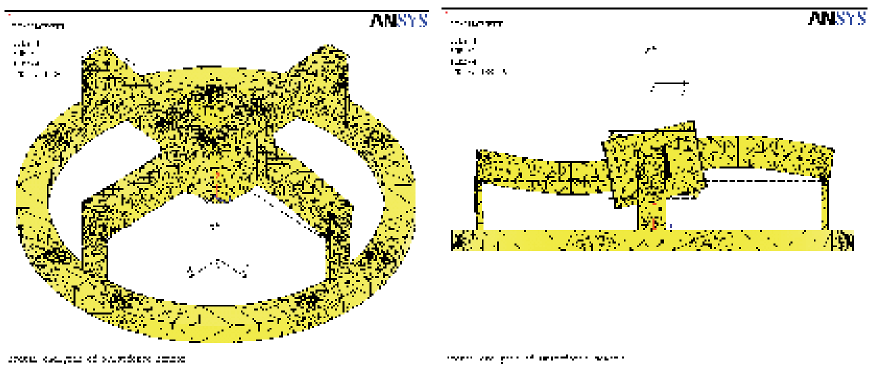
(a) F_y=20N



(b) F_z=20N



(c) M_z=20×4.5 N.mm



(d) M_x =20×4.5 N.mm

Fig. 12. Deformation of the elastic body under each single force/torque

	F_y=20N	F_z=20N	M_z=90 N.mm	M_x=90N.mm
s_1	0.29	76.16	0.02	0.67
s_2	-74.06	1.20	-13.84	3.08
s_3	0.39	-71.94	0.01	0.60
s_4	74.03	1.30	13.83	3.14
s_5	-4.1	1.07	-10.34	0.22
s_6	-4.1	1.16	-13.84	0.24
s_7	-4.02	0.92	10.42	0.22
s_8	-4.02	0.97	13.91	0.25
s_9	1.07	74.20	-0.19	0.75
s_{10}	73.84	1.12	-13.82	3.11
s_{11}	0.48	-74.66	-0.08	0.72
s_{12}	-73.82	0.81	13.77	3.00
s_{13}	4.04	1.21	-10.54	-0.17
s_{14}	4.04	1.27	-13.60	-0.20
s_{15}	4.05	1.08	10.65	-0.18
s_{16}	4.05	1.16	13.55	-0.21

Table 4. The strain outputs under each single force/torque

6.3 Coupled error analysis of the 4 DOF force/torque Sensor

Substituting the strain outputs under each single force/torque in Table 4 into the equation (10) yields the output matrix as

$$\begin{bmatrix} U_{Fx} \\ U_{Fy} \\ U_{Fz} \\ U_{Mz} \end{bmatrix} = \begin{bmatrix} 295.75K_1 & -0.07K_1 & 0.08K_1 & -0.02K_1 & 0.17K_1 & -0.6K_1 \\ -0.07K_2 & 297.5K_2 & 0.31K_2 & 0.17K_2 & -0.02K_2 & 0.08K_2 \\ 0.49K_3 & 0.49K_3 & 296.96K_3 & 0.01K_3 & 0.01K_3 & -0.1K_3 \\ 0.09K_4 & 0.09K_4 & -0.28K_4 & 41.95K_4 & -0.01K_4 & -0.01K_4 \end{bmatrix} \begin{bmatrix} F_x \\ F_y \\ F_z \\ M_x \\ M_y \\ M_z \end{bmatrix} \tag{11}$$

Therefore, the coupled interference under each single force/torque can be easy to calculated. Under a single force F_y=20N, the coupled interference caused by F_y are calculated as follows

$$\begin{aligned} Er(F_x \mid F_y) &= K_1(s_6 + s_{16} - s_8 - s_{14}) = -0.07K_1 \\ Er(F_z \mid F_y) &= K_3(s_1 + s_9 - s_3 - s_{11}) = 0.49K_3 \\ Er(M_z \mid F_y) &= K_4(s_7 + s_{15} - s_5 - s_{13}) = 0.09K_4 \end{aligned}$$

Under a single force F_z=20N, the coupled interference caused by F_z are calculated as follows

$$\begin{aligned} Er(F_x \mid F_z) &= K_1(s_6 + s_{16} - s_8 - s_{14}) = 0.08K_1 \\ Er(F_y \mid F_z) &= K_2(s_4 + s_{10} - s_2 - s_{12}) = 0.31K_2 \\ Er(M_z \mid F_z) &= K_4(s_7 + s_{15} - s_5 - s_{13}) = -0.28K_4 \end{aligned}$$

Under a single torque M_z=20×4.5 $N.mm$, the coupled interference caused by M_z are calculated as follows

$$\begin{aligned} Er(F_x \mid M_z) &= K_1(s_6 + s_{16} - s_8 - s_{14}) = -0.6K_1 \\ Er(F_y \mid M_z) &= K_2(s_4 + s_{10} - s_2 - s_{12}) = 0.08K_2 \\ Er(F_z \mid M_z) &= K_3(s_1 + s_9 - s_3 - s_{11}) = -0.1K_3 \end{aligned}$$

Under a single torque M_x=20×4.5 $N.mm$, the coupled interference caused by M_x can be calculated as follows

$$\begin{aligned} Er(F_x \mid M_x) &= K_1(s_6 + s_{16} - s_8 - s_{14}) = -0.02K_1 \\ Er(F_y \mid M_x) &= K_2(s_4 + s_{10} - s_2 - s_{12}) = 0.17K_2 \\ Er(F_z \mid M_x) &= K_3(s_1 + s_9 - s_3 - s_{11}) = 0.1K_3 \\ Er(M_z \mid M_x) &= K_4(s_7 + s_{15} - s_5 - s_{13}) = -0.01K_4 \end{aligned}$$

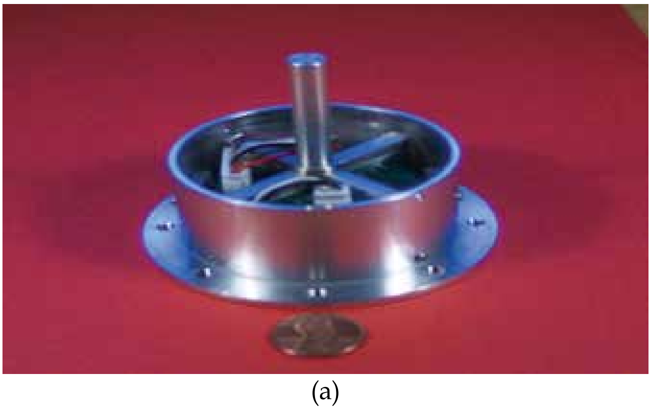
For each axis force/torque measurement, the maximum error caused by coupled interference from other 5 axes is usually expressed as a percentage of full scale.

$$Er(F_x) = \frac{|Er(F_x \mid F_y)| + |Er(F_x \mid F_z)| + |Er(F_x \mid M_x)| + |Er(F_x \mid M_y)| + |Er(F_x \mid M_z)|}{full \quad scale \quad of \quad F_x}$$
$$= \frac{K_1(0.07 + 0.08 + 0.02 + 0.17 + 0.60)}{K_1(74.06 + 74.03 + 73.84 + 73.82)} = 0.32\% F.S.$$
$$Er(F_y) = \frac{|Er(F_y \mid F_x)| + |Er(F_y \mid F_z)| + |Er(F_y \mid M_x)| + |Er(F_y \mid M_y)| + |Er(F_y \mid M_z)|}{full \quad scale \quad of \quad F_y}$$
$$= \frac{K_2(0.07 + 0.31 + 0.17 + 0.02 + 0.08)}{K_2(74.06 + 74.03 + 73.84 + 73.82)} = 0.22\% F.S.$$
$$Er(F_z) = \frac{|Er(F_z \mid F_x)| + |Er(F_z \mid F_y)| + |Er(F_z \mid M_x)| + |Er(F_z \mid M_y)| + |Er(F_z \mid M_z)|}{full \quad scale \quad of \quad F_z}$$
$$= \frac{K_3(0.49 + 0.49 + 0.01 + 0.01 + 0.10)}{K_3(76.16 + 71.94 + 74.20 + 74.66)} = 0.37\% F.S.$$
$$Er(M_z) = \frac{|Er(M_z \mid F_x)| + |Er(M_z \mid F_y)| + |Er(M_z \mid F_z)| + |Er(M_z \mid M_x)| + |Er(M_z \mid M_y)|}{full \quad scale \quad of \quad M_z}$$
$$= \frac{K_4(0.09 + 0.09 + 0.28 + 0.01 + 0.01)}{K_4(10.42 + 10.65 + 10.34 + 10.54)} = 1.14\% F.S.$$

Thus, the maximum coupled error of the 4 DOF force/torque sensor is 1.14%F.S. The analysis results of the FEM, which show the proposed elastic body has merit of low coupled interference, are consistent with the theory analysis results in section 3.

7. Calibration test results

Figure 13 shows the prototypes of 4 DOF force/torque sensor fabricated in our Lab, which is designed with force measurement range ±20N, and torque measurement range ±20×4.5 *N.mm*.



(a)

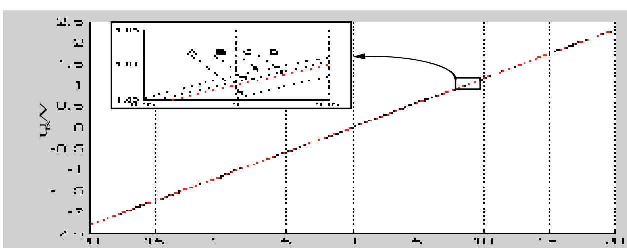
(b)

Fig. 13. Photographs of the 4 DOF force/torque sensors. (a) inner mechanical structure and circuits of the 4 DOF force/torque sensor, (b) two prototypes of 4 DOF force/torque sensor

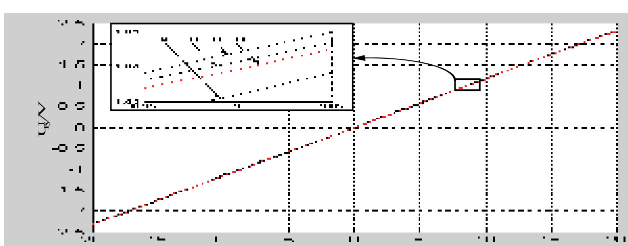
The relationship between the measurand and the output signal is usually obtained by calibration tests.

The calibration procedure of the force/torque sensor is performed as follows. We apply single one of the 4 DOC force/torques on the sensor with a series of values changed from the minimum to the maximum, respectively. And in the meantime, we set other five axes force/torques to be fixed values. After a round of measurement of 4 DOC force/torques, we set the other five axes force/torques to be new fixed values, and do the same measurement again.

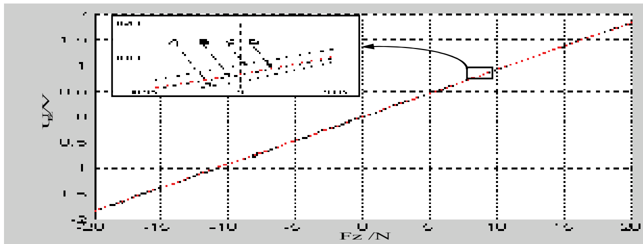
Figure 14 shows the calibration test results. Here, when one axis force/torque is calibrated, the other five axes force/torques are set at zero, half of their full scale values, full scale values, respectively.



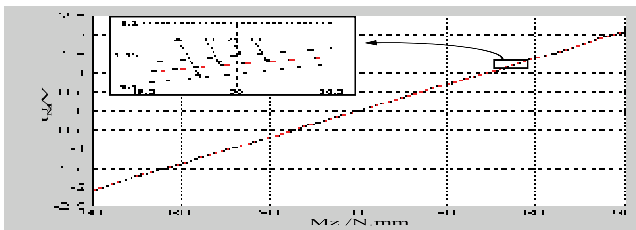
(a)



(b)



(c)



(d)

Fig. 14. Calibration test results. (a) the other five axes force/torques are set at zero, (b) the other five axes force/torques are set at half of full scale values, (c) the other five axes force/torques are set at full scale values, (d) fit curve of average measured values

The calibration results indicate the error of F_x measurement is 0.5%F.S., the error of F_y measurement is 0.5%F.S., the error of F_z measurement is 0.7%F.S., and the error of M_z measurement is 1.3%F.S. Thus the measurement error of the 4 DOF less than 1.5%F.S. (The measurement error of existing commercial 6 DOF force/torque sensors is usually large than 10% F.S. if without decoupling matrix calculation). Although the above error consists not only the coupled interference, but also some other interferences such as error of strain gauge sticking, circuit noise, etc., it is clear that the calibration test results are well correspond with the analysis results of FEM.

The result of impulse response experiment on the 4 DOF force sensor indicates its bandwidth is 210 Hz [Qin, 2004]. Although it is lower than that of the commercial 6 DOF force sensor, it completely meets the bandwidth requirement of HapHCI (>100Hz).

8. Conclusion

A new type multi-dimensional force sensor design for HapHCI is described in this chapter. We build the dynamics model of the force sensor in the HapHCI and give the general principles of force/torque sensor design for HapHCI. According to the proposed design principles, a novel 4 DOF force/torque sensor for HapHCI is developed, which is designed to measure three axis forces Fx, Fy, Fz, and one axis torque Mz by ignoring the other two axis torques. In this chapter, the mechanical structure of the 4 DOF force/torque sensor is presented, and the strain of the elastic body is analyzed in theory and by FEM analysis software ANSYS, respectively. The FEM analysis and calibration results show the new force/torque sensor has low cross sensitivity without decoupling matrix calculation, which means the new force/torque sensor is mechanically decoupled. This new 4 DOF force/torque sensor can be made much smaller owing to the number of the glued strain gauges is greatly reduced and is easier to construct with much lower cost than the existing commercial force/torque sensors. It is well suitable for measuring multi-dimensional interactive force between human hand and interaction device in HapHCI systems.

9. Acknowledgement

This work was supported by National Basic Research and Development Program of China (No.2002CB312102), National Nature Science Foundation of China (No.60775057), and 863 High-Tec Plan of China (No. 2006AA04Z246).

10. References

Anderson, R.J., Spong M.W. (1989). Bilateral control of teleoperators with time delay. *IEEE Transactions on Automatic Control*, 34, 5, 494-901.

Emplus Corporation (1991). Emplus Technical Note: RIXEN EFS-202

Gabriel, R. (2006). The importance of the sense of touch in virtual and real environments. *IEEE MultiMedia*, 7, 24-30.

Huang, W.Y., Jiang, H.M., Zhou, H.Q. (1993). Mechanical analysis of a novel six-degree-of freedom wrist force sensor. *Sensor and Actuators A: Physical*, 35, 203-208

Kaneko, M. (1993). Twin-head type six-axis force sensors. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Yokohama, Japan, pp.26-30.

Kim, G.S. (2001). The design of a six-component force/moment sensor and evaluation of its uncertainty. *Measurement Science and Technology*, 12, 1445-1455.

Kim, G.S. (2007). Development of a three-axis gripper force sensor and the intelligent gripper using it. Sensors and Actuators A: Physical, 137, 2, 213-222.

Lawrence, D.A. (1993). Stability and transparency in bilateral teleoperation. *IEEE Transactions on Robotics and Automation*, 9, 5, 624-637.

Lord Corporation (1985). Force/torque wrist sensing systems. *Technical Note F/T series*. 6-12.

Lorenze, W.A., Peshkin, M.A., Colgate, J.E. (1999). New sensors for new applications: force sensors for human/robot interaction. In *Proceedings of IEEE International Conference on Robotics and Automation*, Detroit, MI, USA, pp.2855-2860.

Nagarajan, R., Sazali, Y., Muralindran, (2003). A design methodology of wrist force sensor for a robot with insufficient degree of freedom. In *Proceedings of IEEE Sensors*, Toronto, Ont., Canada, 22-24 Oct, pp.578-583

Nakamura, Y., Yoshikawa, T., Futamata, I. (1987). Design and signal processing of six axis force sensor. In *Proceedings of 4th International Symposium on Robotics Research*, Santa Barbara, the MIT press, Cambridge, MA

Qin, G., (2004). Research on a novel multi-dimensional wrist force/torque sensor, *Master thesis*, Southeast University, China, (in Chinese)

Song, A.G., Wu, J., Qin G., Huang, W.Y. (2007). A novel self-decoupled four degree-of-freedom wrist force/torque sensor. *Measurement*, 40, 883-891.

Watson, P.C., Drake S.H. (1975). Pedestal and wrist force sensors for automatic assembly. In *Proceedings of 5th International Symposium on Industrial Robot*, pp.501-512.

16

Softness Haptic Display Device for Human-Computer Interaction

Aiguo Song, Jia Liu, Juan Wu
Department of Instrument Science and Engineering, Southeast University
P.R.China

1. Introduction

In the field of virtual reality and teleoperation, haptic interaction between human operator and a computer or telerobot plays an increasingly important role in performing delicate tasks, such as robotic telesurgery, virtual reality based training systems for surgery, virtual reality based rehabilitation systems (Dario et al, 2003) (Taylor, Stoianovici, 2003) (Popescu, et al, 2000), etc. These applications call for the implementation of effective means of haptic display to the human operator. Haptic display can be classified into the following types: texture display, friction display, shape display, softness display, temperature display, etc. Previous researches on haptic display mainly focused on texture display (Lkei et al, 2001), friction display (Richard, Cutkosky, 2002) and shape display (Kammermeier et al, 2000). Only a few researches dealt with softness display, which consists of stiffness display and compliance display. The stiffness information is important to the human operator for distinguishing among different objects when haptically telemanipulating or exploring the soft environment. Some effective softness haptic rendering methods for virtual reality have already been proposed, such as a finite-element based method (Payandeh, Azouz, 2001), a pre-computation based method (Doug et al, 2001), etc. An experimental system for measuring soft tissue deformation during needle insertions has been developed and a method to quantify needle forces and soft tissue deformation is proposed (Simon, Salcudean, 2003). However, there are no effective softness haptic display devices with a wide stiffness range from very soft to very hard for virtual reality yet. The existing PHANToM arm as well as some force feedback data-gloves are inherently force display interface devices, which are unable to produce large stiffness display of hard object owing to the limitation of output force of the motors.

This chapter focuses on the softness haptic display device design for human-computer interaction (HCI). We firstly review the development of haptic display devices especially softness haptic display devices. Then, we give the general principles of the softness haptic display device design for HCI. According to the proposed design principles, a novel method to realize softness haptic display device for HCI is presented, which is based on control of deformable length of an elastic element. The proposed softness haptic display device is composed of a thin elastic beam, an actuator for adjusting the deformable length of the beam, fingertip force sensor, position sensor for measuring the movement of human fingertip, and USB interface based measurement and control circuits. By controlling deformable length of the elastic beam, we can get any desirable stiffness, which can tracks the stiffness of a virtual object with wide range from very soft to hard, to display to a fingertip of human operator. For the convenience of user, a portable softness haptic display device is also developed, which is easy to be connected with a mouse. At last, we build a softness haptic human-computer interaction demo system, which consists of a computer with softness virtual environment, softness haptic modelling element, and the proposed softness haptic display device.

2. Review of haptic display device development

Haptic display devices (or haptic interfaces) are mechanical devices that allow users to touch and manipulate three-dimensional objects in virtual environments or tele-operated systems. In human-computer interaction, haptic display means both force/tactile and kinesthetic display. In general, haptic sensations include pressure, texture, softness, friction, shape, thermal properties, and so on. Kinesthetic perception, refers to the awareness of one's body state, including position, velocity and forces supplied by the muscles through a variety of receptors located in the skin, joints, skeletal muscles, and tendons. Force/tactile and kinesthetic channels work together to provide humans with means to perceive and act on their environment (Hayward et al, 2004).

One way to distinguish among haptic devices is their intrinsic mechanical behavior. Impedance haptic devices simulate mechanical impedance —they read position and send force. Admittance haptic devices simulate mechanical admittance — they read force and send position. Being simpler to design and much cheaper to produce, impedance-type architectures are most common. Admittance-based devices are generally used for applications requiring high forces in a large workspace (Salisbury K., Conti F., 2004).

Examples of haptic devices include consumer peripheral devices equipped with special motors and sensors (e.g., force feedback joysticks and steering wheels) and more sophisticated devices designed for industrial, medical or scientific applications. Well-known commercial haptic devices are the PHANToM series from Sensable Technology Corporation, and the Omega.X family from Force Dimension Corporation. These haptic devices are impedance driven.

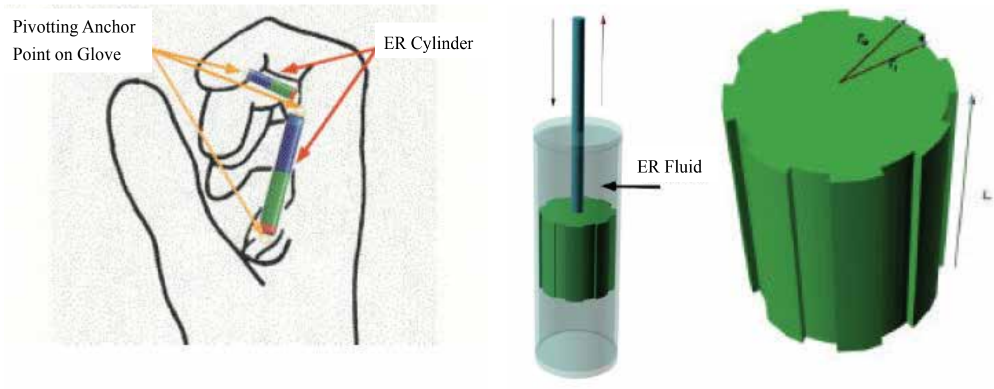
Fig. 1. The PHANTOM desktop device

Fig. 2. The Omega.X device

In recent years, different research groups have developed laboratory prototypes of haptic display devices based on different principles. Haptic display devices previously developed explore servomotors (Wagner et al, 2002), electromagnetic coils (Benali-Khoudja et al, 2004), piezoelectric ceramics (Pasquero, Hayward, 2003) (Chanter, Summers, 2001) (Maucher et al, 2001), pneumatics (Moy et al, 2000), shape memory alloys (SMA) (Kontarinis et al, 1995) (Taylor, Creed, 1995) (Taylor et al,1997) (Taylor, Moser, 1998), Electro-magnetic (Fukuda et al,1997) (Shinohara et al, 1998), polymer gels (Voyles et al, 1996) and fluids as actuation technologies (Taylor et al, 1996).
A softness haptic display is important to distinguish between the different objects. This haptic information is essential for performing delicate tasks in virtual surgery or tele-surgery. However, at present only a few literatures have researched on the softness display device design. The existing softness display device design approaches can be divided into four categories of approaches as follows.

2.1 Softness haptic display device based on electro-rheological fluids

Mavroidis et al developed a softness haptic display device that could enable a remote operator to feel the stiffness and forces at remote or virtual sites (Mavroidis et al, 2000). The device was based on a kind of novel mechanisms that were conceived by JPL and Rutgers University investigators, in a system called MEMICA (remote Mechanical Mirroring using Controlled stiffness and Actuators) which consisted of a glove equipped with a series of electrically controlled stiffness (ECS) elements that mirrors the stiffness at remote/virtual sites, shown in Figure 3. The ECS elements make use of Electro-Rheological Fluid (ERF), which was an Electro-Active Polymer (EAP), to achieve this feeling of stiffness. The miniature electrically controlled stiffness (ECS) element consisted of a piston that was designed to move inside a sealed cylinder filled with ERF. The rate of flow was controlled electrically by electrodes facing the flowing ERF while inside the channel. To control the stiffness of the ECS, a voltage was applied between electrodes that are facing the slot and the ability of the liquid to flow was affected.



(a) MEMICA system (b) ECS element and its piston

Fig. 3. Softness haptic display device based on electro-rheological fluids

2.2 Softness haptic display device based on the fingertip contact area control

It has been reported that softness in the cutaneous sense can be produced by controlling contact area corresponding to contact force (Fujita et al, 2000).

Fujita and Ikeda developed a softness haptic display device by dynamically controlling the contact area (Ikeda, Fujita, 2004) (Fujita, Ikeda, 2005). The device consisted of the pneumatic contact area control device and the wire-driven force feedback device, shown in Figure 4. The contact area was calculated using Hertzian contact theory using the Young's modulus, which is converted from the transferred stiffness. The air pressure to drive the pneumatic contact area control device was controlled using the pre-measured device property. The reaction force was calculated based on the stiffness using Hook's law.

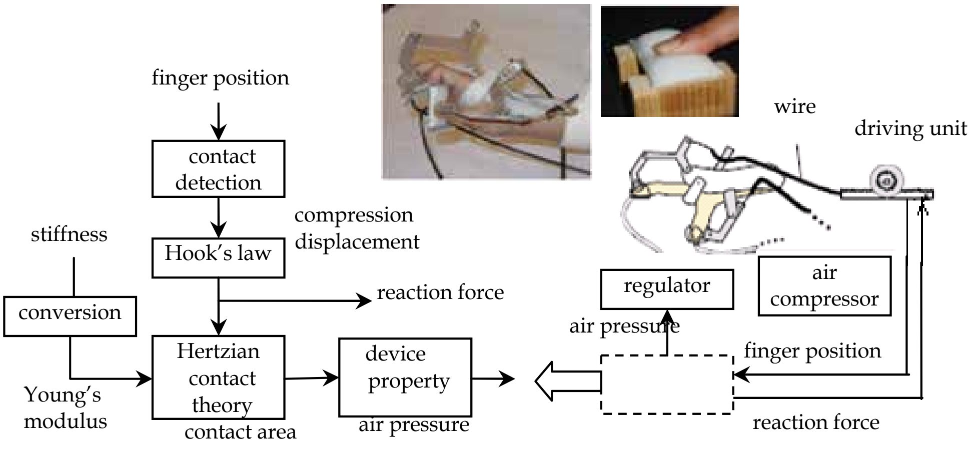


Fig. 4. Fingertip contact area control system

Fujita and Ohmori also developed a softness haptic display device which controlled the fingertip contact area dynamically according to the detected contact force, based on the human softness recognition mechanism (Fujita, Ohmori, 2001). A fluid-driven vertically moving cylinder that had rubber sheet at its top surface was utilized, because of the simplicity of development and the spacial resolution as shown in Figure 5. The piston of the device was installed on a loadcell for contact force detection. The inside of the piston was designed as empty, and fluid was pumped into the piston through the pipe at the side wall of the piston. The pumped fluid flaws out from twelve holes at the top of the piston, and the fluid push-up the rubber-top cylinder. Because the center of the rubber is pushed by the fingertip, the peripheral part is mainly pushed up. Therefore the contact area between the fingertip and the rubber increases. The pressure distribution within the contact area becomes constant because of the intervention of the fluid. The softness was represented as the increase rate of the contact area. The fluid volume control pump consisted of a motor-driven piston, a cylinder and a potentiometer to detect the piston position. The fluid volume in the device was indirectly measured and controlled by controlling the piston position of the pump. A DC servo control circuit was utilized for the pump control.

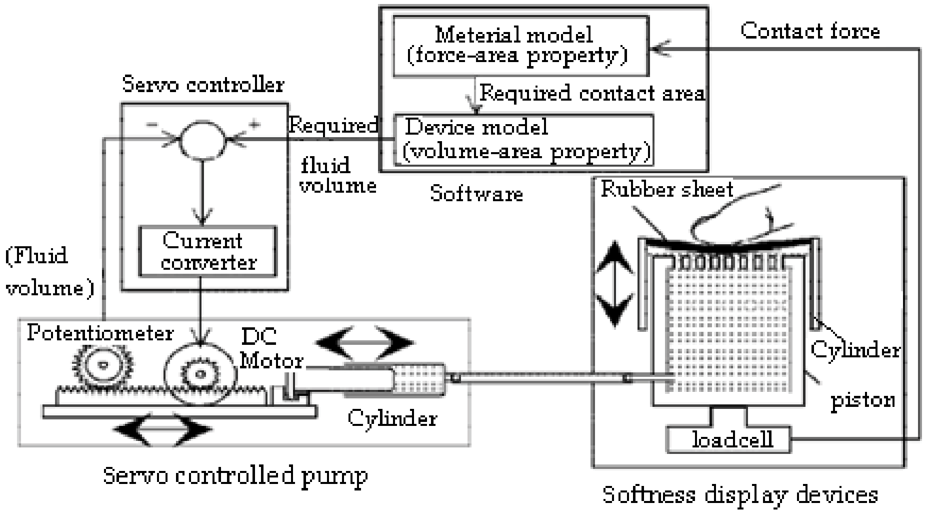


Fig. 5. Softness display system by controlling fingertip contact area based on detected contact force

Fig. 6. Close-up the device and the finger

2.3 Softness haptic display device based on pneumatic array

Moy et al at University of California presented a softness haptic display device using pneumatically actuator, which consisted of two parts, the contact interface and the pneumatic valve array of tactor elements (Moy et al, 2001), Shown in Figure 7. A 5x5 array of tactor elements were spaced 2.5 mm apart and were 1 mm in diameter. The working frequency was 5 Hz. The contact interface was molded from silicone rubber in a one-step process. Twenty-five stainless steel pins were soldered to the back of the baseplate. Silicone tubing was placed around each of the pins. The silicone rubber bonds with the silicone tubing to form an airtight chamber. The contact interface was connected to the pneumatic valve array by hoses and barbed connectors. The pulse width modulated (PWM) square wave controlled the pressure in the chamber.

Fig. 7. The softness haptic display attached to the finger

2.4 Softness haptic display device based on elastic body

Takaiwa and Noritsugu at Okayama University developed a softness haptic display device that can display compliance for human hand aiming at the application in the field of virtual reality (Takaiwa, Noritsugu, 2000). Pneumatic parallel manipulator was used as a driving mechanism of the device, consequently, which yielded characteristic that manipulator worked as a kind of elastic body even when its position/orientation was under the control.

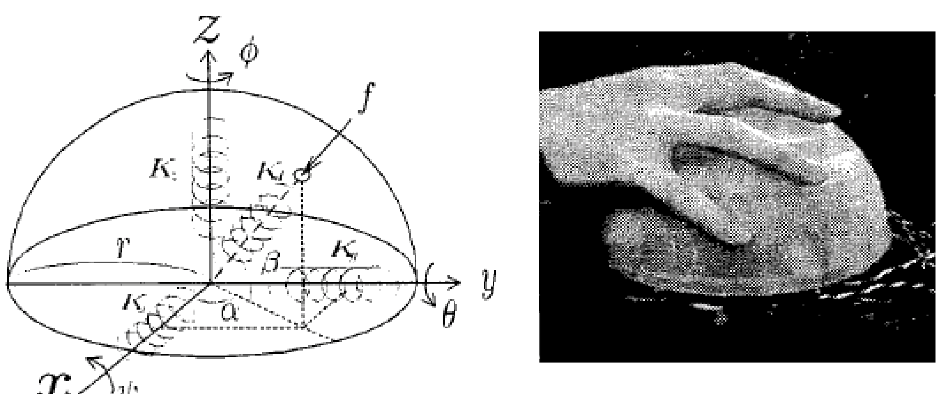


Fig. 8. The softness haptic display device based on elastic body

3. The general principles of the softness haptic display device for HCI

Each of these approaches has its own advantages/disadvantages. Humans use two different forms of haptic display devices: active and passive. Active haptic display devices have joints with motors, hydraulic actuators, or some other form of actuator that creates motion, adds energy, and reflects virtual forces. Passive haptic display devices have brakes or dampers that provide the user with feedback forces. The passive haptic display devices cannot force a user in a certain direction - it can only prevent or slow a user's motion. The benefit of a passive haptic display device over an active haptic display device is that force spikes generated by the virtual environment cannot do any damage to the human operator. Electrorheological (ER) fluids suspensions show swift and reversible rheological changes when the electric or magnetic field is applied. However, there are such defects as a restriction on usable temperatures so as to avoid evaporation or freezing of the water, an extreme increase in the electric current flow as the temperature raises, inferior stability caused by transfer of water, etc. The method based on the fingertip contact area control is easy to implement. However to different objects, confirming the relation between the dynamic changes of contact area and stiffness needs lots of psychophysiological experiments, and real time contact area control with high precision is difficult to guarantee. Pneumatically actuated haptic display devices have to overcome leakage, friction and non-conformability to the finger.

In this section we present four principles of designation of the softness haptic display devices as follow.

(a) Because the active haptic display devices are unable to produce very high stiffness, and the large force directly provided by the active element, such as electric motors, pneumatic drivers, hydraulic drivers, etc., sometimes may be harmful to the human operator. Passive haptic display devices are recommended for safety.

(b) The softness haptic display devices must be able to produce continuous stiffness display in wide range.

(c) The softness haptic display devices should be controlled accurately and rapidly.

(d) The size and weight are very important to the softness haptic display device design. To guarantee the high transparency of the softness haptic human-computer interaction system, small size and light weight is required. It is necessary to seek a portable haptic display device that can be taken easily.

4. A novel softness display device designation method

The environment dynamics is usually expressed by a mass-spring-damp model as follows:

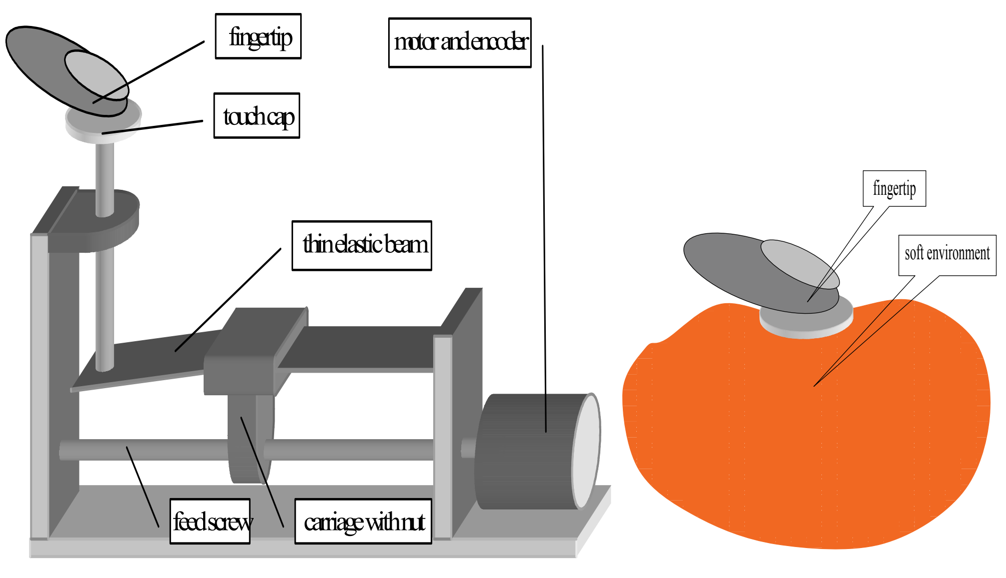
$$f_e = m_e \ddot{x}_e + b_e \dot{x}_e + k_e x_e \tag{1}$$

where f_e is force acted on the environment, x_e is displacement of the environment, and m_e, b_e, k_e are mass, damp and stiffness of the environment, respectively. As to the soft environment discussed here, the displacement x_e represents local deformation of its surface, and m_e represents the local mass of its surface, which is relatively very small and usually can be omitted. If the damp is notable and the stiffness is small, the soft object is characterized by the compliance. If the reverse is the case, the soft object is characterized by the stiffness.

In this chapter, our research mainly focuses on the stiffness display, because for a lot of soft objects, such as most of the tissues of human body, stiffness is not only inherent, but also notable by comparison with damp or viscous. So that how to replicate the sense of stiffness to the user as if he directly touches with the virtual or remote soft environment is a primary issue in the softness display of the virtual environment and of the teleoperation.

We design and fabricate a novel haptic display system based on control of deformable length of an elastic element (CDLEE) to realize the stiffness display of the virtual environment, which is shown in schematic form in Figure 9(a). It consists of a thin elastic beam, feed screw, carriage with nut, and motor. The stiffness of the thin elastic beam is the function of deformable length of the beam l seen in Figure 10. So the stiffness can be easily and smoothly changed to any value by controlling the deformable length of the thin beam l. Here, a motor, together with a feed screw and a nut, is used to control the position of the carriage, which determines the deformable length l.

In ideal case, when the human operator's fingertip pushes or squeezes the touch cap of the softness haptic display interface device, he will feel as if he directly pushes or squeezes the soft environment with a small pad, seen in Figure 9(b).



(a) (b)

Fig. 9. Softness display of virtual soft environment

Figure 10 shows the principle of the softness display based on CDLEE. Where, y is vertical displacement of the end of the thin elastic beam when force f acted on that point.

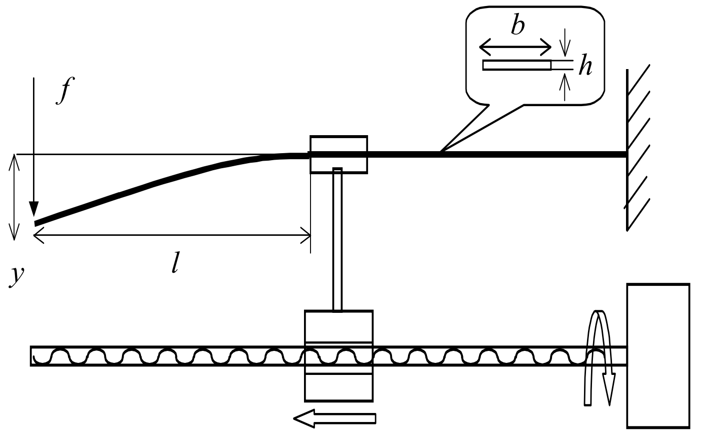


Fig. 10. Principle of the softness display based on CDLEE

According to the theory of Mechanics of materials, the deformation of the thin elastic beam under the force *f* can be given as:

$$y = \frac{fl^3}{3EI} \tag{2}$$

where *E* is Young's modulus, and *I* is moment of inertia of the thin elastic beam.

$$I = \frac{bh^3}{12} \tag{3}$$

b and *h* are width and thickness of the thin elastic beam, respectively. Substituting equation (3) into equation (2) gives:

$$y = \frac{4fl^3}{Ebh^3} \tag{4}$$

Thus, the stiffness of the thin elastic beam, which is felt by the human fingertip at the touch cap of the device, can be expressed by an elastic coefficient as

$$k = \frac{f}{y} = \frac{Ebh^3}{4l^3} = \rho\frac{1}{l^3} \tag{5}$$

$\rho = \dfrac{Ebh^3}{4}$ is the gain of the stiffness. Equation (5) shows the stiffness at the free end of the cantilever k is proportional to the third power of reciprocal of the deformable length l, which indicates that the stiffness k can be changed with wide range as l is changed.
Differentiating both sides of equation (5) with respect to time yields stiffness change ratio as

$$r_k = \frac{dk}{dt} = \frac{dk}{dl}\frac{dl}{dt} = -3\rho\frac{1}{l^4} \times v_{motor} \tag{6}$$

From the above formula, we know r_k is proportional to the fourth power of reciprocal of the deformable length l, which indicates that the stiffness k can be changed very quickly as l is changed, especially when $l \rightarrow 0$, $r_k \rightarrow \infty$. Therefore the above formula means the ability of real time stiffness display based on CDLEE in our device.

5. Position control for real time softness display

Section 4 implies the key issue of the real time softness display actually is how to realize the real time position control of the carriage, which determines the deformable length l of the elastic beam. Here, PD controller is employed for the real time position control. The control structure for the real time softness display is seen in Figure 11.

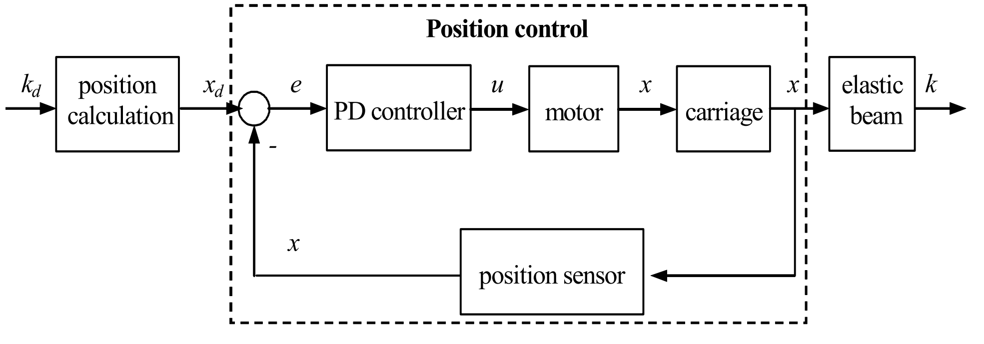


Fig. 11. Control structure for real time softness display

where k_d is a destination stiffness to display, which comes from the virtual or remote soft environment. x_d is a destination position of the carriage, which equals to the destination deformable length of the thin elastic beam l_d. Rewriting equation (5), we have

$$l_d = \sqrt[3]{\rho / k_d} \tag{7}$$

$$x_d = l_d = \sqrt[3]{\rho / k_d} \tag{8}$$

ρ can be estimated by calibrating the stiffness change with respect to the deformable length of the thin elastic beam l. To simplify the estimation of ρ, let $z = 1/l^3$, and substitute it into equation (5), so that the power function in equation (5) can be transformed into a linear function as

$$k = \rho \cdot z \tag{9}$$

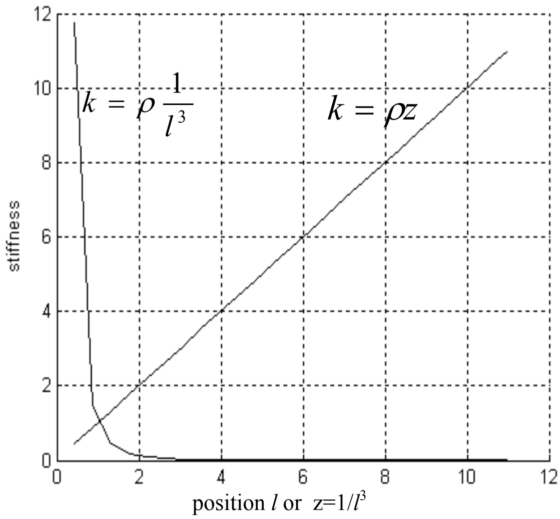


Fig. 12. Transform the power function into linear function

LMS method is used to estimate the parameter ρ as follows

$$\frac{\partial \sum_{i=1}^{n} |e_i|^2}{\partial \hat{\rho}} = 0 \tag{10}$$

where e_i is error of each measurement point.

$$e_i = k_i - \hat{\rho} \cdot z_i \qquad i = 1,\Lambda ,n \tag{11}$$

where k_i is the ith measurement value of stiffness at the ith point z_i.
So that,

$$\sum_{i=1}^{n}(k_i - \hat{\rho}\cdot z_i)z_i = 0$$

$$\hat{\rho} = \frac{\sum_{i=1}^{n} k_i z_i}{\sum_{i=1}^{n} z_i^2} \tag{12}$$

The PD controller used here for position control of the carriage can be expressed as

$$u = K_p e + K_d \frac{de}{dt} \tag{13}$$

$$e = x_d - x \tag{14}$$

where K_p is proportional control gain, K_d is differential control gain, and e is error between the destination position x_d and the current real position x.

6. Real time softness haptic display device

The real time stiffness display interface device based on CDLEE method is shown in Figure 13, which is composed of a thin elastic beam, a motor with an encoder, feed screw, carriage with nut, force sensor, position sensor, and a touch cap.
The material of the thin elastic beam in the stiffness display interface device is spring steel, whose Young's modulus of elasticity is $E = 180 \times 10^9 N / m^2$. The size of the thin elastic beam is set as 80mm long × 0.38mm thick × 16.89mm wide.
Substituting the above parameters into equation (5) can yield the minimum stiffness of the device:

$$k_{\min} = 0.1287 \times 10^3 N / m = 0.13 N / mm$$

K_{min} is the minimum stiffness of the softest object. Thus, the stiffness display range of the device is from 0.13×10^3N/m to infinite, which almost covers the stiffness range of soft tissues in human body.

The position of the carriage is measured by an encoder with resolution of 8000 CPR. The displacement of the touch cap, which equals to the deformation of the end point of the thin elastic beam, is measured by a resistance based position sensor with 1% linearity. And the force acted by a fingertip on the touch cap is measured by a full bridge arrangement of resistance strain gauges with 0.05N accuracy. The range of up-down movement of the touch cap when human fingertip jiggles it is from 0 to 2 cm.

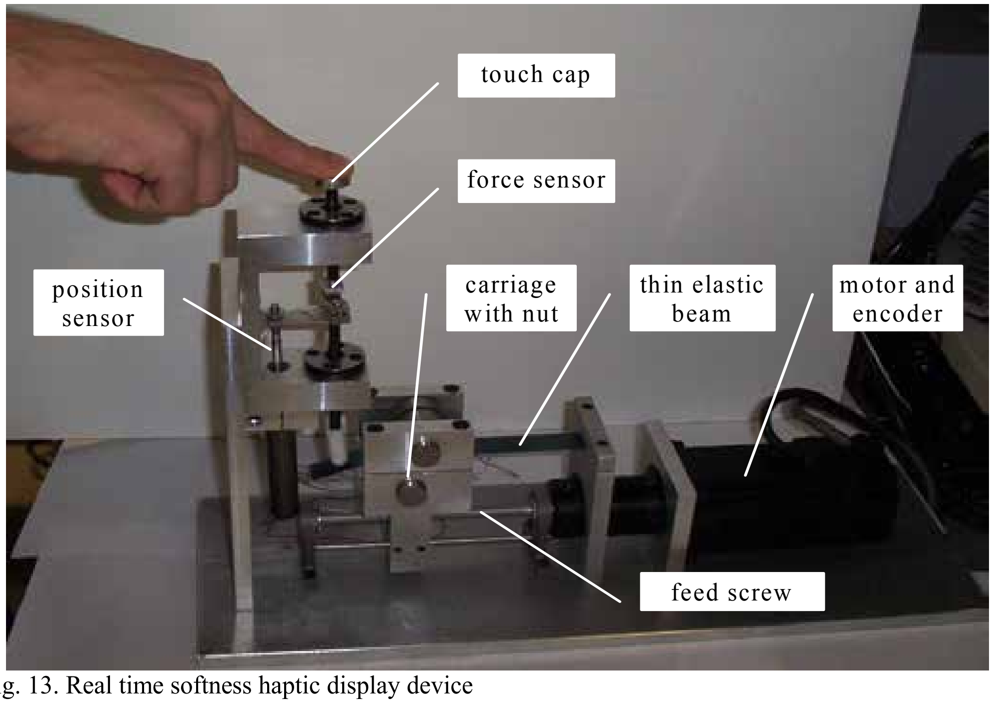


Fig. 13. Real time softness haptic display device

7. Calibration results

The results of stiffness calibration of the softness haptic display device are shown in Figure 14. According to equation (12), the fitting curve of the relation between stiffness and $z = 1/l^3$ is shown in Figure14, and the $\hat{\rho}$ is estimated as

$$\hat{\rho} = 4.05 \times 10^4 (N \cdot mm^2)$$

The Figure 14 and Figure 15 demonstrate the validity of the equation (5), although there exists some difference between experimental curve and fitting curve. The difference mainly comes from the effect of friction between the cantilever beam and the carriage, and from the effect of nonlinear property when the length of the cantilever beam becomes small and the ratio of end point deformation to the length of the cantilever beam becomes large.

In order to overcome the bad effects of friction and nonlinear property so as to control the deformable length of the thin elastic beam precisely, we make a table to record the relationship between the stiffness and the deformable length of the beam point by point based on calibration data. And a table-check method is used for transforming a destination stiffness to a destination length of the cantilever beam.

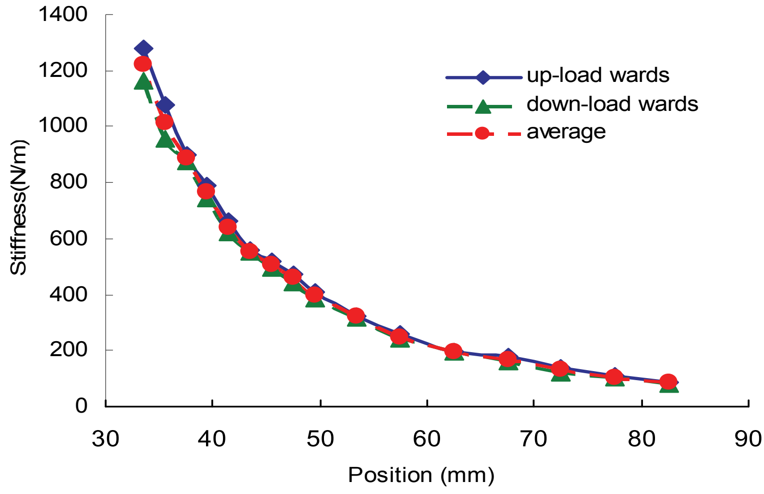


Fig. 14. Results of stiffness calibration

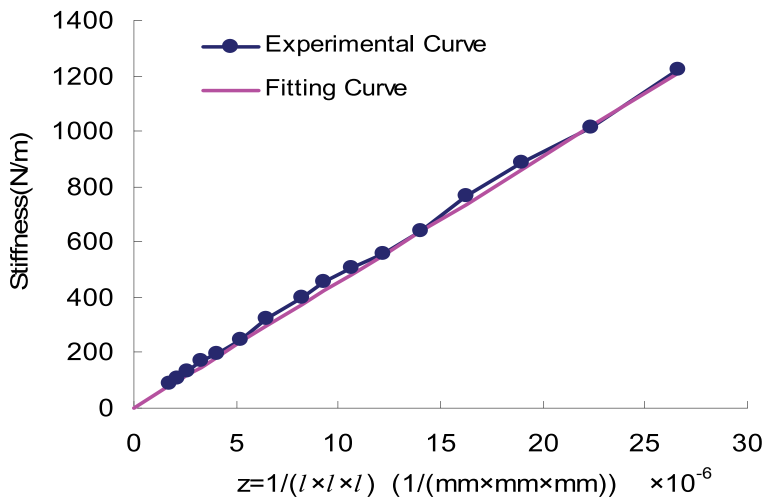


Fig. 15. Fitting curve of characteristic of stiffness

The result of the position control of the carriage is shown in Figure 16. Here, the proportional control gain and the differential control gain of the PD controller are set as

$$K_p = 5 \times 10^{-3}$$

$$K_d = 1.6 \times 10^{-4}$$

The above setting is based on experience and some experiment results.

Figure 16 implies the control of deformable length of the thin elastic beam is real time control.

The trajectory of stiffness display which tracks the destination stiffness change of a virtual soft object is shown in Figure 17. Note that the destination stiffness is set as step square pulses, which corresponds to the typical change of stiffness of some soft tissues with blood vessels beneath the surface.

The stiffness display experiment results demonstrate that the stiffness display interface device is able to replicate the stiffness of the virtual soft object quickly and accurately.

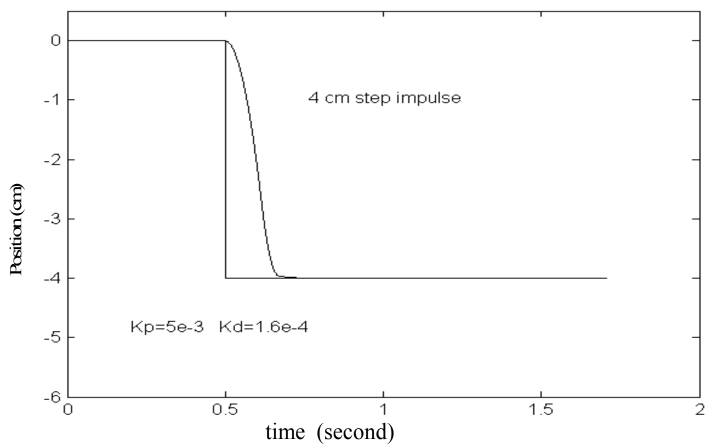


Fig. 16. Position control result

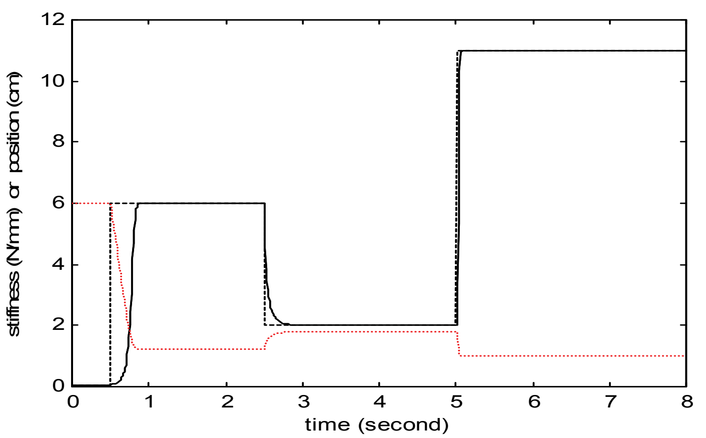


Fig. 17. Stiffness display experiment results. The solid line represents displayed stiffness, the dashed line represents destination stiffness, and the dotted line represents position of the carriage controlled by PD controller.

8. Portable softness display device

During the past decade, many haptic display devices have been developed in order to address the somatic senses of the human operator, but only a few of them have become widely available. There are mainly two reasons for that. Firstly, the costs of devices are too

expensive for most people to afford. Secondly, most of the devices are not easy to carry around. It is necessary to seek a more efficient implementation in terms of cost, performance and flexibility.
Based on the softness display device proposed in section 7, a new low-cost, truly lightweight and highly-portable softness haptic display device is presented shown in Figure 18. This device can be easily carried in the user's hand with compact dimensions (10cm x 7cm x 15 cm). Its total expense is less than 150 US Dollars. Thus it will encourage people to use haptic devices.
The material of the elastic thin beam is spring steel, whose Young's modulus of elasticity is $E=180\times10^9 N/m^2$. The size of the thin elastic beam is chosen as 9 mm long, 1 mm thick, and 0.3 mm wide. The stiffness display range of this device is from 25N/m to 1500N/m.
The position of the carriage is measured by a step motor. The displacement of the touch cap, which is equal to the deformation of the end point of the thin elastic beam, is measured by a Hall Effect position sensor fixed under the touch cap with 0.1 mm accuracy. And the force applied by a human fingertip on the touch cap is measured by a touch force sensor fixed on the top of the touch cap with 9.8 mN accuracy.
The most important advantage of this device is that a computer mouse can be assembled at the bottom of the device conveniently. Two shafts are designed and installed on each side of the touch cap and contact to the left and right mouse buttons, respectively, which is used for transferring the press of human fingertip to the left and right mouse buttons, respectively, so the human finger is easy to control the left and right mouse buttons when he use the portable softness haptic display device. The device is a good interface that succeeded to combine both pointing and haptic feature by adding stiffness feedback sensation.

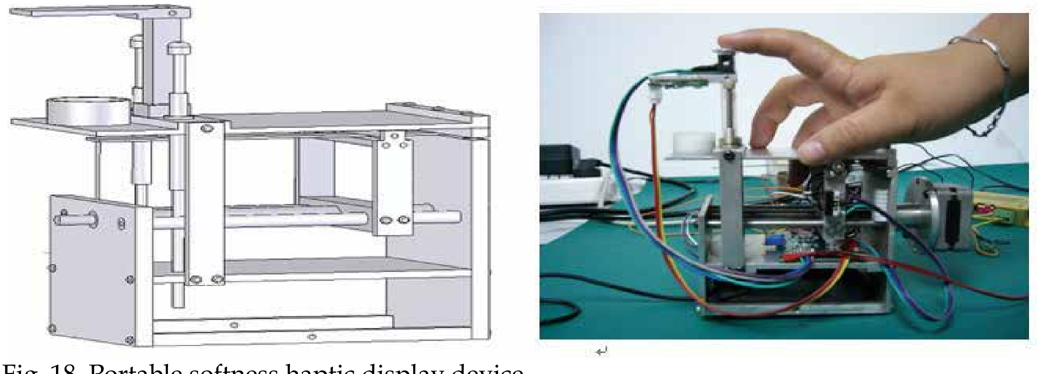


Fig. 18. Portable softness haptic display device

9. Softness haptic human-computer interaction demo system

Most human–computer interaction systems have focused primarily on the graphical rendering of visual information. Among all senses, the human haptic system provides unique and bidirectional communication between humans and their physical environment. Extending the frontier of visual computing, haptic display devices have the potential to increase the quality of human-computer interaction by accommodating the sense of touch. They provide an attractive augmentation to visual display and enhance the level of understanding of complex data sets. In case of the palpation simulator, since the operator wants to find an internal feature of the object by touching the object, the haptic information is more important than the visual information. In this section, we construct a softness haptic human-computer interaction demo system by using the softness haptic display device. The haptic human-computer interaction system is shown in Figure 19, which provides visual and haptic feedback synchronously allowing operators to manipulate objects in the virtual environment. The virtual environment consists of 3D virtual object models, a visual feedback part and a stiffness feedback part.

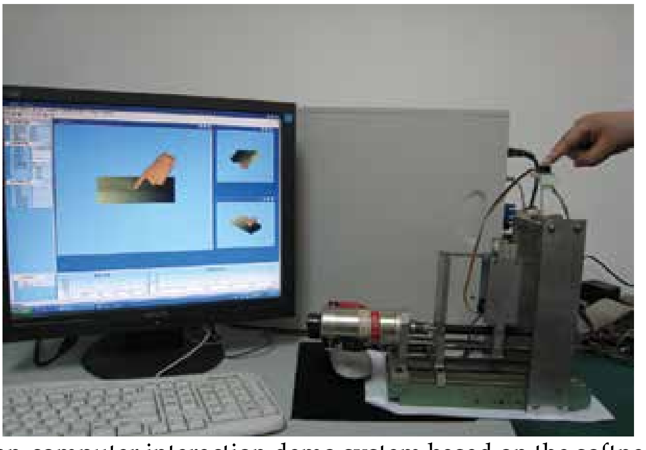


Fig. 19. Haptic human-computer interaction demo system based on the softness haptic display device

The software of the demo system is implemented by Visual C + + MFC and OpenGL programming based on MVC (Model-View-Controller) pattern. The MVC pattern divides an interactive application into three parts. The model contains the core functionality and data. Views display information to the user. Controllers handle user input. Views and controllers together comprise the user interface. A change propagation mechanism ensures consistency between the user interface and the model. Figure 20 illustrates the basic Model-View-Controller relationship. The purpose of the MVC pattern is to separate the model from the view so that changes to the view can be implemented or even additional views created, without affecting the model.

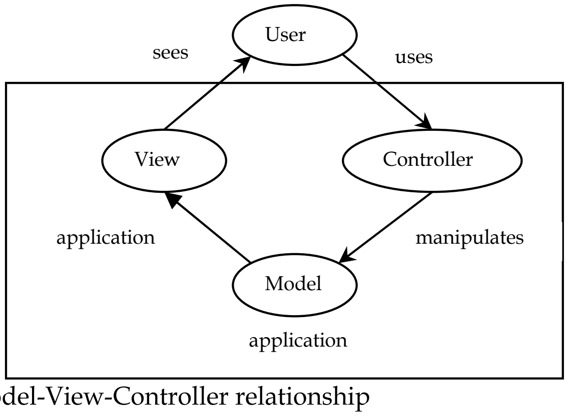


Fig. 20. The basic Model-View-Controller relationship

A 3D virtual model plays an important role in many simulators. Due to the computational burden, the main type of virtual objects for various stimulators is a surface model. We adopt a shortcut method of three dimensional simulated realization combining OpenGL programming technology and 3DS MAX software. The simulated surfaces are divided into small triangles. The Gauss deformation model is used to simulate the deformation of virtual objects. Figure 21 shows the sequence diagram of the system.

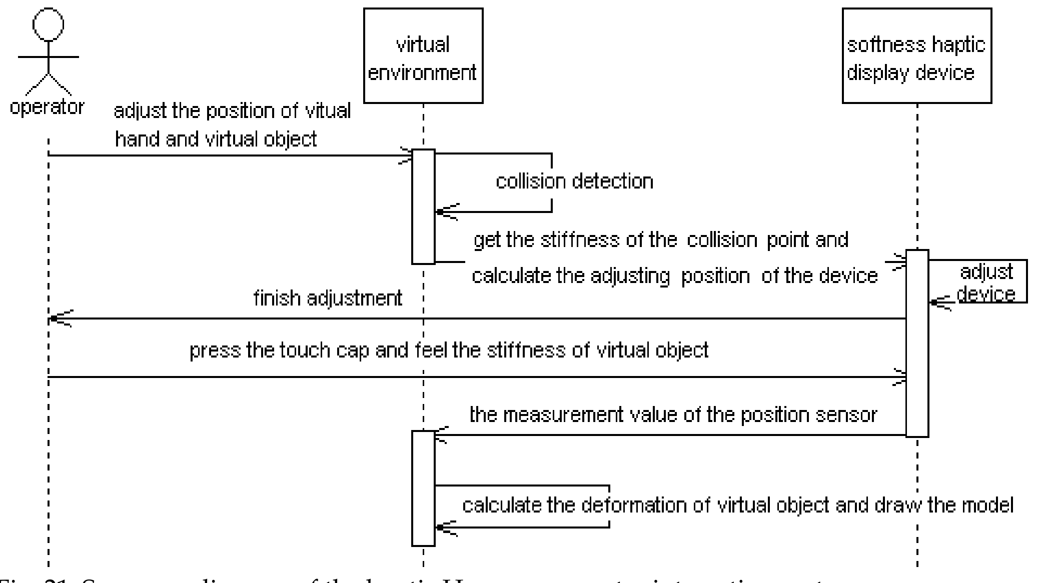


Fig. 21. Sequence diagram of the haptic Human-computer interaction system

A human operator controls the position of the virtual hand by mouse and keyboard. When the virtual hand contacts with the virtual object, the stiffness of the virtual object at the touch point is calculated and fed back to the softness display haptic device. Then by controlling the elastic beam deformable length based on PD controller, its stiffness tracks the stiffness of a virtual object, which is directly felt by the fingertip of human operator. The up-down displacement of the operator's fingertip is measured by the position sensor as command to control the movement of virtual fingertip up-down. At the same time, the deformation of the virtual object is calculated by deformation algorithm. The human operator could feel the stiffness of the virtual object via a softness haptic display device and observe a real time graphics in the screen simultaneity.

We use two virtual objects for simulation. A virtual cube with different stiffness distribution (nonhomogeneous object) in the surface is modeled using 5600 triangular meshes with 3086 nodes. And a liver with same stiffness distribution (homogeneous object) has 6204 triangular meshes with 3104 nodes. Figure 22 and Figure 23 show the deformation simulation. According to the softness haptic model, when a virtual hand finger contacts with the virtual object, the softness haptic display device is able to replicate the stiffness of the virtual object quickly and accurately.

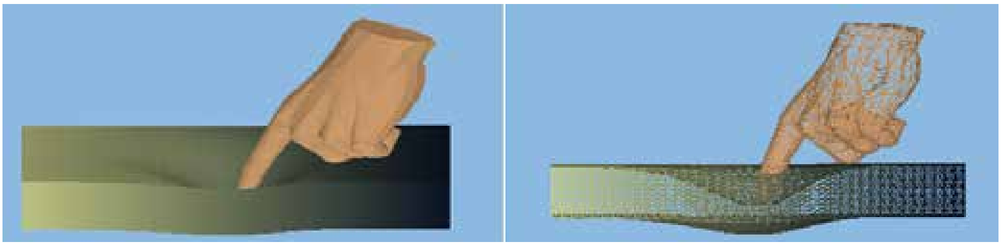


Fig. 22. Deformation simulation of a virtual soft cube with different stiffness distribution

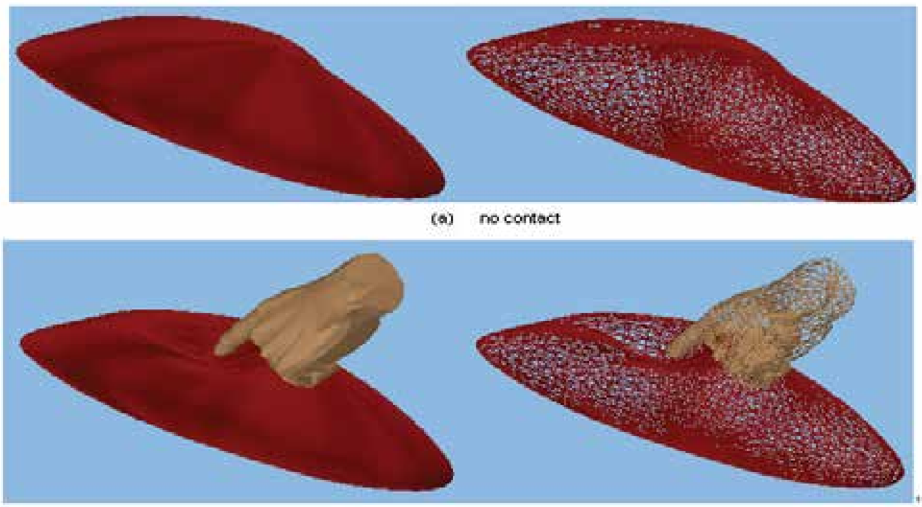


Fig. 23. Deformation simulation of a virtual liver with constant stiffness distribution

To establish the realism to the human operator, the softness haptic display device must be kept operating at 100Hz at least. But an acceptable refresh rate for stable visual feedback is 30Hz. This can be accomplished by running different threads with different servo rates. In our program, three main threads exist. The visual-rendering thread is typically run at rates of up to 30 Hz. The acquisition thread is run as fast as possible congruent with the simulated scene's overall complexity. A collision-detection and deformation thread, which computes a local representation of the part of the virtual object closest to the user avatar (e.g. virtual hand), is run at slower rates to limit CPU usage.

10. Conclusion

This chapter reviews the development of haptic display devices especially softness haptic display devices, and give the general principles of the softness haptic display device design for HCI. According to the proposed design principles, a novel method based on control of deformable length of elastic element (CDLEE) to realize the softness haptic display for HCI is proposed. The proposed softness haptic display device is composed of a thin elastic beam and an actuator to adjust the deformable length of the beam. The deformation of the beam under a force is proportional to the third power of the beam length. By controlling the deformable length of the beam, we can get the desirable stiffness quickly. And a portable softness haptic display device is also developed, which is convenient to be connected with a mouse. The softness haptic human-computer interaction demo system based on the proposed device demonstrates the softness haptic display device is well suitable for haptic human-computer interaction.

11. Acknowledgement

This work was supported by National Basic Research and Development Program of China (No.2002CB312102), National Nature Science Foundation of China (No.60775057), and 863 High-Tec Plan of China (No. 2006AA04Z246). Thanks to Prof. J. Edward Colgate, Prof. Michael A. Peshkin, Mr. Mark Salada and Mr. Dan Morris for their good advice.

12. References

Benali-Khoudja, M., Hafez, M., Alexandre, J.M., Kheddar, A., Moreau, V., (2004). VITAL: A New Low-Cost Vibro-TActiLe Display System, *Proceedings of IEEE International Conference on Robotics and Automation*, pp.721-726, ISBN: 0-7803-8232-3, New Orleans, USA, April-May 2004

Bicchi,A.; De Rossi,D. & Scilingo,P.(2000). Haptic Discrimination of Softness in Teleoperation: The Role of the Contact Area Spread Rate, *Robotics and Automation*, vol.16, no.5, 496-504

Chanter, C.M., Summers, I.R. (2001). Results from a tactile array on the fingertip. In *Proceedings of Eurohaptics*, pp.26-28

Dario, P., Hannaford, B., Menciassi, A. (2003). Smart surgical tools and augmenting devices. *IEEE Transaction on Robotics and Automation*, 19, 5, 782-792

Darwin, G., Caldwell, N., Tsagarakis, C.G. (1999). An integrated tactile/shear feedback array for stimulation of finger mechanoreceptor. In *Proceedings of the 1999 IEEE International Conference on Robotics & Automation*, Detroit, Michigan, pp.187-192

Doug, L.J, Dinesh, K.P. (2001). A unified treatment of elastostatic contact simulation for real time haptics. *Haptics-e*, 2, 1, (http://www.haptics-e.org)

Fujita, K., Ohmori, H. (2001). A new softness display interface by dynamic fingertip contact area control. In *Proceedings of the 5th World Multiconference on Systemics, Cybernetics and Informatics*, pp.78-82

Fujita, K., Ikeda, Y. (2005). Remote haptic sharing of elastic soft objects. http://reality.ei.tuat.ac.jp/papers/haptic05.pdf

Fukuda, T., Morita, H., Arai, F., Ishihara, H., Matsuura, H. (1997). Micro resonator using electromagnetic actuator for tactile display. In *Proceedings of International Symposium on micromechatronics and human science*, pp.143-148

Hayward, V., Astley O.R., Manuel C.H. (2004). Haptic interfaces and devices. http://www.roblesdelatorre.com/gabriel/VH-OA-MC-DG-GR-04.pdf

Ikeda, Y., Fujita, K. (2004). Display of soft elastic object by simultaneous control of fingertip contact area and reaction force. *Transactions of the Virtual Reality Society of Japan*, 9, 2, 187-194

Kammermeier, P., Buss, M., Schmidt, G. (2000). Dynamic display of distributed tactile shape information by a prototypical actuator array. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pp.1119-1124

Kontarinis, D.A., Son, J.S., Peine, W., Howe, R.D. (1995). A tactile sensing and display system for teleoperated manipulation. In *Proceedings of the IEEE International Conference on Robotics & Automation*, pp. 641-646

Lkei, Y., Yamada, M., Fukuda, S. (2001). A new design of haptic texture display and its preliminary evaluation. In *Proceedings of Virtual Reality Annual International Symposium*, pp.21-28

Maucher, T., Meier, K., Schemmel, J.(2001). An interactive tactile graphics display. In *Proceedings of International Symposium on Signal Processing and its Applications*, pp.190-193, ISBN: 0-7803-6703-0, Kuala Lumpur, Malaysia

Mavroidis, C., Pfeiffer, C., Celestino, J., Bar-Cohen, Y. (2000). Controlled compliance haptic interface using electro-rheological fluids. In *Proceedings of SPIE - The International Society for Optical Engineering*, pp.300-310

Moy, G., Wagner, C., Fearing, R. (2000). A compliant tactile display for teletaction. In *Proceedings. of IEEE International Conference on Robotics and Automation*, pp.3409-3415, Piscataway, USA

Payandeh, S., Azouz, N. (2001). Finite elements, mass-spring-damper systems and haptic rendering. In *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pp.224-229

Pasquero, J., Hayward, V. (2003). STReSS: A practical tactile display system with one millimeter spatial resolution and 700 Hz refresh rate, In *Proceedings of Eurohaptics 2003*, pp.94-110, Dublin, Ireland

Popescu, V.G., Burdea, G.C., Bouzit, M., Hentz, V.R. (2000). A virtual-reality-based tele-rehabilitation system with force feedback. *IEEE Transaction on Information Technology in Biomedicine*, 4, 1, 45-51

Richard, C., Cutkosky, M.R. (2002). Friction modeling and display in haptic applications involving user performance. In *Proceedings of IEEE International Conference on Robotics and Automation*, pp.605-611

Salisbury, K., Conti, F., Barbagli, F. (2004). Haptic rendering: introductory concepts, *Computer Graphics and Applications*, 24, 2, 24- 32

Shinohara, M., Shimizu, Y., Mochizuki, A. (1998). Three-dimensional tactile display for the blind. *IEEE Transactions on Rehabilitation Engineering*, 6, 3, 249-256

Simon, P.D., Salcudean, S.E. (2003). Needle insertion modeling and simulation. *IEEE Transaction on Robotics and Automation*, 19, 5, 864-875

Takaiwa, M., Noritsugu, T. (2000). Development of pneumatic human interface and its application for compliance display. In *Proceedings of 26th IEEE Annual Conference on Industrial Electronics*, pp.806-811, ISBN: 0-7803-6456-2, Nagoya, Japan

Taylor, P.M., Creed, A., Sianaki, A.H., Varley, C.J., Moser, A. (1995). Creating the illusion of shape for virtual environments. In *Proceedings of IEE Colloquium on Innovative Actuators for MechatronicSystems*, pp.7/1 -7/7

Taylor, P.M., Hossein, A., Varley, C.J. (1996). An electrorheological fluid based tactile array for virtual environments. In *Proceedings of IEEE International Conference on Robotics and Automation*, pp.18-23, Minneapolis,USA

Taylor, P.M., Moser, A., Creed, A. (1997). The design and control of a tactile display based on shape memory alloys. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp.1/1-1/4

Taylor, P.M., Moser, A., Creed,A. (1998). A sixty-four element tactile display using Shape Memory Alloy wires. *Displays,* 18, 3, 163-168

Taylor, R.H., Stoianovici, D. (2003). Medical robot in computer-integrated surgery. *IEEE Transaction on Robotics and Automation*, 19, 5, 765-781

Voyles, R., Fedder, G., Khosla, P. (1996). Design of a modular tactile sensor and actuator based on an electro-rheological gel. In *Proceedings of IEEE International Conference on Robotics and Automation,* pp.13-17, Minneapolis, USA

Wagner, C., Lederman, S., Howe, R. (2002). A tactile shape display using RC servomotors. In *Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp.354-355, ISBN: 0-7695-1489-8, Orlando, FL, USA

17

3D User Interfaces for Collaborative Work

Frank Steinicke, Gerd Bruder, Klaus Hinrichs, Timo Ropinski
Westfälische Wilhelms-Universität Münster, Institut für Informatik
Einsteinstraße 62, 48149 Münster
Germany
Marco Lopes
TRIDELITY Display Solutions GmbH
Leopoldstrasse 1, 78112 St. George
Germany

Abstract

Desktop environments have proven to be a powerful user interface and are used as the de facto standard human-computer interaction paradigm for over 20 years. However, there is a rising demand on 3D applications dealing with complex datasets, which exceeds the possibilities provided by traditional devices or two-dimensional display. For these domains more immersive and intuitive interfaces are required. But in order to get the users' acceptance, technology-driven solutions that require inconvenient instrumentation, e.g., stereo glasses or tracked gloves, should be avoided. Autostereoscopic display environments equipped with tracking systems enable users to experience 3D virtual environments more natural without annoying devices, for instance via gestures. However, currently these approaches are only applied for specially designed or adapted applications without universal usability. Although these systems provide enough space to support multi-user, additional costs and inconvenient instrumentation hinder acceptance of these user interfaces.

In this chapter we introduce new collaborative 3D user interface concepts for such setups where minimal instrumentation of the user is required such that the strategies can be easily integrated in everyday working environments. Therefore, we propose an interaction system and framework, which allows displaying and interacting with both mono- as well as stereoscopic content in parallel. Furthermore, the setup enables multiple users to view the same data simultaneously. The challenges for combined mouse-, keyboard- and gesture-based input paradigms in such an environment are pointed out and novel interaction strategies are introduced.

1. Introduction

In recent years 3D user interfaces (UIs) have become more and more popular and widespread due to the requirements of several application areas, where two-dimensional desktop systems lack immersive and intuitive interaction. In addition the user's ability to perform complex interaction tasks increased, since bi-manual interactions or six degrees of freedom (DoFs) manipulations do not require much effort and are easy to learn even for non-experts.

Current 3D UIs are technology-driven solutions providing more immersive exploration of and interaction with complex datasets, in particular by using stereoscopic projection and tracked six DoFs input devices. Although the costs for such a setup have reached a moderate level, experts just like ordinary users rarely uses these systems – even when 3D tasks have to be accomplished [3]. One reason for this is the inconvenient instrumentation required allowing immersive interactions, i.e., the user is forced to wear stereo glasses, tracked devices, gloves etc. [12]. Furthermore the most effective ways for humans to interact with synthetic 3D environments have not finally been resolved [3, 6]. Devices that enable control over multiple DoFs simultaneously still involve problems, which are often avoided by the usage of their 2D counterparts – as a matter of fact 2D interactions are performed best with 2D devices [3, 18, 9]. However, while in real life humans are able to move and turn objects freely in a single motion, this natural interaction is absent in two-dimensional interfaces; the user is forced to decompose 3D tasks into several 2D tasks. In addition, shortage of spatial input in typical 3D applications leads to the need to switch modes. This procedure results in ineffectiveness, in particular when switching between manipulation and navigation techniques is required in a repetitive manner. Most desktop-based 3D applications include three-dimensional content in combination with two-dimensional elements for graphical user interface (GUI) interaction. While 3D content usually benefits from stereoscopic display, 2D GUI items often do not require immersive visualization.

For such a system current autostereoscopic (AS) displays can be used to view 3D data stereoscopically without wearing any devices [8]. Thus the user is able to perceive a stereoscopic image in a fixed area called sweet spot. When the AS display features an optical head tracker, the user can even move in front of the display, while the tracking system can be further exploited to allow gesture-based interaction [11]. Even multiple users can view the stereoscopic content in different horizontally neighbouring sweet spots. However, the separation of the stereo half images performed by an AS display (see Section 3.1) influences viewing of monoscopic content in such a way that essential elements of the GUI are distorted. Although some displays allow displaying monoscopic content on the display, simultaneously display of mono- as well as stereoscopic content is not supported. Thus, simultaneous viewing requires an additional conventional display to show the monoscopic content. But only few applications support rendering of a stereoscopic window on a different display. Nevertheless, problems arise from decoupling interaction and visualization; interactions with 2D GUI elements have to be performed on the 2D screen, whereas 3D content is displayed stereoscopically on an AS display.

In this chapter we introduce new collaborative 3D user interface concepts as a solution to the lack of spatial input and intuitive interaction techniques for direct manipulation of mono- as well as stereoscopic content in multi-user desktop environments. We propose an AS display environment and present a framework that enables to display arbitrary shaped areas of the GUI either in a mono- or in a stereoscopic way. Furthermore, the framework allows interaction between both "worlds" and thus opens up new vistas for human-computer interaction (HCI). Hence, the user can interact with any 2D or 3D application via familiar mouse/keyboard devices in combination with natural gestures. The remainder of this chapter is organized as follows. Section 2 summarizes related work. In Section 3 we describe the proposed setup, while Section 4 introduces interaction strategies for such everyday working environments. Section 5 presents implementation details. The results of an experimental evaluation are discussed in Section 6. Section 7 concludes the chapter and gives an overview about future work.

2. Related Work

AS Display Environments In 2000, the Heinrich-Hertz-Institute built an AS display system consisting of a gaze tracker, a head and a hand tracker [11]. The head tracker gives the user a look-around capability, while the gaze tracking activates different applications on the desktop. The hand tracker enables the user to navigate and manipulate objects in 3D space via simple gestures, where computer vision is the major technological factor influencing the type of gesture that are supported. Similar approaches support gesture-based interactions by tracking the users hand and fingers with magnetic fields [24] or optical-based solutions [2]. These approaches rather address tracking technologies than advanced 3D user interfaces. Although, these systems potentially support novel forms of interaction they are restricted to specific applications designed for these setups [2]; simultaneous display of and interaction between mono- and traditional devices with stereoscopic content is not considered.

2.1 Simultaneous Mono- and Stereoscopic Display

Although, current stereo-in-a-window systems [5, 24] show stereoscopic content either in one window time-sequentially or using filtering techniques, these technologies are restricted to only one rectangular window and glasses are still required. Hardware-based approaches have been proposed to display monoscopic and stereoscopic content simultaneously on one AS display [13]. However, interaction concepts have not yet been developed for these displays and these systems only exist as prototype solutions. Due to the lack of simultaneous display most interaction approaches only propose improvements for interactions either in 2D using monoscopic display or in 3D using stereoscopic display, but they do not combine both worlds. The interaction with stereoscopic content using two-dimensional strategies involves further problems, for instance, monoscopic representation of the mouse cursor disturbs stereoscopic perception, whereby precise interactions are impeded.

2.2 Three-dimensional User Interfaces for Individual and Collaborative Work

In recent years, many frameworks have been proposed which extend 2D GUIs for operating systems (OSs) to so called 3D desktops, but also existing OSs evolve to 3D and include depth information [1, 16]. These approaches provide a virtual 3D space in which three-dimensional counterparts replace 2D GUI elements. Hence, more space is available to display further information. Although these environments provide a fancy visualization, it has not been investigated in how far they improve the interaction process, since they force the user to perform 3D interactions where 2D interactions are intended. Due to the mentioned shortcomings of virtual reality (VR) interfaces, hybrid approaches have been proposed which combine 2D and 3D interaction using different display or interaction technologies [4, 21]. For example, Benko et al. have discussed techniques to grab monoscopically-displayed objects from a projection screen in order to view them stereoscopically using a head mounted display [4]. However, an instrumentation of the user is still required.

2.3 Both-handed and Cooperative Interactions

When interacting with the hands numerous factors have to be considered. With respect to the tasks, the hands need to be moved symmetrically or asymmetrically, some tasks can be performed better with the dominant, others with the non-dominant hand. Also the used input devices have a major impact on how bi-manual interactions are performed. For instance, the used devices can be equal (e.g., keyboard and keyboard), different (e.g., mouse and keyboard), and they can support different DoFs or involve constraints.
These approaches are applied in everyday tasks as well as in most user interfaces. Writing on a paper, when one hand holds the pencil while the other clamps the paper, involves asymmetrical interactions. In many computer games the dominant hand using the mouse performs navigation tasks, whereas status changes are accomplished with the non-dominant hand via keyboard shortcuts. Interactions techniques for large-screen displays or VR environments often involve symmetrical bi-manual manipulation in order to scale, or rotate virtual objects. However, the combination of traditional devices and gestures in AS display environments that run ordinary 3D applications has not been considered until now. The aim of this chapter is not to debate the validity of desktop-based interaction concepts – there is no need to throw away 40 years of 2D UI research – neither the benefits of technology-driven VR approaches.
The objective is to explore in how far these concepts can mutually adapt to each other in order to provide efficient interfaces that will be accepted by users as setups for their daily working environments.

3. System Setup for Single and Multi-User Interaction

In this section we present the setup that we believe has the potential to be accepted by the users since natural as well as immersive interactions are supported, whereas instrumentation of the user is avoided.

3.1 Autostereoscopic Display Environment

On current AS displays users can see 3D data without wearing any instruments, for example by using lenticular rasters [8]. The lenticular screen is a plastic sheet molded to have the form of dozens of tiny lenses per inch. This raster operates as a beam splitter and ensures that the pixels displayed in each odd column are seen by the user's left eye, while the pixels displayed in each even column are perceived with the right eye. If the viewer positions her head in certain viewing positions, she perceives a different image with each eye giving a stereo image. To support multiple users there are up to eight different neighbouring sweet spots where users perceive stereoscopic images correctly. When a user leaves a sweet spot slightly to one side, the stereo half images for this user have to be swapped in order to maintain the stereoscopic effect. When the user further moves to the same side, she gets into the next sweet spot and views from the perspective of the neighbouring region.

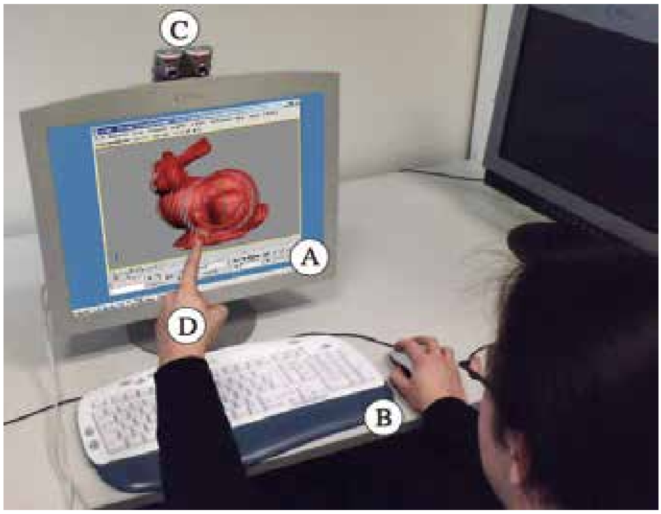


Fig. 1. 3D user interface setup includes (A) an AS display, (B) traditional mouse and keyboard, and (C) stereo-based camera setup. (D) The user applies gestures in order to perform 3D manipulations of a 3D scene.

The separation of the stereo half images influences viewing of monoscopic content in such a way that the most essential elements of the GUI are distorted. Therefore, we have implemented a software framework (see Section 5), which provides full control over the GUI of the OS. Thus, any region or object can be displayed either mono- or stereoscopically. Furthermore, we are able to catch the entire content of any 3D graphics application based on OpenGL or DirectX. Our framework allows changing the corresponding function calls such that visualization can be changed arbitrarily. The interaction performed in our setup is primarily based on mouse and keyboard (see Figure 1). However, we have extended these devices with more natural interfaces.

3.2 Stereo-based Tracking System

AS displays can be equipped with eyes or head tracking systems to automatically adjust the two displayed images and the corresponding raster. Thus, the user perceives a stereo image in a larger region. Vision-based trackers enable non-intrusive, markerless computer vision based modules for HCI. When using computer vision techniques several features can be tracked, e.g., the eyes for head tracking, but it is also possible to track fingers in order to interpret simple as well as intuitive gestures in 3D. Pointing with the fingertip, for example, is an easy and natural way to select virtual objects. As depicted in Figure 1 we use a stereo-based camera setup consisting of two USB cameras each having a resolution of 640 × 480 pixels. They are attached on the top of the AS display in order to track the position and orientation of certain objects. Due to the known arrangement of the cameras, the pose of geometric objects, e.g., user's hands can be reconstructed by 3D reprojection. Besides pointing actions, some simple gestures signalling stop, start, left and right can even be recognized. These gesture input events can be used to perform 3D manipulations, e.g., to rotate or translate virtual objects (see Figure 1). Furthermore, when different coloured fingertips are used even multiple fingers can be distinguished (see Figure 3 (right)).

4. Collaborative 3D User Interface Concepts

Due to the availability of the described setup, traditional input devices can be combined with gesture-based paradigms. There are some approaches that use similar setups in artificial environments consisting of applications exclusively designed or even adapted therefore. Hence, these concepts are not applicable in daily working environments with ordinary applications. With the described framework we have full control over the GUI of the OS, in particular any arbitrarily shaped region can be displayed either mono- or stereoscopically, and each 3D application can be modified appropriately. The implementation concepts are explained in Section 5. In the following subsections we discuss implications and introduce several universal interaction techniques that are usable for any 3D application and which support multiple user environments.

4.1 Cooperative Universal Exploration

As mentioned in Section 3.1 our framework enables us to control any content of an application based on OpenGL or DirectX. So-called display lists often define virtual scenes in such applications. Using our framework enables us to hijack and modify these lists. Among other possibilities this issue allows us to change the viewpoint in a virtual scene. Hence, several navigation concepts can be realized that are usable for any 3D application.

Head Tracking Binocular vision is essential for depth perception; stereoscopic projections are mainly exploited to give a better insight into complex three-dimensional datasets. Although stereoscopic display improves depth perception, viewing static images is limited, because other important depth cues, e.g., motion parallax phenomena, cannot be observed. Motion parallax denotes the fact that when objects or the viewer move, objects which are farther away from the viewer seem to move more slowly than objects closer to the viewer. To reproduce this effect, head tracking and view-dependent rendering is required.

This can be achieved by exploiting the described tracking system (see Section 3.2). When the position and orientation of the user's head is tracked, this pose is mapped to the virtual camera defined in the 3D scene; furthermore the position of the lenticular sheet is adapted. Thus, the user is able to explore 3D datasets (to a certain degree) only by moving the tracked head. Such view-dependent rendering can also be integrated for any 3D application based on OpenGL. This concept is also applicable for multi-user scenarios. As long as each collaborator is tracked the virtual scene is rendered for each user independently by applying the tracked transformation. Therefore, the scene is rendered in corresponding pixels, the tracked transformation is applied to the virtual camera registered to the user.

4.2 Universal 3D Navigation and Manipulation

However, exploration only by head tracking is limited; object rotation is restricted to the available degrees of the tracking system, e.g. 60 degrees. Almost any interactive 3D application provides navigation techniques to explore virtual data from arbitrary viewpoints. Although, many of these concepts are similar, e.g., mouse-based techniques to pan, zoom, rotate etc., 3D navigation as well as manipulation across different applications can become confusing due to various approaches.

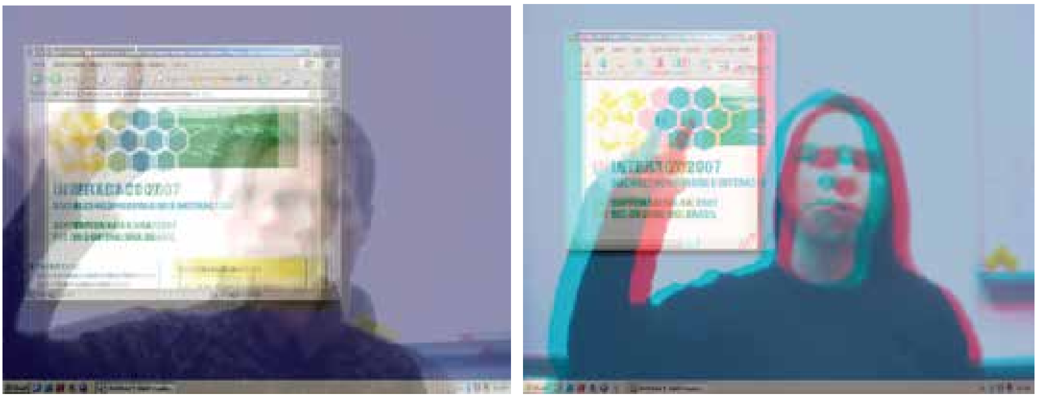


Fig. 2. Screenshot of an AS desktop overlaid with a transparent image of the user in (left) vertical interlaced mode and (right) anaglyph mode.

The main idea to solve this shortcoming is to provide universal paradigms to interact with a virtual scene, i.e., using the same techniques for each 3D application. Therefore, we use gestures to translate, scale, and rotate objects, or to move, fly, or walk through a virtual environment. These techniques are universal since they are applicable across different 3D applications. Moreover, individual strategies supported by each application can be used further on, e.g., by mouse- or keyboard-based interaction. We have implemented these navigational concepts by using gestures based on virtual hand techniques [6]. Therefore, a one-to-one mapping in terms of translational and rotational mappings between the movements of the user's hand and the virtual scene is applied. Thus the user can start an arbitrary 3D application, activate gesture recognition and afterwards, the user can manipulate the scene by the combination of mouse, keyboard and gestures. Other concepts, such as virtual flying, walking etc. can be implemented, for instance, by virtual pointer approaches [6].

4.3 Stereoscopic Facetop Interaction

Besides depth information regarding the user's head and hand pose, we also exploit the images captured by the stereo-cameras mounted on top of the AS display (see Figure 1). Since the cameras are arranged in parallel, while their distance approximates the eye base of ≈ 65mm, both images compose a stereoscopic image of the user. Due to the full control over the GUI, we are able to display both half images transparently into the corresponding columns of the AS display – one image into the even columns, one into the odd ones. Hence, the user sees her image superimposed on the GUI as a transparent overlay; all desktop content can still be seen, but users appear to themselves as a semi-transparent image, as if looking through a window in which they can see their own reflection. This visualization can also be used in order to enable stereo-based face-to-face collaboration. Hence users can see stereoscopic real-time projections of their cooperation partners. The technique of superimposing the user's image on top of the display has been recently used in the Facetop system [21]. More recently, Sony has released the Eyetoy that enables gesture interaction. In both approaches the user is able to perform 3D gestures in order to fulfill 2D interactions on the screen, where a visual feedback is given through captured images of the user. However, besides gesturing multiple DoFs for two-dimensional control, e.g., moving the mouse cursor by pointing, a stereo-based camera setup allows to use multiple DoF to enable 3D interaction. Furthermore, we use the stereoscopic projection of the user. This provides not only visual feedback about the position of the cursor on the screen surface, but also about its depth in order to simplify 3D interaction. A 3D representation of the mouse cursor is displayed at the tracked 3D position. A mouse click might be emulated if the position of the real finger and the visual representation of the finger stereoscopically displayed overlap in space. Alternatively, other gestures might be predefined, e.g., grab gestures. The depth information is also used when interacting with 2D GUIs. When using our framework, a corresponding depth is assigned to each window and it is displayed stereoscopically. In addition shadows are added to all windows to further increase depth perception. When finger tracking is activated, the user can arrange windows on the desktop in depth by pushing or pulling them with a tracked finger. Figure 2 shows screenshots of two stereoscopic facetop interaction scenarios. Each user arranges windows on the desktop by pushing them with the finger. This face-to-face cooperation has the potential to increase performance of certain collaborative interaction that requires cooperation between at least two partners. Figure 3 shows such a procedure for remote collaborative interaction. In Figure 3 (left) two users use the same screen to interact in a co-located way. In Figure 3 (right) two users collaborate remotely. The user wears a red thimble in order to simplify vision-based tracking.

Fig. 3. Illustration of a collaborative interaction setup in which (left) two users collaborate co-locatedly and (right) a user cooperates with another user in a remote way [21].

4.3 Combining Desktop-based and Natural Interaction Strategies

By using the described concepts we are able to combine desktop devices with gestures. This setup is beneficial in scenarios where the user holds a virtual object in her non-dominant hand using universal exploration gestures (see Section 4.1), while the other hand can perform precise interactions via the mouse (see Figure 1). In contrast to use only ordinary desktop devices, no context switches are required, e.g., to initiate status switches between navigation and manipulation modi. The roles of the hands may also change, i.e., the dominant hand can be used for gestures, whereas the non-dominant interacts via the keyboard.

4.4 Stereoscopic Mouse Cursor

When using the described setup we experienced some drawbacks. One shortcoming, when interacting with stereoscopic representations using desktop-based interaction paradigms is the monoscopic appearance of the mouse cursor, which disturbs the stereoscopic perception. Therefore we provide two different strategies to display the mouse cursor. The first one exploits a stereoscopic mouse cursor, which hovers over 3D objects. Thus the mouse cursor is always visible on top of the objects surface, and when moving the cursor over the surface of a three-dimensional object, the user gets an additional shape cue about the object. The alternative is to display the cursor always at the image plane. In contrast to ordinary desktop environments the mouse cursor gets invisible when it is obscured by another object extending out of the screen.

Thus the stereoscopic impression is not disturbed by the mouse cursor, indeed the cursor is hidden during that time. Figure 7 (left) shows a stereoscopic scene in Google Earth where the mouse cursor is rendered stereoscopically on top of the building.

4.5 Monoscopic Interaction Lens

Many 2D as well as 3D applications provide interaction concepts which are best applicable in two dimensions using 2D interaction paradigms. 3D widgets [7] are one example, which reduce simultaneously manipulated DoFs. Since these interaction concepts are optimized for 2D interaction devices and monoscopic viewing we propose a monoscopic interaction lens through which two-dimensional interactions can be performed without loosing the entire stereoscopic effect. Therefore we attach a lens at the position of the mouse cursor. The content within such an arbitrary lens shape surrounding the mouse cursor is projected at the image plane. Thus the user can focus on the given tasks and tools to perform 2D or 3D interactions in the same way as done on an ordinary monoscopic display. This can be used to read text on a stereoscopic object, or to interact with 3D widgets. Figure 7 (right) shows the usage of a monoscopic interaction lens in a 3D modelling application. Potentially, this lens can be visualized to one user who can manipulate the three-dimensional content by using a 3D widget, another user can view the 3D objects, whereas the lens is not visible in her sweet spot.

5. Implementation

To provide a technical basis for the concepts described above, we explain some implementation details of our 3D user interface framework [17, 20]. To allow simultaneous viewing monoscopic content need to be modified in order to make it perceivable on AS displays, while a stereo pair need to be generated out of the 3D content. Since these are diverse image processing operations first 2D is separated from 3D content. To achieve this separation, our technique acts as an integrated layer between 3D application and OS. By using this layer we ensure application operating system that the operating system takes care about rendering 2D GUI elements in a native way (see Figure 4 (step 1)).

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Fig. 4. Illustration of the interscopic user interface framework showing 2D and 3D content simultaneously.

5.1 Processing of 2D Content

When viewing unadapted 2D content on AS displays two separated images are perceived by the eyes that do not match. This leads to an awkward viewing experience. To make this content perceivable we have to ensure that left and right eye perceive almost the same information, resulting in a flat two-dimensional image embedded in the image plane. To achieve this effect with (vertical-interlaced) AS displays the 2D content has to be scaled (see Figure 4 (step 2) in order to ensure that in the odd and even columns almost same information is displayed. With respect to the corresponding factor, scaling content can yield slightly different information for both half images. However, since differences in both images are marginal, the human vision system can merge the information to a final image, which can be viewed comfortably. Since we achieve proper results for a resolution of 1024×768 pixels we choose this setting for a virtual desktop from which the content is scaled to the AS displays native resolution, i.e., 1600×1200 pixels. Therefore, we had to develop an appropriate display driver that ensures that the OS announce an additional monitor with the necessary resolution and mirrors the desktop content into this screen.

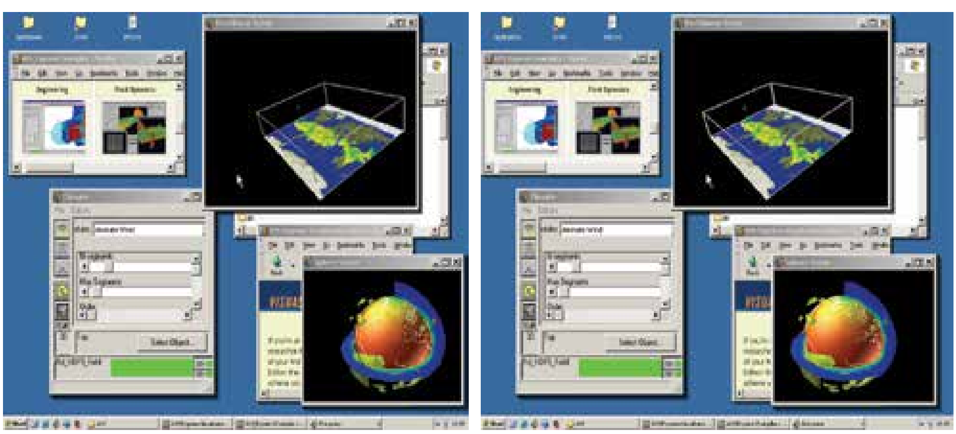


Fig. 5. Two stereoscopic half images arranged side-by-side, i.e., (left) for the left eye and (right) for the right eye.

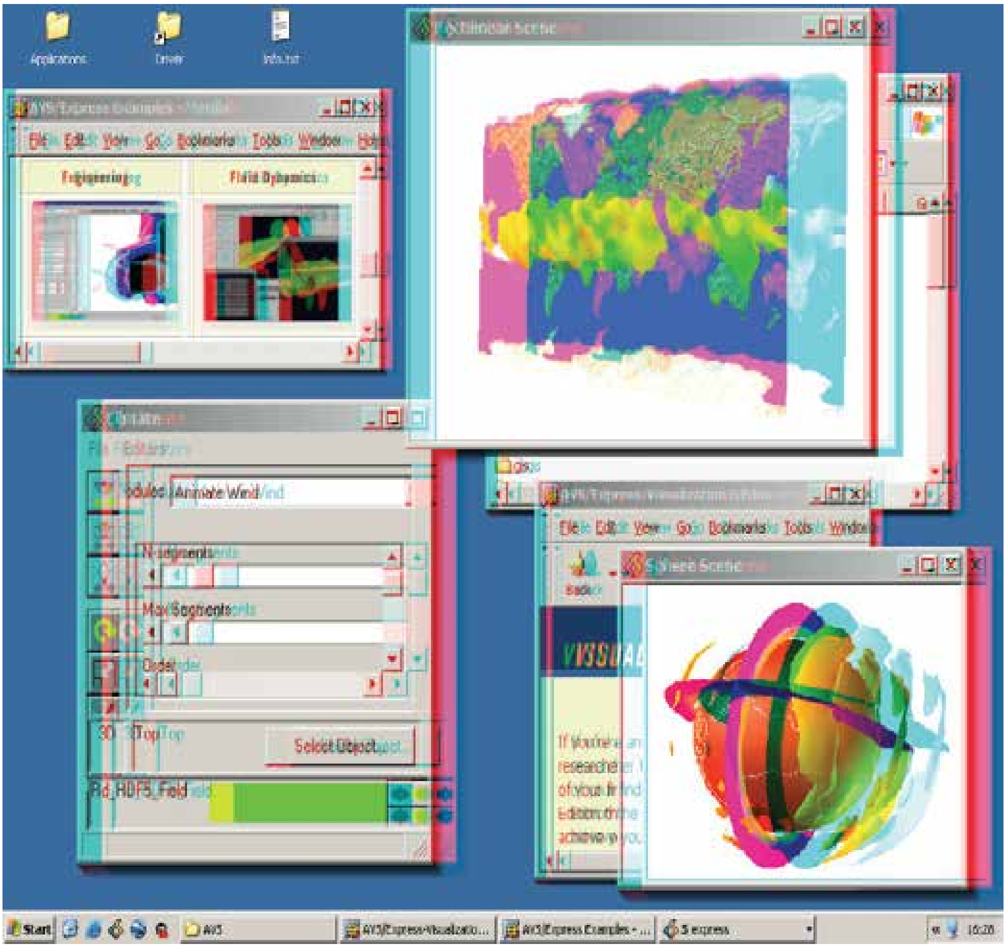


Fig. 4. Screenshot of the 3D user interface showing mono- and stereoscopic content simultaneously.

5.2 Generating Stereoscopic Images

Since only a few 3D applications natively support stereoscopic viewing on AS displays, in most cases we have to adapt also the 3D content in order to generate stereoscopic images (see Figure 4 (step 3)). There are two techniques for making an existing 3D application stereoscopic. The first one is to trace and cache all 3D function calls and execute them twice, once for each eye. The alternative exploits image-warping techniques. This technique performs are projection of the monoscopic image with respect to the values stored in the depth `buffer`. Image warping has the shortcoming that not all the scene content potentially visible from both eyes is presented in a single monoscopic image, and thus pixel filling approaches have to be applied [10]. Hence, we use the first approach, catch all 3D function calls in a display list, apply off-axis stereographic rendering, and render the content in the even and odd columns for the left respectively right eye. We generate a perspective with respect to the head position as described in Section 4. Figure 5 shows an example of a pair of stereoscopic half images. The images can be viewed with eyes focussed at infinity in order to get a stereoscopic impression. Figure 6 shows a screenshot of a desktop with mono- as well as stereoscopic content in anaglyph mode.

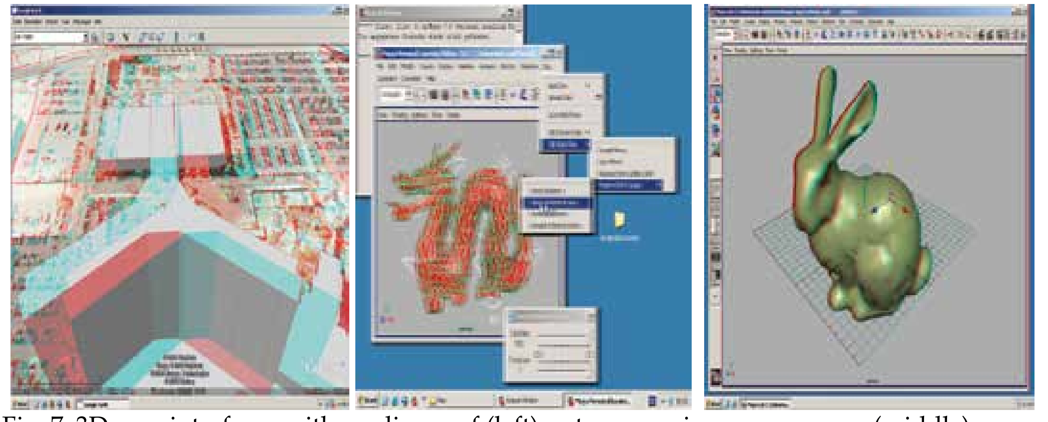


Fig. 7. 3D user interfaces with appliance of (left) a stereoscopic mouse cursor, (middle) several context menus and (right) monoscopic interaction lens.

Embedding Mono- and Stereoscopic Display To separate 2D and 3D content, we have to know which window areas are used for stereoscopic display. This can be either determined manually or automatically. When using the manual selection mechanism, the user is requested to add a 3D window or region and selects it to be displayed stereoscopically with the mouse cursor. When using automatic detection, our framework seeks for 3D windows based on OpenGL and applies stereoscopic rendering. The final embedding step of 2D and 3D content is depicted by step 3 in Figure 4. An obvious problem arises, when 2D and 3D content areas overlap each other. This may happen when either a pull-down menu or a context menu overlaps a 3D canvas. In this case the separation cannot be performed on the previous 3D window selection process only. To properly render overlaying elements we apply a masking technique. This is for example important, when dealing with 3D graphics applications, whereas context menus provide convenient access to important features. When merging 2D and 3D content the mask ensures that only those areas of the 3D window are used for stereoscopic display, which are not occluded by 2D objects. Figure 5 shows two resulting screenshots in anaglyph respectively interlaced stereoscopic mode, where 3D content is shown in stereo. The windows appear at different distances to the user (see Section 4.2). The task bar and the desktop with its icons are rendered monoscopically.

6. Experiments

In several informal user tests, all users have evaluated the usage of stereoscopic display for 3D applications as very helpful. In particular, two 3D modelling experts revealed stereoscopic visualization for 3D content in their 3D modelling environments, i.e., Maya and Cinema4D, as extremely beneficial. However, in order to evaluate the 3D user interface we have performed a preliminary usability study. We have used the described experimental environment (see Section 3). Furthermore, we have used a 3D mouse to enable precise 3D interaction.

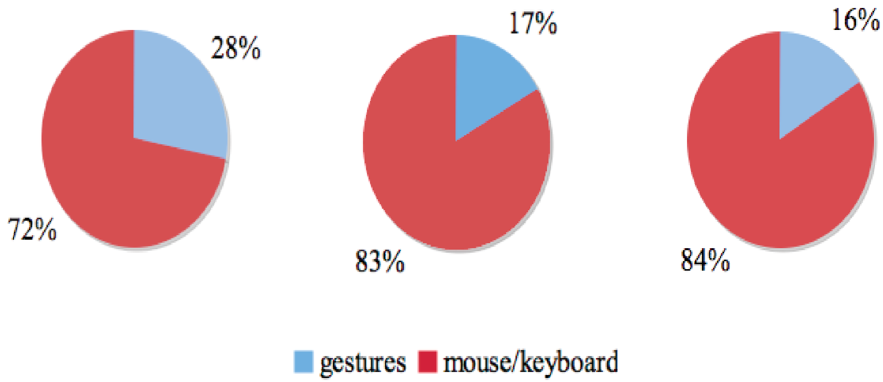


Fig. 8. Usage of gestures in comparison to traditional input devices constrained to (left) three DoFs, (middle) two DoFs and (right) one DoFs.

6.1 Experimental Tasks

We restricted the tasks to simple interactions in which four users had to delete several doors and windows from a virtual building. The building consisted of 290 triangles, where windows and doors (including 20 triangles) were uniformly separated. We have conducted three series. In the first series the user could use all provided input paradigms, i.e., mouse, keyboard, and gestures via a 3D mouse, in combination with stereoscopic visualization. In this series we have also performed sub-series, where gestures were constrained to three, two and one DoFs. In the second series, only the mouse and keyboard could be used, again with stereoscopic display. In the last series, interaction was restricted to traditional devices with monoscopic visualization.

6.2 Results

We have measured the required time for the entire task and we have measured how long each input modality has been used. Figure 6 shows that the less DoFs are available the less gestures have been used. When three DoFs were supported (left), one-third of the entire interaction time was spent on 3D manipulation by gestures with the objective to arrange the virtual building. With decreasing DoFs the required time for 3D manipulation also decreases. This is due to the fact that constraint-based interaction supports the user when arranging virtual objects. As pointed out in Figure 7 using gestures in combination with mouse and keyboard enhances performance, in particular when 3D manipulation is constrained approriatly. Participants accomplished the task fastest, when all devices could be used and only one DoFs was supported. Monoscopic display was advantageous in comparison to stereoscopic display. This is not unexpected since exploration of 3D objects was required only marginal; the focus was on simple manipulation where stereoscopic display was not essential.

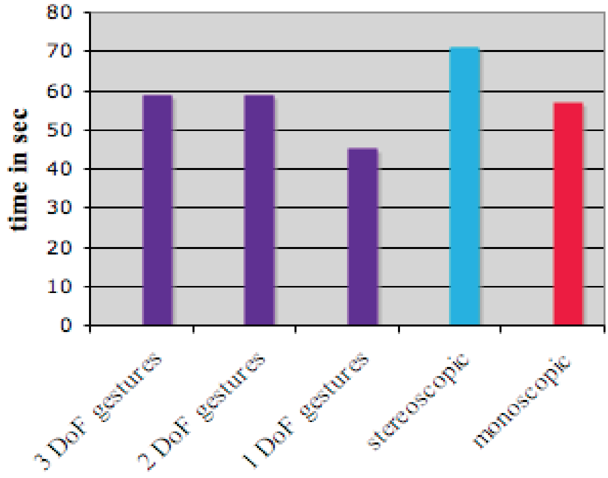


Fig. 9. Required time for the interaction task with stereoscopic display and gestures supporting three, two and one DoFs, and stereoscopic as well as monoscopic display only supporting mouse and keyboard without gesture.

7. Discussion and Future Works

In this chapter we have introduced 3D user interface concepts that embed in everyday working environments providing an improved working experience. These strategies have the potential to be accepted by users as new user interface paradigm for specific tasks as well as for standard desktop interactions. The results of the preliminary evaluation indicate that the subjects are highly motivated to use the described framework, since as they remarked instrumentation is not required. Moreover, users like the experience of using the 3D interface, especially the stereoscopic facetop approach. They evaluated the stereoscopic mouse cursor as clear improvement. The usage of the monoscopic interaction lens has been revealed as very useful because the subjects prefer to interact in a way that is familiar for them from working with an ordinary desktop system.

In the future we will integrate further functionality and visual enhancements using more stereoscopic and physics-based motion effects. Moreover, we plan to examine further interaction techniques, in particular, for domain-specific interaction tasks.

8. References

A. Agarawala and R. Balakrishnan. Keepin' It Real: Pushing the Desktop Metaphor with Physics, Piles and the Pen. *In Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1283–1292, 2006.

Z. Y. Alpaslan and A. A. Sawchuk. Three-Dimensional Interaction with Autostereo- scopic Displays. In A. J. Woods, J. O. Merritt, S. A. Benton, and M. R. Bolas, editors, *Proceedings of SPIE*, Stereoscopic Displays and Virtual Reality Systems, volume 5291, pages 227–236, 2004.

R. Balakrishnan. A Grant of 3D. *Keynote speach Symposium on 3D User Interfaces*, 2006.

H. Benko, E. W. Ishak, and S. Feiner. Cross-Dimensional Gestural Interaction Techniques for Hybrid Immersive Environments. *In Proceedings of the Virtual Reality*, pages 209–216. IEEE, 2005.

P. Bourke, Autostereoscopic Lenticular Images (http://local.wasp.uwa.edu.au/pbourke).

D. Bowman, E. Kruij, J. LaViola, and I. Poupyrev. *3D User Interfaces*: Theory and Practice. Addison-Wesley, 2004.

D. B. Conner, S. C. Snibbe, K. P. Herndon, D. C. Robbins, R. C. Zeleznik, and A. van Dam. *Three-Dimensional Widgets*. In Symposium on Interactive 3D Graphics, 1992.

N. A. Dodgson. Autostereoscopic 3D Displays. *In Computer*, volume 38, number 8, pages 31–36, 2005.

A. J. Hanson and E. Wernert. Constrained 3D Navigation with 2D Controllers. *In Proceedings of Visualization '97*, pages 175–182. IEEE Computer Society Press, 1997.

P. Kozankiewicz. Fast Algorithm for Creating Image-based Stereo Images. *In Proceedings of WSCG*, pages 59–66, 2002.

J. Liu, S. Pastoor, K. Seifert, and J. Hurtienne. *Three Dimensional PC toward novel Forms of Human-Computer Interaction*. In Three-Dimensional Video and Display Devices and Systems SPIE, 2000.

J. D. Mulder and R. van Liere. Enhancing Fish Tank VR. *In Proceedings of Virtual Reality*, pages 91–98. IEEE, 2000.

T. Peterka, R. L. Kooima, J. I. Girado, J. Ge, D. J. Sandin, A. Johnson, J. Leigh,

J. Schulze, T. A. DeFanti. Dynallax: Solid State Dynamic Parallax Barrier Autostereoscopic VR Display. *In Proceedings of the IEEE Virtual Reality Conference 2007*, pages 91–98. IEEE, 2007.

Philips 42-3D6C01. (http://www.inition.co.uk).

Project Looking Glass. (http://www.sun.com/software).

G. Robertson, M. van Dantzich, D. Robbins, M. Czerwinski, K. Hinckley, K. Risden, D. Thiel, and V. Gorokhovsky. The Task Gallery: A 3D Window Manager. *In Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 494–501, 2000.

T. Ropinski, F. Steinicke, G. Bruder, and K. Hinrichs. Simultaneously Viewing Monoscopic and Stereoscopic Content on Vertical-Interlaced Autostereoscopic Displays. *In Poster-Proceedings of the 33rd International Conference on Computer Graphics and Interactive Techniques* (SIGGRAPH06), 2006.

T. Salzman, S. Stachniak, and W. Stürzlinger. Unconstrained vs. Constrained 3D Scene Manipulation. *In 8th IFIP International Conference on Engineering for Human-Computer Interaction*, volume 2254, 2001.

G. Smith, T. Salzman, and W. Stürzlinger. 3D Scene Manipulation with Constraints. In B. Fisher, K. Dawson-Howe, and C. O'Sullivan, editors, *Virtual and Augmented Architecture*, pages 35–46, 2001.

F. Steinicke, T.Ropinski, G.Bruder, K. Hinrichs. Towards Applicable 3D User Interfaces for Everyday Working Environments International Conference on Human-Computer Interaction (INTERACT2007), *Lecture Notes in Computer Science*, Volume 4662, page 546-559, 2007.

D. Stotts, J. C. Smith, and K. Gyllstrom FaceSpace: Endo- and Exo-Spatial Hyper- media *in the Transparent Video Facetop Proceedings of the 15th ACM Conference on Hypertext and Hypermedia*, pages 48 - 57, 2004.

Z. Szalavari and M. Gervautz. Using the Personal Interaction Panel for 3D Interaction. *In Proceedings of the Conference on Latest Results in Information Technology*, page 36, 1997.

H. Tramberend. A Display Device Abstraction for Virtual Reality Applications. *In Proceedings of Afrigraph*, pages 75–80, 2001.

C. van Berkel. Touchless Display Interaction. *In SID 02 DIGEST*, 2002.

18

Context-aware Mobile AR system for Personalization, Selective Sharing, and Interaction of Contents in Ubiquitous Computing Environments

Youngjung Suh, Youngmin Park, Hyoseok Yoon, and Woontack Woo
GIST U-VR Lab
500-712, S.Korea

1. Introduction

Now, the augmented reality (AR) systems have been developed to faithfully realize the basic concept that their main purpose is to supplement the real world with the addition of virtual objects (or information) to enhance a user's perception of and interaction with the real world (Azuma,1997; Azuma & Baillot, 2001). To this end, researchers in AR technologies have emphasized the technical challenges involved in providing accurate augmentation, natural interaction, and realistic rendering. With current advances in tracking and increased computing power, there have been corresponding developments in mobile AR systems. There have also been a number of studies pertaining to mobile AR technologies attempting to overcome these technical challenges (Pasman & Woodward, 2003; Wither et al., 2006; Billinghurst et al., 2000; Farbiz et al., 2005). However, we should consider more than just such immediate technical questions, but rather work to resolve issues related to possible interfaces and contents for user interaction in ubiquitous computing environments.

Several studies have adopted mobile AR technology to develop systems that offer potential interfaces and contents to users carrying a mobile device (Matsuoka et al., 2002; Geiger et al., 2001). Despite this provision, just providing every user with a uniform interface using mobile AR technology does not actually give much consideration to an individual user's desires, needs, and preferences. Also, with too much content, the use of mobile AR technology may cause users to become easily confused. In ubiquitous computing environments, various aspects related to the context of users and their environment can be utilized to provide a user with a personalized user interface and filtered contents.

There have been a number of efforts to support AR-enabling techniques using the notion of context-awareness (Henrysson & Ollila, 2004; Long et al., 1996; Feiner et al., 1997). However, these efforts have been mainly restricted to exploiting spatial contexts in adopting the notion of context-awareness in mobile AR systems. Nevertheless, there is a broad range of contexts available, ranging from physical to user contexts. This development has in turn led to a number of other efforts attempting to bridge between ubiquitous computing environments and mobile AR technology.

For instance, aPost-it was proposed as a context-based information augmentation and sharing system (Oh et al., 2004). However, this system only personalizes media contents through a webpage obtained from an object, and also lacks an efficient interface for controlling smart appliances in an intuitive and personalized manner. However, intuitive and personalized is the type of interface required to enable user's easy access to and customized control of pervasive and invisible smart appliances in a ubiquitous computing environment. In addition, for selective sharing of media contents in aPost-it, users have to explicitly inform the system whether they would like to share the contents through the user interface on the web page. Explicitly requesting the selective sharing of contents, however, is quite inconvenient for users.

Thus, in this paper, we propose a Context-Aware Mobile Augmented Reality (CAMAR) system. Besides enabling new ways of taking pictures of smart appliances in our daily life to control them, it brings innovative ways of associating context with multimedia-based interaction by enabling the interface and media contents to not only be personalized, but shared selectively and interactively among a group of people.The system is based on the ability to access contexts in a user's mobile device in a ubiquitous computing environment through a mobile AR technology-enabled interface.

On the one hand, it is generally agreed that "point-and-click" is the action of a computer user moving a cursor to a certain location on a screen (point) and then clicking a mouse button. In our system, we propose a point-and-click interface (Beigl, 1999) as the user interface, where users only need to take a picture of a smart appliance with a built-in camera in a mobile device to indicate their intention to control the appliance. The action of "taking a picture" is very similar to the one of "pointing" to indicate a user's interest. Similarly, the "click" of a mouse button to execute commands corresponds to the "control" of smart appliances. In this way, the system allows users to interact with smart appliances through a personalized control interface displayed on their mobile devices.

On the other hand, CAMAR supports enabling media contents to be not only personalized, but also shared selectively and interactively among a group of people based on mobile user's profile and context in ubiquitous computing environments. Even if separate users look into the same AR marker with a built-in camera in a mobile device, different media contents are augmented on their mobile device. Here, media contents are personalized as they are processed with context information. Then, the personalized media contents can be selectively shared within a community that our system implicitly constructs by analyzing context information.

Thus, by bridging a variety of contexts in ubiquitous computing environments and mobile AR technologies, the proposed system overcomes limitations in existing Information Technologies, which tend to provide the same information to all end-users. Applicable areas of the proposed system include mobile AR applications, such as a meeting system that supports information augmentation to a real environment for collaborations, a universal remote control for controlling various kinds of smart objects, and a mobile service agent that utilizes a user's location and activity to diversify and expand its use for mobile AR-based services.

This paper is organized as follows. In section 2, an overview of the CAMAR system including potential applicable scenarios is provided. We deal with implementations in Section 3, and Section 4 describes usability tests and observations. Finally, we conclude our work along with a brief outline of remaining work in Section 5.

2. System Overview

In this paper, we present a Context-Aware Mobile Augmented Reality (CAMAR) system that supports two main functionalities. One is the intuitive and personalized control of smart appliances with mobile AR devices. The other is enabling media contents to be not only personalized, but also shared selectively and interactively among a group of people. The important foci of our system are as follows:
1) Controller augmentation (translate appearance to controller) –when a user takes a picture of a marker attached to a smart appliance through a mobile AR device, the system offers the user a personalized visual interface that can be used to control the smart appliance.
2) Multimedia augmentation (translate marker to contents) – when a user reads a marker embedded with contents (a digital map) through a mobile AR device, the system offers personalized media contents (photos) to the user.

2.1 System Features

Before specifying a concrete scenario, it is perhaps more useful to categorize the functionalities that such a scenario implies. A Context-Aware Mobile AR system aims at enabling the personalization and selective sharing of media contents as well as the mobile AR-based control of smart appliances in a ubiquitous computing environment. These aims are achieved as follows.
1) Support for easy and intuitive accessibility – the ubiquitous computing environment includes a large number of pervasive and invisible computing resources. Due to the multitude and invisibility of resources, it is difficult for users to make a use of those computing resources. Moreover, as the computing resources such as information appliances become smarter with more features, the accompanying user interfaces get complicated and becomes burdens for users (Badami & Chbat, 1998). In our approach, CAMAR supports the user-centered access of computing resources, especially smart appliances with the interface that is intuitive and easy to use. The user only needs to take a picture of computing resources with a built-in camera of the mobile device to indicate his intentions to access as in a "point-and-click interface."
2) Support for a personalized control interface – if computing resources become ubiquitous, then devices will be used in a wide range of dynamically changing environments. For these devices to be truly helpful to a user's interaction with smart services in this type of environment, they must be aware of both the environment as well as the user who is interacting with the services in the environment. To this end, CAMAR supports the personalization of a smart appliance controller that adapts based on what the user's interface usage patterns are. Our system supports a personalized control interface by generating a customized control menu that best suits the user's needs in controlling smart appliances. It allows the system to provide an intuitive and transparent user interface such that the user can concentrate on the original task.
3) Support for context-based media content provision and selective sharing – a user will be supplied with a potentially overwhelming amount of media contents in a ubiquitous computing environment. To assist the user in avoiding confusion, the development of a personalized content provider is needed, one that adapts based on who is present and what their preferences are. For this reason, CAMAR allows different media contents to be augmented to individuals even if they look at the same marker. Another key theme of

ubiquitous computing systems is the support for groups of people (Kohno & Rekimoto, 2005). For selective sharing of media contents in the ubiquitous computing systems, users generally do not like to explicitly inform the system of whether or not they would like to share the contents through conventional user interfaces. Thus, this system allows personalized media contents to be selectively shared within a community that our system implicitly constructs by analyzing context information. It specifically enables selective sharing of common knowledge and experiences of contents interactively and collaboratively among a group of people. Fig. 1 shows the concept diagram for a CAMAR system supporting mobile AR-based control, personalization, and community-based selective sharing.

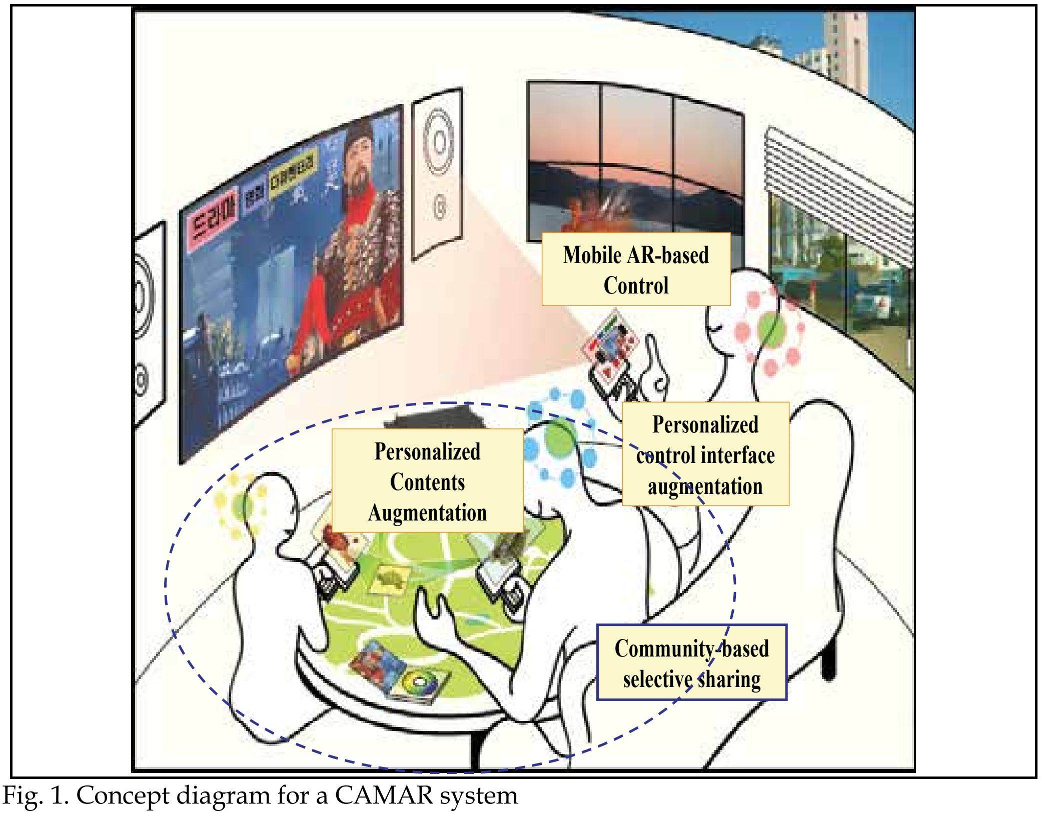


Fig. 1. Concept diagram for a CAMAR system

2.2 Applicable Scenarios

As a test-bed for a ubiquitous computing environment, we used a smart home which has been established in our laboratory. In addition, as a test-bed for content, we used information on a specific domain, the Unju Temple (Lee et al., 2005), to allow our test data to include photos and detailed 3D models of places (Lee et al., 2005). Fig. 2 shows a conceptual diagram of the system, suggesting potential scenarios applicable to the CAMAR system. A typical scenario is outlined as follows.

Fig. 2. Visualization of overall scenarios applicable to the CAMAR system

1) Easy and intuitive accessibility: Embedded marker-based and personalized smart appliance control using a mobile device.

When a user takes a picture of a specific smart appliance with a built-in camera in a mobile device, a controller for the captured smart appliance can be augmented on the mobile device: a smart TV controller, a smart table controller, a smart window controller, and/or a smart light controller.

Scene #1	*Hyoseok comes home from work. He becomes aware of the fact that his wife Hyejin went to a market in the neighborhood. As soon as he comes into the living room, the lighting service prepares the green lighting he usually prefers at this time of day. He changes the color of lighting to a blue one since he feels the weather is so hot, and wants to make the room feel a little more refreshing.*
Description	A user takes a picture of a light switch covered with an embedded marker in consonance with an on-off switch. After that he can control the functions of the lighting service. Also, he can confirm other available service lists from the service discovery on his mobile device.
Scene #2	*Then, Hyoseok approaches a smart TV. The smart TV recognizes that he might approach close to it. Hyoseok selects a sports channel using the smart TV controller on his mobile device and begins to watch TV.*

Description	A user can take a picture of the visible marker displayed on the TV as a kind of screen saver.
Scene #3	*Hyejin then returns home from the market. After she confirms that Hyoseok has selected a sports channel to watch on TV, she moves to a smart window. She begins to control a service of the smart window that allows her to navigate a virtual view of Unju Temple. After she entertains herself by viewing the virtual navigation system, she comes over to the smart TV. When she is in front of the smart TV, a control hand-over button appears on Hyoseok's smart TV controller on his mobile device. As soon as Hyoseok transfers his control of the smart TV to Hyejin, a visual marker indicating control transfer completion is displayed in the upper-right corner of the smart TV. Hyejin then takes a picture of the visual marker. Simultaneously, the smart TV recognizes that Hyejin has obtained control from Hyoseok and provides a recommendation menu based on Hyejin's preferences. Hyejin selects a cooking channel using the smart TV controller on her mobile device and begins to watch TV.*
Description	In this situation, the second user is supplied with a personalized control menu interface that suits her best. She can control the smart appliance, in a personalized manner, by using the personalized menu interface on her mobile device.

Table 1. Description of the scenario applicable to the personalized smart appliance controller

Table 1 and Fig. 3 show a description and visualization of the scenario applicable to the Personalized Smart Appliance Controller, respectively.

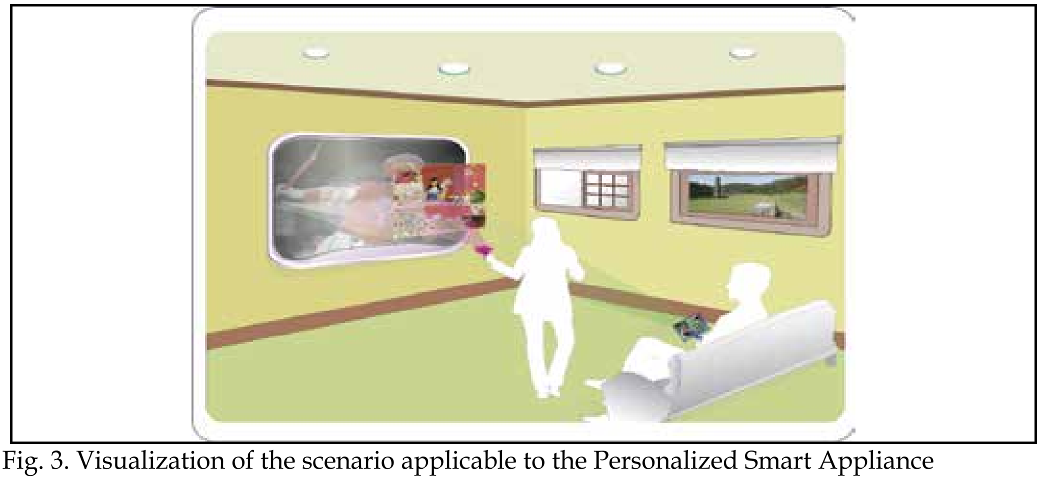


Fig. 3. Visualization of the scenario applicable to the Personalized Smart Appliance Controller

2) Context-based media content augmentation and sharing

Scene #1 (*Hyoseok returned from visiting Unju Temple last week)*	*Hyoseok visited the Unju Temple with his family last week. He really loves taking photographs, so he took a number of pictures of the precincts of the Unju Temple. The photos were spatially indexed through a GPS-receiver in the mobile device that was used to take the photographs, allowing Hyoseok to later know where a certain photo was taken. Hyoseok decides to recommend that his friend Youngmin visit the Unju Temple, so he invites his friend over to his home to show him the pictures of his visit last week. Because it has already been a week since he visited, it is difficult for Hyoseok to remember exactly when and where he took which shots. For this reason, Hyoseok likes to reminisce about the Unju Temple through a smart table and navigate the paths of the virtual Unju Temple. He then recognizes at a glance the places he took the pictures through the map displayed on the smart table. Hyoseok moves his mobile device, just like looking through a magnifying glass, to a specific region on the map to see the pictures he took in the region. Thus, Hyoseok can enjoy the pictures augmented on the screen of his mobile device.*
Description	In this situation, only the media content with a high degree of similarity to his specified preferences are recommended and augmented.
Scene #2 (*Hyoseok with his son Hyunmin*)	*Hyunmin visited the Unju Temple with his dad last week, and so cares to join his dad visiting a virtual Unju Temple, after coming back from school. The places at which Hyoseok took the pictures are indicated through makers displayed on the map on a smart table. "Wabul" in the map is indicated by a distinguished marker. Hyoseok decides to look into that place moving his mobile device with his hands to that location. Hyunmin also gets interested in the "Wabul"and looks at it on the map since he remembers he had taken some pictures with his dad at the rock. They begin to talk about how they felt at the time while in that place. Then, Hyoseok wishes to share pictures, of which he is thumbing through, with his son in which they appear together. On Hyoseok's pressing a button in his mobile device for sharing, Hyunmin receives and enjoys the pictures as well.*
Description	In this situation, Hyoseok's photos of "Wabul" tower are delivered to *Hyunmin's* mobile device. In this way, Hyunmin can share Hyoseok's experiences via the photos augmented on his mobile device.
Scene #3 (*Youngmin visits Hyoseok's home*)	*Youngmin has never visited the Unju Temple, and so decides to visit a virtual Unju Temple using Hyoseok's smart table, after hearing from Hyoseok that the Unju Temple is quite unlike other temples. The places at which Hyoseok took his pictures are indicated through markers displayed on the map on the smart table; for instance, "Wabul" tower in the map is indicated by a distinguished marker. Youngmin decides to look at that place by moving his mobile device closer to that marker. Hyoseok wishes to share his experiences of "Wabul" tower with Youngmin by sending him pictures of his previous visit, so he presses the sharing button on his mobile device. Youngmin looks at the pictures and feels that "Wabul" rock in the Unju Temple is indeed impressive. Thus, he becomes interested in the temple with the help of content augmentation on his mobile device.*
Description	In this situation, Hyoseok's photos of "Wabul" tower are delivered to Youngmin's mobile device. In this way, Youngmin can share Hyoseok's experiences via the photos augmented on his mobile device.

Table 2. Description of the scenario applicable to a context-based content augmentation and sharing service

Table 2 and Fig. 4 show a description and visualization of the scenario applicable to a context-based content augmentation and sharing service, respectively.

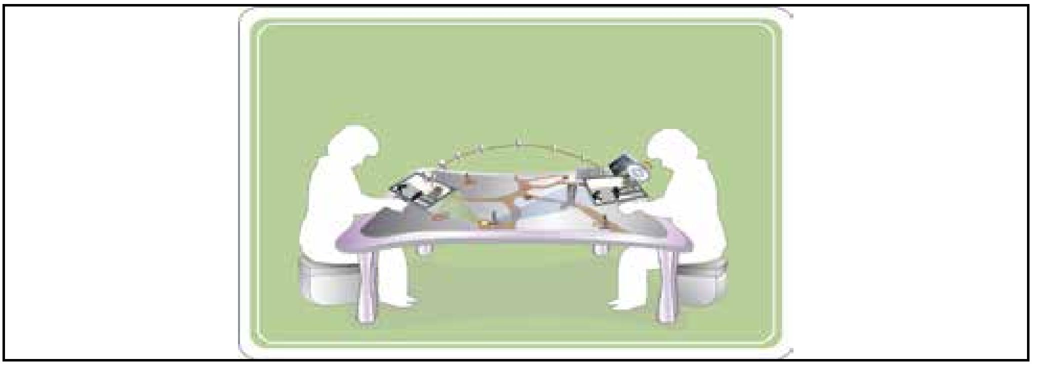


Fig. 4. Visualization of the scenario applicable to a context-based content augmentation and sharing service

3. System Details

Here, we identify the main components for a system targeting the main challenges of smart appliance control, personalized content, and selective sharing of personalized contents. To this end, Fig. 5 shows the overall system block diagram of CAMAR.

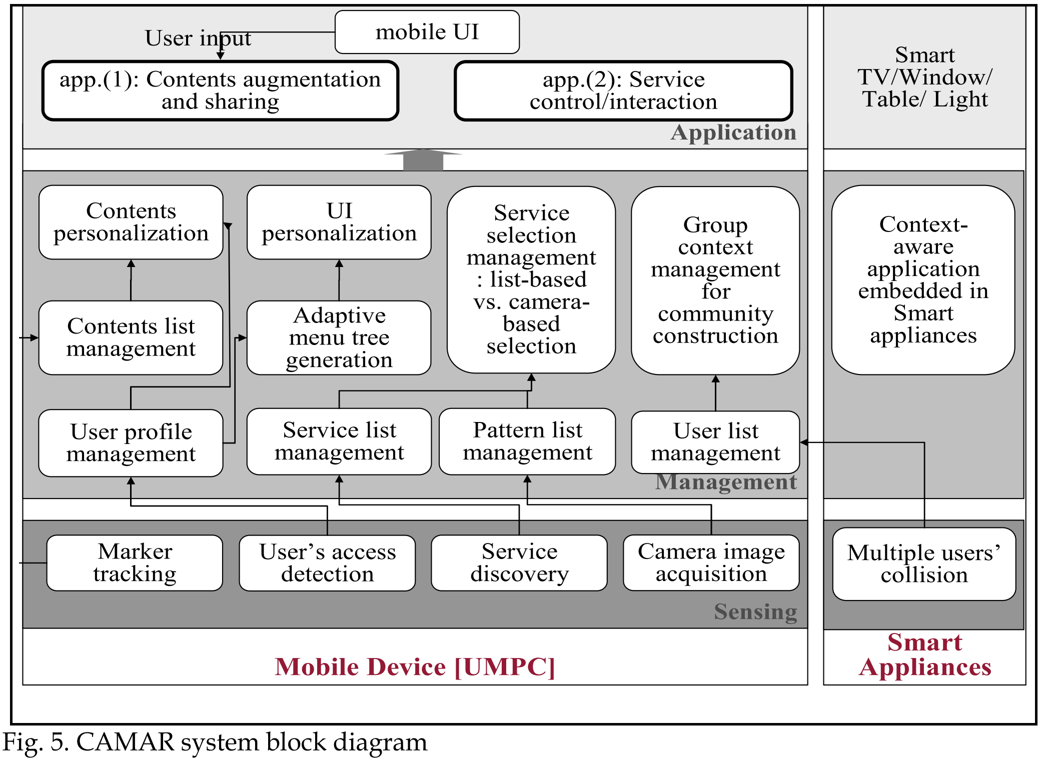


Fig. 5. CAMAR system block diagram

3.1 Easy and intuitive accessibility for ubiquitous computing resources

To support smart appliance control, we have designed a camera-based method along with a set of logical modules for service discovery, service selection, and service interaction. In the sensing layer, a user's personal mobile device can discover and visualize potential services in the environment. Specifically, the built-in camera is used to recognize and identify smart appliances that can later be personalized. In order to recognize smart appliances, we either embed markers into the appliances or use the physical features of a smart appliance as the marker. In this way, smart appliances with a display such as ubiTV, MRWindow, and ARTable (Oh et al., 2005; Park & Woo, 2006) display a screen saver while in ready-mode, which later changes into a visible marker when the user is in the effective service area. In terms of the use of physical features, features of a light switch can be used as an embedded marker for the light service. Then, in the management layer, after the user selects a service through either a list-based or camera-based method, the user's context is exploited to personalize the user interface. Subsequently, the same logical flow can be further developed to control a number of smart appliances, which can then be included in one universal remote controller. Fig. 6 shows the procedural diagram for smart appliance control.

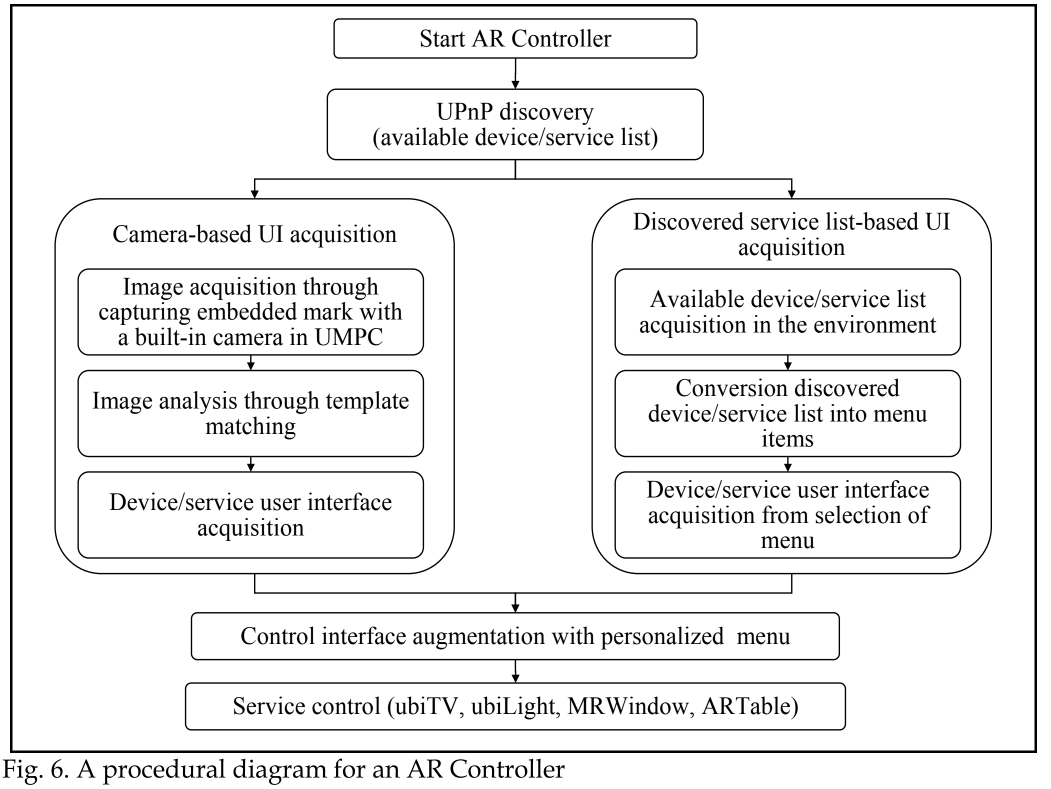


Fig. 6. A procedural diagram for an AR Controller

The smart appliances and services in this environment are implemented with ubi-UCAM (Unified Context-aware Application Model for Ubiquitous Computing Environment) (Oh et al., 2005) to enable context-awareness. Context-aware services are deployed into smart appliances to collect, integrate, interpret, and/or manage the user context to provide personalized services. For effective use of computing resources, context-aware services recognize each user's preferences and service status to detect conflicts. Then, when/if conflict occurs, the service profile of the conflicting service and the user profiles of the users in the conflict situation can be utilized to form a recommendation for a unified context (Jang & Woo, 2005). Thus, a conflict-free context can be delivered to other service providers and into the environment.

3.2 Personalization of control interface

When providing users with a smart appliance AR controller, menu tree organization in the control interface was our main consideration in terms of the personalization of the user interface. Mobile devices are becoming ever smaller, allowing users to make the most out of their portability and convenient manipulation. Thus, it should be possible to display a large amount of data in such a way that a users' satisfaction of data selection can be enhanced. In our work, personalization of the control interface aims to relieve the user from scrolling through screens and exploring a multi-layered menu structure by providing a simpler menu interface that best suits the user. This simplification is achieved by analyzing the user's usage pattern of the menu interface in his mobile device; first, the information pertaining to the menu items used as part of the user's interaction with his mobile device is collected and managed as a history.

We can then simplify the menu structure in the control interface using the history data collected over a certain period of time. Note that the menu items we simplify may involve the full range of items on the menu. The basic method we consider is to learn the frequency of selection of the menu item. However, this method may give rise to wrong results because the frequency in selecting the upper and middle levels of the menu gradually increases regardless of the user's intention. To overcome this inaccuracy, we used a reciprocal scoring method that includes the position information of the menu in the interface. In equation (1), indicates the score of menu x, the depth of the menu, and the constant.

$$q_x^{d-1} = \sum_x \frac{K}{q_x^d} \; (if \;\; q_x^d = 0 \;\; then \;\; set \;\; \frac{1}{q_x^d} = 0) \tag{1}$$

The order of the menu item is determined by the score calculated by (1) and then displayed in the user's mobile device.

3.3 Context-based contents augmentation and sharing

To consider the personalization of media contents, we developed a photo content recommendation module based on the user profile and the symbolic location information of the virtual heritage map displayed in a smart table. We exploit the metadata of the photos that were already spatially-indexed by the GPS-receiver in the mobile device that was used to take the photograph. First of all, we filter the photo contents shot at the sites that correspond with the symbolic location information of the virtual map in the smart table. Then, we use the user preferences with respect to the photo contents to draw up a list of photo contents and recommend them to the user. As a method for recommending photo contents based on user preferences, we use a similarity measuring equation (2) (Yu et al., 2006). Here, the user preference (P) and the metadata of the photo contents (C) are described in terms of the vectors, $P = (w_1, \ldots, w_n)$ and $C = (u_1, \ldots, u_n)$, respectively. In this representation, w_i is the weight value of the users on a certain property of the photo contents, and u_i is the weight value of the photo contents on the property corresponding to w_i.

$$Similarity(C,P) = \frac{C \cdot P}{\| C \| \times \| P \|} = \frac{\sum_{i=1}^{n} u_i w_i}{\sqrt{\sum_{i=1}^{n} u_i^2 \sum_{i=1}^{n} w_i^2}} \quad (2)$$

The fundamental context information we can acquire in this aspect includes the time when we took the photograph, and the location of where we took the photograph.

Context Element	Meaning	Description
Who	Who the user is	User ID
When	When the interaction takes place	The time the picture is taken at. The time the user interacts with contents in ARTable.
Where	Where the user is	Spatial context when taking the picture at the real Unju Temple.
		Spatial context when exploring the virtual Unju Temple in the virtual world displayed on ARTable.
What	What the interaction is about	Service ID Content ID
Why	What the user is interested in	Preferences: mathematics, photography, history, etc.

Table 3. Context information used in the system

Types of 5W1H context information (Jang & Woo, 2005) used in our system are described in Table 3. Additionally, the system generates and manages group context by extracting common preferences through an analysis of multiple users' integrated contexts and their relationships. After managing the group context, it selectively allows users with common interests to share contents. Fig. 7 shows the overall process of the phase in which one user has priority over another in enjoying context-based content augmentation and sharing in ARTable.

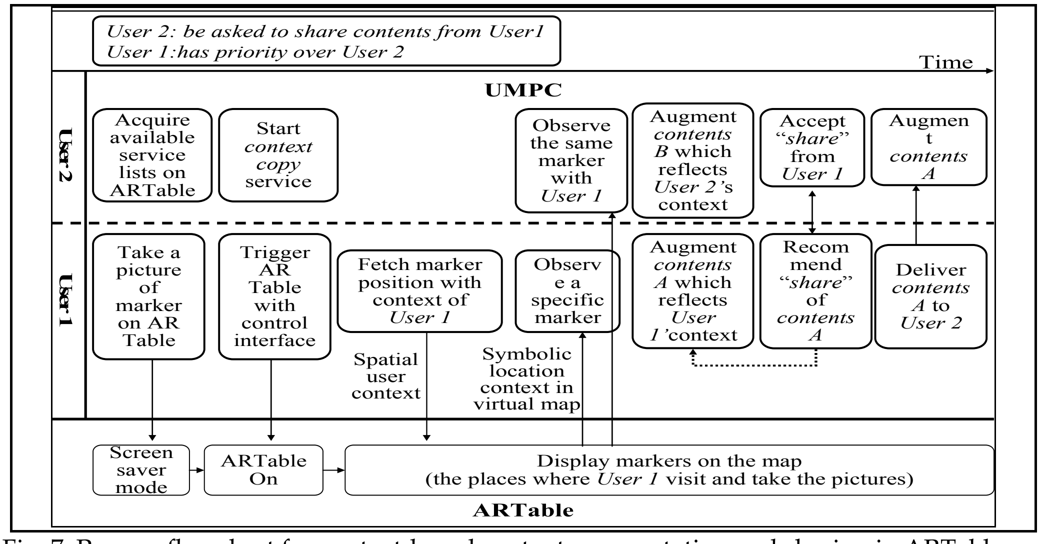


Fig. 7. Process flowchart for context-based content augmentation and sharing in ARTable

4. Implementation

We implemented a Context-Aware Mobile Augmented Reality (CAMAR) system that supports two main functionalities. One is the control of smart appliances with mobile devices. The other is enabling media contents to not only be personalized, but also selectively and interactively shared among a group of people. Table 4 describes the CAMAR system platform.

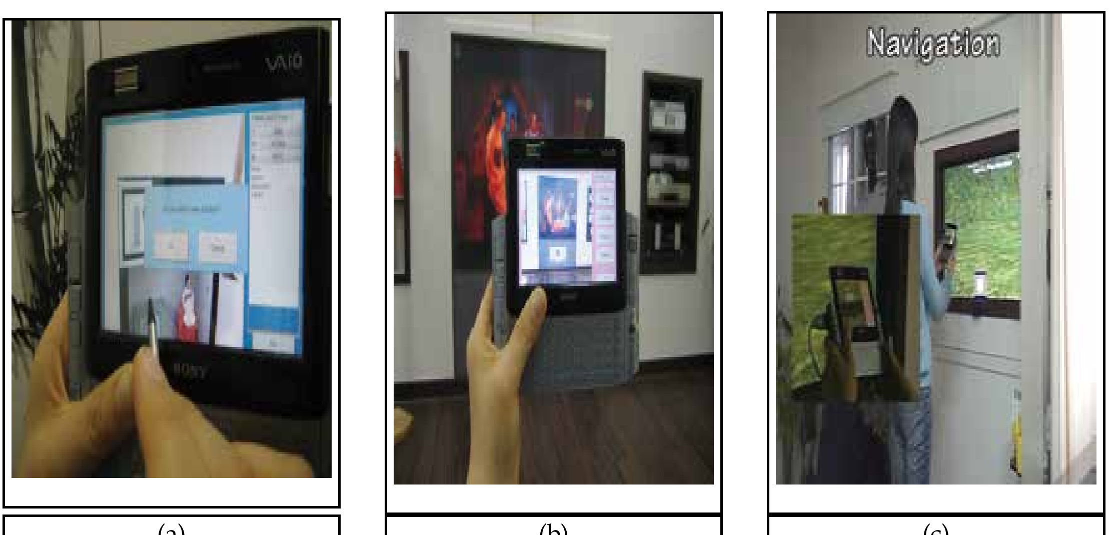
HW/SW	Specification
UMPC	SONY VAIO VGN-UX-17LP[1]
OS	Microsoft Windows XP Professional
Software Development IDE	Microsoft Windows Visual Studio 2005
Camera Library	OpenCV beta 5[2]
Image Processing Library	ARToolkit 0 Glut 3.7.6[3]
UPnP SDK	Intel® Authoring Tools for UPnP Technologies (Build1825) Intel® Tools for UPnP Technologies (Build 1768)

Table 4. System Platform

[1] http://vaio-online.sony.co.kr/
[2] http://sourceforge.net/projects/opencvlibrary
[3] http://www.xmission.com/~nate/glut.html

1) Smart appliance control: Controlling smart appliances with mobile devices

A Personalized Smart Appliance AR Controller is a mobile user interface that enables users to control smart appliances in a personalized manner using a mobile device with a built-in camera. Most mobile devices have only one dedicated user, making it easy for a mobile device to provide a personalized interface. For example, a mobile device provides an interface that is consistent with the interface for smart appliance that the user is familiar with or prefers. In this way, we designed and implemented the Personalized Smart Appliance AR Controller to allow users to control smart appliances in the environment through a personalized user interface. When a user wishes to control a smart appliance, he or she only needs to take a picture of it with the built-in camera. Then, the personalized and tailored service interface is automatically augmented on the mobile device. The controller device processes the pattern matching, obtains contextual information that contains an abstract functional description from the smart appliance, and uses the contextual information of the description to properly generate a personalized user interface.
The Personalized Smart Appliance AR Controller provides four main functions. The first function is the personalization of the mobile user interface. The second function pertains to service notification in terms of discovering devices and services in a user's home network environment. The third function is that a single mobile device such as PDA functions as a universal remote control for multiple devices and services. Lastly, when/if a service conflict occurs, i.e., you are prohibited from using a certain service because it is pre-occupied by another user, service recommendations and service control hand-over functions can be used to resolve any conflicts. Fig. 8 and Fig. 9 show smart appliance control functions in a UMPC version and a PDA version, respectively. The personalized controllers for smart appliances augmented on the mobile devices include ubiTV Controller, MRWindow Controller, ubiLight Controller, and ARTable Controller.



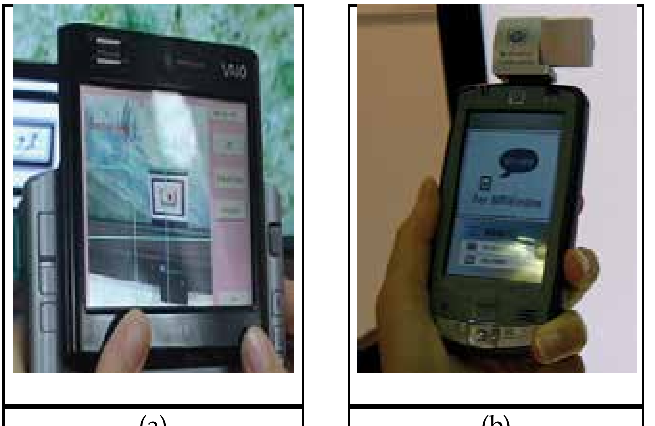
(a) (b) (c)

Fig. 8. Smart appliance control with AR Controller embedded in UMPC, (a) taking a picture of the ubiLight service, (b) controlling the ubiTV service, and (c) navigating in VR contents

(a) (b) (c)

Fig. 9. Smart appliance control with AR Controller embedded in a PDA, (a) controlling the ubiTV service, (b) taking a picture of the ubiLight service, and (c) taking a picture of the ubiTV service.

As prototypes, the AR Controller is implemented in both PDA and UMPC platforms. Here, the PDA platform has an advantage because it is relatively smaller, cheaper, and more portable. However, the UMPC platform performs better in image processing for pattern matching. Thus, for research purposes, we used both platforms interchangeably to develop compatible components for both the PDA and UMPC platforms. Fig. 10 shows two different versions of the AR Controller.



(a) (b)

Fig. 10. AR Controller in (a) UMPC platform and (b) PDA platform

2) Context-based content augmentation and sharing

To present content augmentation and sharing with mobile AR technologies, we implemented an edutainment system that augments photos taken in a site and allows users to share them. As shown in Fig. 11, we used ARTable to display a navigation map and AR markers, and UMPCs to augment and share the photo contents. We selected Unju Temple as the site and temple photos as the contents to augment and share.

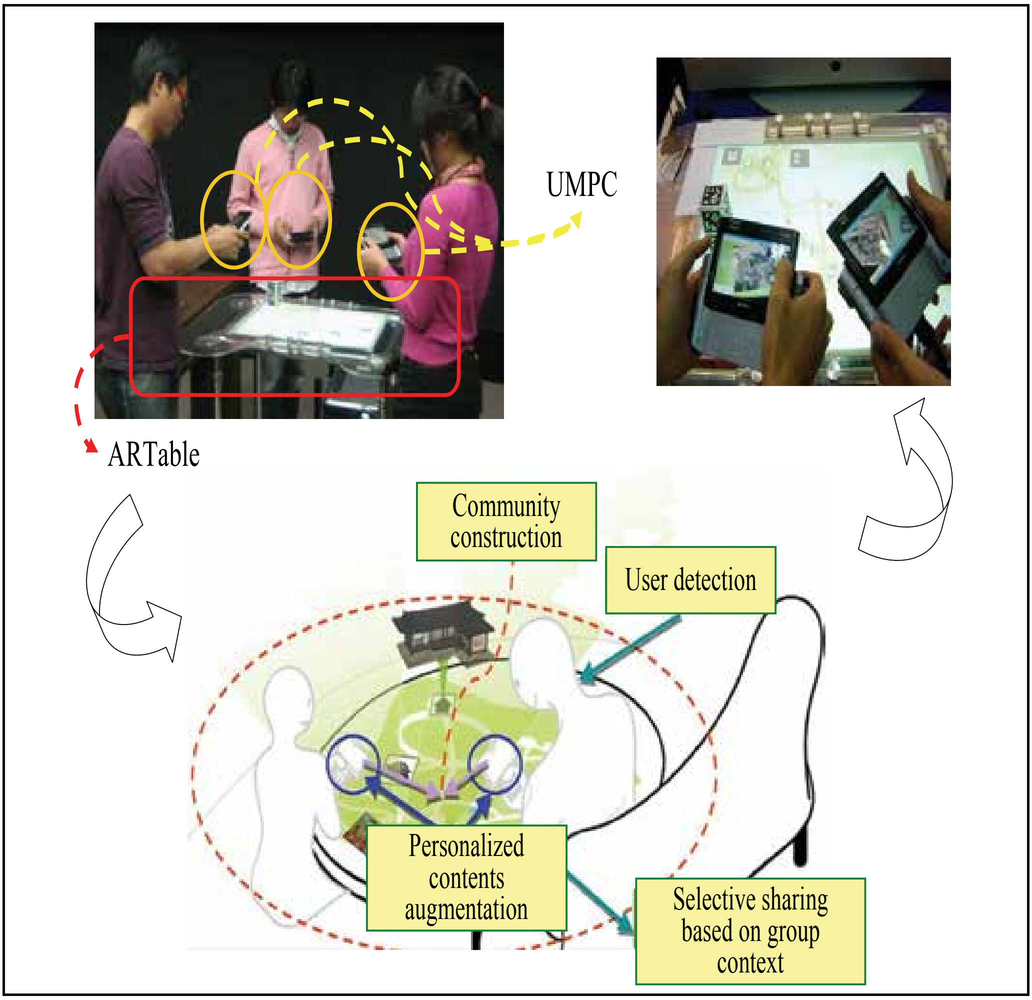


Fig. 11. System overview of context-based content augmentation and sharing system

When we explore a cultural site, we tend to take pictures of cultural assets and save them in our mobile device to record our memory or experience. Then, from the photos or video data the visitor can revisit the moment of visiting a cultural heritage site. Our system aims to realize Context Copy by extracting contexts, providing personalized media contents, and having them shared through mobile AR techniques. To this end, Fig. 12 shows a context-based photo content augmentation and sharing system for realizing the concept of Context Copy.

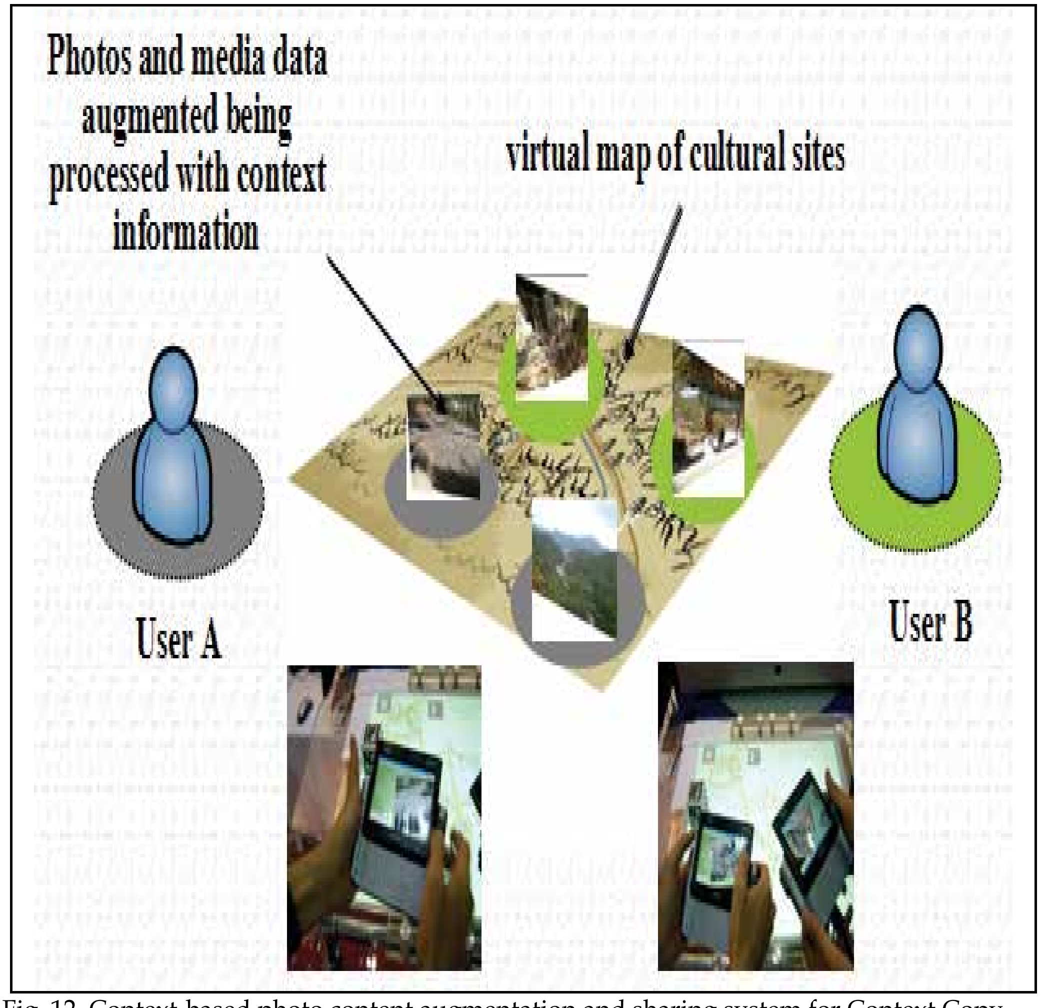


Fig. 12. Context-based photo content augmentation and sharing system for Context Copy

The personal context information such as time, location, etc. is managed in an individual mobile device. We allowed our system to show the distinct personalized media contents of two users depending on whether or not the user has visited Unju Temple. Here, we assumed that User A has been to Unju Temple, and that the pictures he took at several places around the temple are saved in his UMPC. When User A looks at a specific AR marker on a map on ARTable with his UMPC in his hands, the pictures he took at that place are augmented. Then, User A can flick through the pictures one by one. Here, the pictures are augmented in order of User A's preferences. ARTable (Park & Woo, 2006) is a smart table in which contents are dynamically displayed as a reaction to a user's interaction. In our system, ARTable constructs a space that allows multiple users to interact with services; here, a map with the paths around Unju Temple is displayed. To indicate a specific site on the map, we designed a particular AR marker in consonance with the surroundings in the map, and the system allows these markers to be detected. We extracted context appropriate information for application to this kind of system. The context information that is brought into play in the first phase is the location information on the sites the user has visited and taken pictures of. As shown in Fig. 13, ARTable displays a customized map of Unju Temple with the sites that the user has already visited indicated by AR markers. In addition, the larger the number of photos taken at a specific place, the larger the size of the marker indicating the place.

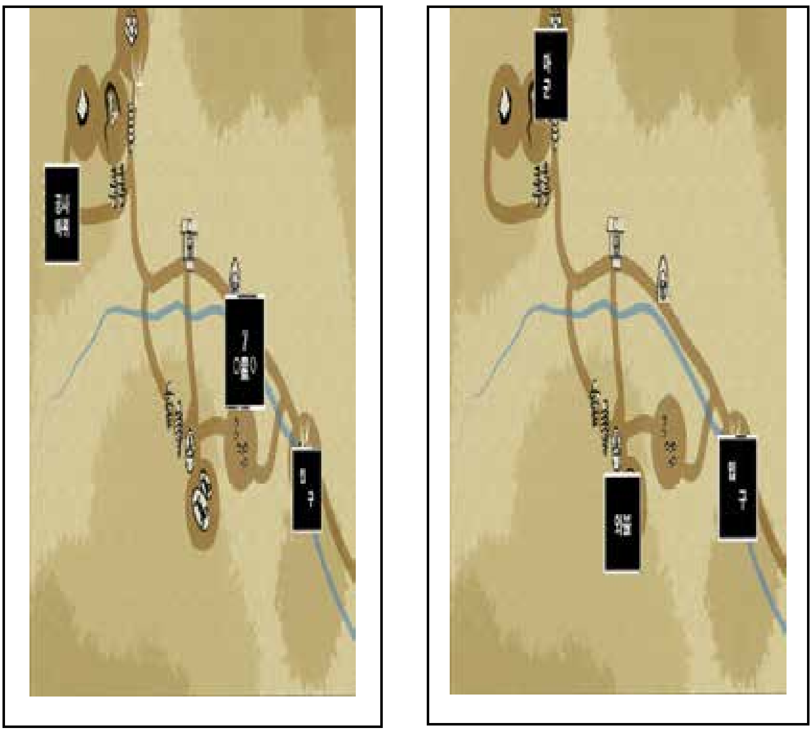


Fig. 13. Customized map of Unju Temple with the sites visited by the user indicated by AR markers

In the user's mobile device, the photos based on his or her context are then augmented. To determine the photos to augment, our system investigates if they are taken at the corresponding places as the specific marker that the user looks at with the built-in camera from among the markers displayed on ARTable. If we assume that User B has never been to Unju Temple, then the photos of Unju Temple would not exist in his UMPC. Thus, when User B uses his UMPC to view the same photo as the AR marker that User A sees in the map, the photos relating to that place are not augmented. Our system just provides the general information relating to that place without any personalized content augmented on the marker, as User B has no personal experience of visiting Unju Temple. In this context, User A has the option of then delivering the augmented photos from his UMPC to User B, allowing User B an opportunity to look at these photos. In addition, User B can receive abundant information related to the place through sound or animation augmentation. To this end, Fig. 14 illustrates context-based photo content sharing

Fig. 14. Context-based photo content sharing

5. Conclusion and Future works

With this exploratory study on CAMAR, we have shown a novel way of using context in a multimedia-based interaction that breaks from the preconceptions originating from the limitations of conventional AR applications. We believe that our work demonstrated the feasibility of personalized smart appliance AR controllers as well as context-based content augmentation and sharing systems. Besides enabling new ways of taking pictures of smart appliances in our daily life to control them, CAMAR broadened the possibilities of using context as a resource in new multimedia-based interaction techniques that naturally bridge the interface between human and mobile devices. Nevertheless, we need to conduct further user studies to see how useful our system is and accepted as a means of interaction for potential consumers by comparing with similar existing systems. Then, by considering the compatibility of the embodied technologies and contents, we will try to determine better contents for supporting the concretizing of user's experiences and sharing something meaningful to families through a CAMAR-embedded device. Moreover, to increase the satisfaction of the embodied technology, we should investigate the possibilities for controlling various smart but complex appliances with an intuitive interface on CAMAR-embedded devices. Subsequently, it will be valuable that we design CAMAR core platform to meet the system requirements and integrate them into an extensive framework.

6. References

R. Azuma (1997), “A Survey of Augmented Reality,” *Presence: Teleoperators and Virtual Environments*. vol. 6, no. 4, pp. 355-385, Aug. 1997.

R. Azuma, Y. Baillot (2001), R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *IEEE Computer Graphics and Applications*, vol. 21, Issue 6, pp. 34 - 47, Nov.-Dec. 2001.

W. Pasman and C. Woodward (2003), "Implementation of an augmented reality system on a PDA," *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 276 – 277, 2003.

J. Wither, S. Diverdi, and T. Hollerer (2006), "Using aerial photographs for improved mobile AR annotation," *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp.159 - 162, Oct. 2006.

M. Billinghurst, I. Poupyrev, H. Kato, and R. May(2000), "Mixing realities in shared space: An augmented reality interface for collaborative computing," in *Proc. IEEE Int. Conf. Multimedia and Expo.*, New York, Jul. 2000.

F. Farbiz, A.D. Cheok, L.Wei, Z. Zhiying, X. Ke, S. Prince, M. Billinghurst, H. Kato (2005), "Live three-dimensional content for augmented reality", *Multimedia, IEEE Transactions on*, Vol. 7, No. 3., pp. 514-523, 2005.

H. Matsuoka, A. Onozawa, and E. Hosoya(2002), "Environment Mapping for Objects in the Real World: A Trial Using ARToolkit," *Proc. First IEEE Intl. Augmented Reality Toolkit Workshop (ART02)*, Darmstadt, Germany, Sep. 2002.

C. Geiger, B. Kleinjohann, C. Reimann, and D. Stichling (2001), "Mobile AR4All," *Proc. The Second IEEE and ACM International Symposium on Augmented Reality (ISAR'01)*, New York, Oct. 2001.

A. Henrysson and M. Ollila (2004), "UMAR: Ubiquitous Mobile Augmented Reality," *Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia (MUM 2004)*, vol. 83, pp. 41-45, Oct. 2004.

S. Long, D. Aust, G. Abowd, and C. Atkeson (1996), "Cyberguide: Prototyping context-aware mobile applications," *In CHI '96 Conference Companion*, pp. 293–294, Apr. 1996.

S. Feiner, B. MacIntyre, T. Höllerer, and T. Webster (1997), "A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment," *In Proc ISWC '97 (Int. Symp. on Wearable Computing)*, pp. 74–81, Oct. 1997.

Y. Oh, M. Lee, S. Jung, and W. Woo (2004), "Dynamic Contents Provision of Context-based Information Augmentation & Sharing System," *ACM/IEEE ICAT04*, pp. 594-597, 2004.

M. Beigl (1999), "Point & Click - Interaction in Smart Environments," *Handheld and Ubiquitous Computing: First International Symposium, HUC'99, LNCS 1707*, pp.311-313, 1999

VV Badami and NW Chbat (1998), "Home appliances get smart," *Spectrum, IEEE*, vol. 35, pp. 36~43, Aug. 1998

M. Kohno, and J. Rekimoto (2005), "Searching Common Experience: A Social Communication Tool Based on Mobile Ad-hoc Networking", *ACM MobileHCI '05* pp.15~22, September 2005

Y. Lee, S. Oh, B. Lee, J. Park, Y. Park, Y. Oh, S. Lee, H. Oh, J. Ryu, K. H. Lee, H. Kim, Y. Lee, J. Kim, Y. Ho, and W. Woo (2005), "Responsive Multimedia System for Context-based Storytelling," *LNCS (PCM)*, vol.3767, pp.361-372, 2005.

Y. Lee, S. Oh, and W. Woo (2005), "A Context-based Storytelling with Responsive Multimedia System (RMS)," *LNCS(ICVS)*, 3805, pp. 12-21, 2005.

Y. Oh, C. Shin, W. Jung, and W. Woo (2005), "The ubiTV application for a Family in ubiHome," *2nd Ubiquitous Home workshop*, pp. 23-32, 2005.

Y. Park and W. Woo (2006), "The ARTable: An AR-based Tangible User Interface System", *LNCS (Edutainment)*, vol.3942, pp.1198-1207, 2006.

H. Kato and M. Billinghurst (1999), “Marker tracking and HMD calibration for a video-based augmented reality conferencing system,” *Proceedings of 2nd IEEE and ACM International Workshop on Augmented Reality, (IWAR '99)*, pp. 85~94, 1999.

S. Jang and W. Woo (2005), “Unified Context Representing User-Centric Context: Who, Where, When, What, How and Why,” *ubiComp workshop* (*ubiPCMM*), pp. 26-34, 2005.

Z. Yu, X. Zhou, D. Zhang, C. Chin, X. Wang, J. Men (2006), "Supporting Context-Aware Media Recommendations for Smart Phones", Pervasive Computing, IEEE, Vol. 5, No. 3, pp. 68-75, 2006.

19

User Experience in Digital Games

Jari Takatalo*, Jukka Häkkinen+, Jyrki Kaistinen* & Göte Nyman*
University of Helsinki, +Nokia Research Centre
Finland

1. Introduction

This chapter concentrates on the psychological analysis of the user experience (UX). We present three different psychological frameworks that have been used in the study of complex UX environments. They have been widely applied in different game contexts studied by the PDL (Psychology of Digital Life) research group at the University of Helsinki. The frameworks work their way from very general and theoretical aspects towards more specific and empirically validated model of the UX in games. Both theoretical and methodological issues are presented, including the empirical data collected (n=2200) and the directional results that demonstrate how the psychological UX framework works in practice. Some other technology environments are also considered within the fundamental psychological concepts that all these frameworks share, that is, cognitions, emotions and motivations. It is shown, how basic psychology can be applied in order to understand the rich variety of human experiences. There are profound reasons to study games. They provide interactive entertainment for millions of users around the world. Economically they have exceeded traditional media, such as movies and music. Game companies have become forerunners of the software industry, guiding the future of the whole field and having impact also on hardware development. Gaming technology will be applied in numerous other areas as socially interactive virtual worlds (e.g., 2nd Life) have already shown. New generations familiar with being and interacting in virtual environments are born. The daily activities of these citizens are shifting towards the world wide virtual world. This change has made games and virtual worlds also true topics in cultural discussions. Since games are played to get experiences, they are hard to study and understand with traditional usability methods that are typically applied in the field of human-computer interaction (HCI). This will be problematic when the future user interfaces evolve towards those used in games today. Thus, more holistic approaches to understand and evaluate the inner worlds of the users and their complex functionality in numerous technological contexts are required. When the holistic nature of the human experience is studied systematically with methodologically solid approaches, psychological understanding of the human experience in such environments increases and benefits the designing of value and UX for these environments.

2. User Experience

The study of UX arises from the need to understand the complexity of HCI. Conventionally in the field of HCI, applications are evaluated by task- and work-related usability metrics (Hassenzahl & Tractinsky, 2006). However, concentrating on external behavior, such as, cognitive efficiency and objective measures of bottom-up processes does not reach the relevant psychological phenomena that technological development has brought to people. Computers have moved from offices to homes, telephones have been shrunk to our pockets and demanding customers want experiences in addition to efficiency. Bottom-up usability issues remain relevant but a usable gadget needs to feel good, look smart, sound personal, and bring some added color to its user's inner life. In order to study and evaluate such deep psychological phenomena, a top-down approach to user psychology is needed. There is a simple reason for this: there is no well-defined bottom-up theory of human experience that would substantially contribute to the understanding of complex behavioral contexts, such as playing digital games.

The field of UX research is young and it still faces problems, such as incoherent concepts and lack of empirical research (Hassenzahl & Tractinsky, 2006). Many of the studies are based on authors and designers personal experiences in the field and often the user's point of view is neglected (Fabricatore et al., 2002). However, there are also serious attempts towards consensual definition of UX (e.g., UXSIG, http://www.cost294.org). In these approaches, many essential psychological components are included (e.g., feeling, satisfaction, need, mood, motivation). However, many definitions would do better if they were more clearly aimed at describing what mental compartments are included into analysis and what experiential attributes are evaluated. This is a demanding task. It makes little sense to create lists of numerous mental ingredients that neglect the well established psychological research on human experience that offers scientifically validated, theoretical and empirical material for these purposes. Also in most of the UX definitions a product, service, event, design or system is referred as the object of the interaction responsible for UX. However, each of these cases has its own nature that should be regarded, in order to understand UX.

Naturally, UX in digital games and cellular phones is quite different. The past experience of the user and the context of use bring in variables that are both general and specific to the context. If we have a clear understanding of the dimensions and attributes of these experiences their measurement becomes easier. If we want to understand how two persons experience a red star, for example, it could be quite difficult to assess or even compare these two experiences without an idea of what we want to measure. With a general enough approach we can get an idea of what is the quality, intensity, meaning value, and extensity of the red star for these two persons. These attributes stem from the general psychological compartments -cognition, emotion and motivation -and they make operationalizing, measuring and comparing experiences easier. The same general psychological constructs should be able to use to assess different forms of new technology in order to understand UX. But if we do not know why and how they observe these celestial objects, we cannot understand their experiences, let alone measure them in a valid way. People do similar things for completely different reasons.

2.1 Psychology of User Experience

Human experience is a very complex process. Perhaps, the most well known explanation about it is the notion used by William James: "the content of consciousness is experience" (James, 1890). In order to understand, how experience evolves in consciousness, relevant psychological constructs need to be defined and used, but in a suitable, simplified form. Firstly, we need to know, what are consciousness and awareness. In the introduction part of his book Optimal Experience – Psychological studies of flow in consciousness, Mihaly Csikzentmihalyi (Csikszentmihalyi & Csikszentmihalyi, 1988) gives a general overview of the structure and functioning of the consciousness. Based on, for example, (Pope & Singer, 1978) Csikzentmihalyi suggests that consciousness can be divided into three subsystems: 1) attention 2) awareness and 3) memory. Lots of environmental stimuli are perceived but only a minor proportion of this is interesting enough to draw our attention and to become a content of our consciousness. Most of our daily routines are experienced rather automatically or sub-consciously (Forlizzi & Ford, 2000). Those perceptions that draw our attention enter the consciousness and become interpreted by the awareness. Awareness can be better understood by its three main processes or faculties originally proposed by Moses Mendelsshon in 1795, they are: understanding (cognition), feeling (affection) and will (conation) (Hilgard, 1980). Over decades this trilogy has been considered to concern human cognition, emotion and motivation (Mayer, 2001). Thus, we perceive and focus our attention to stimuli that motivate and interest us (James, 1890). The cause for the interest and motivation may originate from the environment (e.g., survival), but it is often our intrinsic needs that motivate our perceptual processes and focus of attention. Cognitively we recognize and relate these stimuli with each other and with our past experiences stored in our memory (Glenberg, 1997). Such an interpretation process is informational in nature, and it is enhanced by emotional labels that are attached to it (Lazarus, 1991a). Damasio (Damasio, 1994) describes the role of the somatic markers in the body as crucial in emotional labeling of our perceptions. He pointed out that our cognitive reasoning would be impaired without such bodily reference to a certain stimuli. These somatic markers stem from our bodies and they are felt as emotions and feelings in our awareness (Fig 1.).

When we understand our perceptual-attentional processes, cognitive reasoning, emotions, personal relevance, and past experiences related to a certain events and objects only then we can have an idea of the central experiential attributes involved in any human activity. Needless to say, this is a formidable task. However, we can begin our journey of discovery from events that have a clear beginning and an end, such as playing a video game (Dewey, 1934). Then we would need a psychological approach that provides us heuristics (Takatalo et al., 2007), within which we can evaluate different aspects of the experience; its content, intensity, meaning, value, quality, and extensity (James, 1890; Wundt, 1897). Being able to measure these attributes helps us to evaluate and rank anything that a person interacts with, weather it is a digital game or a cellular phone. The application of the presented psychological framework to such use cases is quite complicated and requires a

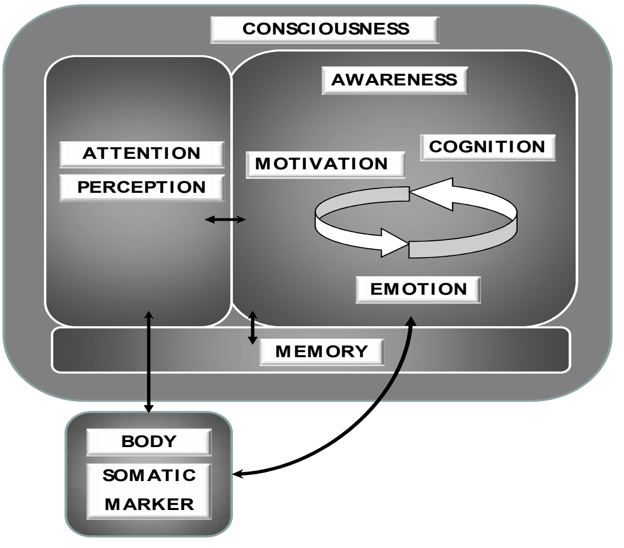


Fig. 1. The reciprocal relationship between body and mind in forming the experience

deep analysis of the used technologies, users and the contexts of use. In our red star example, for example, the experiences of two persons may vary greatly if one is a die-hard fan watching a football game in "Marakana" stadium, the home of the Red Star Belgrade and the other is an American anti-communist, who has lived through the cold war. By analyzing their cognitions, emotions and motivations quite different meaning and value is attached to the red star. There is no "pure" perception of a red star - or any other stimulus. Also the intensity and quality of an experience received may vary quite a lot. When we start analyzing such experiences that we receive from our environment, it becomes obvious that the mind is not composed of boxes that work as a simple linear system. Keeping in mind that these components are essential parts of the mind, we combine them into wider psychological concepts and frameworks so that we can try to utilize and even measure them in practice. This is what we have tried to do.

2.2 Need & Motivation, Adaptation and Appraisal

Here we show how the above psychological components are likely to be connected in an experiential cycle. We use the constructs of need and motivation to refer to interests and curiosity that guide our attention and environmental perceptions. Next phase in this experiential cycle is denoted as adaptation. It means adaptation to, for example, a situation, environment or content and it is considered as one of the central functions of the consciousness (Angell, 1907). Consciousness has been considered to mediate between the needs of the complex being and the demands of the complex environment (James, 1890). In nature, the life has always been easier for those who have been able to adapt into the requirements of the environment. During the stage of adaptation, the environmental features become cognitively interpreted and referred to memory structures of the past experiences. Adaptation is closely related to the third stage - appraisal, which is an evaluation of the knowledge about the current situation of our personal goals and well-being in a particular environmental interaction. Because such knowledge is important and meaningful to us, emotions occur in relation to them (Lazarus, 1991a). In short, appraisal combines cognitive evaluation of the goal-directed interaction in an environment with emotions and feelings evoked by these interactions. In addition, appraisal has an impact on our motivation to either maintain or re-direct our ways of interacting. In this way the experiential cycle circles along the stages need & motivation-adaptation-appraisal. But, there are abundant environmental objects (product, service, event, design, or system), sub-features of them (e.g., cellular phones, software) and different ways and contexts of using them (communicating, taking photos ,word processing, or playing games). All these contribute to UX by affecting the way this cycle is rolling. It is a complex and dynamic process.

In general, technologies are designed to bring added value to our lives and extend our capabilities. Ideally, it would be beneficial to fit variety of technologies into the UX framework such as experiential cycle. This would allow the measurement and understanding of how UX evolves in the human mind in general. Naturally, this would be limited to a very superficial level and when more accurate and practical measurement models are requested, a detailed analysis of the studied technologies, users and contexts is needed. Here we have embedded game software and cellular phone into the experiential cycle in order to demonstrate complex, but still understandable UX-related phenomena. We then continue to explain UX in games in more detail. Users in the experiential cycle example are restricted into experienced ones, who engage with a game or a phone on a daily basis and have already used them for a while. The use context is restricted to communication with a distant friend and a first person driving game played offline against the computer.

Need and motivation to use a phone to communicate with a distant friend stems from its capability to extend our natural resources and to provide a good set of options in order to become connected. Depending on the nature of the communication need and the nature of the friendship, either a phone call or SMS can be used, for example. Quite similarly, also games provide added value to our life. They enclose us into places, stories and activities that elicit rich experiences. The reason to initiate the use, the context of use and the way the phone is used or what game is played are all crucial factors that contribute to UX, and hence to the way it should be measured. They all have an impact on how we attend to and perceive the technology at hand. Need & motivation are also strengthened by the ongoing perceptual process; the way technologies feel, look and sound have a psychological impact every time they are interacted with. Due to this whole process, the focus of the evolving UX shifts towards interpretation and adaptation.

Having the need to use the cellular phone the user needs to adapt to the user interface and to the possible options (call, sms) available. We could say that a good UX with a phone is offered by a technology that allows as good adaptation to the user's needs as possible. An optimal technology should become our natural extension that supports playful and enjoyable interaction with the environment. Also the game software attempts to support playful interaction between a user and a game environment. However, in a game the interaction (its goals) is created by the game itself and this is accepted by the user simply by selecting this specific game software. In addition to interactivity, a game creates an illusion of space in which the interaction takes place. In both our cases system features such as adjustability and personalization of the technology are likely to deepen the adaptation. Motivation and adaptation are also likely to go hand in hand; if we are very motivated to use something then the adaptation is more desired and requirements of the system can be even lower. This was quite typical in the use of the first computer games. Only now it is easy to see how primitive (but still adaptive) they were. On the other hand, motivation and a usable and fancy interface do not carry far without proper functionality and meaningful content. However, there are some differences in the way these two cases, phone and game, are appraised.

We evaluate the current situation of our personal goals and well-being based on the relevant outcomes of person-environment interactions. The goal of the phone use is to get connected with a friend. However, there are two different interactions going on that are appraised: physical interaction between a user and a phone as well as social interaction between a user and a friend. If there are, for example, challenges in the first interaction they are most often due to bad usability or technical problems. On the other hand, challenges in the other interaction are related to the content of the conversation. This makes it difficult to evaluate how technology alone actually affects UX. The critical issue then is the technology's ability to support the chosen way of communicating; does it make it better or worse, does it bring new dimensions to it such as anonymity. In the long run, issues such as the brand, social status, durability, and liability of a phone have their significant contributions to UX.

In our other case, a game is appraised based on its ability to provide experiences. Because games provide strong and real-like experiences by transporting the user psychologically away form the real world, the goals provided by the game become part of the users goal to go and experience what is afforded. Thus, the appraisals of the in-game interactions strongly affect UX in games. Also in playing games, clear technical challenges diminish the UX. On the other hand, in-game challenges are natural and important part of a good UX in games. In both our cases the appraisal of the situation is likely to lead to emotions, which again guide our motivation to play and use cellular phone again or search better ways to fulfil our needs to get experiences or communicate with a friend (Lazarus, 1991a).

We have shown that compressing general psychological components into larger concepts are useful in evaluating evolvement of UX in different kinds of technologies. When more detailed understanding of the UX is requested, more accurate measurements and methods are needed. We continue on analyzing UX in games in more detail.

3. UX in Games

Games are played to get experiences. The way to assess such experiential process may vary, but the selected framework and method should not disregard any part of the mind's trilogy. Analysis of this trilogy reveals the quality, intensity, meaning, value, and extensity of the experience.

Although games are studied in the field of HCI, there is a need for a systematic empirical analysis of the UX. Common to many game studies is that they consider UX in games as fun, which makes players enjoy playing and motivates them to keep on playing (Sweetser & Wyeth, 2005). Challenge (goals and uncertainty), fantasy (emotions and metaphors) and curiosity (novelty and surprise) were probably the first empirically based guidelines to produce more interesting, enjoyable and satisfying computer systems, such as games (Malone, 1981). Since then numerous heuristics to design better games have been introduced (Desurvire et al., 2004). Indeed, such heuristics can be used to suggest, what to include into a game. However, to evaluate such a game a psychologically grounded approach to the problem is needed. Theoretical frameworks such as MDA (Mechanics, Dynamics, and Aesthetics) attempt to integrate game structure (e.g., mechanics) to experiential outcomes (Hunicke et al., 2004). But also they share the same problem with the heuristics; they are validated mostly by the professional game designers and developers (Barr et al., 2007). In order to empirically grasp the psychological core of UX, theoretical models are important starting points on which to lay ground for the study.

Fun is often considered as the main motivation to play, thus there are studies that concentrate on the sources of fun in gaming. For example, in an empirical study the motivation of the 30 players and 15 non-players to play or not to play digital games was studied (Lazzaro, 2004). The results indicated four major emotional motivations to play games. These "four keys" were challenges (overcoming obstacles), grab attention (curiosity, excitement, adventure), altered states (emotions and sensations) and other players (competition, co-operation with others). Similar uses of gratification dimensions were factor analytically extracted in a larger study (n=550) (Sherry et al., 2006). When these two studies were compared, the latter one specified altered state to level of arousal, grab attention to fantasy and other players to two distinct dimensions of competition and social interaction. It also included one new dimension, that is, diversion from other things, such as studying. Larger empirical samples (n=3200) have also been studied in order to understand player's motivations (Yee, 2005). Although these studies were conducted within a one genre (MMORPG), the results of the principal components analysis share similarities with the above cross-game studies. Concentrating on one genre enables a more accurate list of in-game challenges (e.g., advancement, competition). However, to uncover the general experiential laws in games, a cross-game sample and systematic psychological analysis of UX is needed.

Empirical studies in general can provide a reliable insight into the user's inner world. They should show how various experiential phenomena can be mapped, based on the empirical data. However, the above studies appear to fall short of identifying standard set of mental constructs and the way these jointly affect UX. A more dynamic and structural approach is adopted in the Microsoft Game Studios (Pagulayan et al., 2003). In these empirical, user-centered, studies the goal has been to find standard components to assess the formation of fun. These studies show that fun is likely to stem from various sources depending on the complex relationship between the game and the player. One of the key components of this approach is challenge, the level of which is evaluated with subjective methods.

The findings of the gaming studies can be generally summarized in terms of the following components of fun: **Challenges** (overcoming obstacles, clear goals), **Emotions and internal sensations** (arousal, enjoy, relax), **Fantasy** (adventure, escapism), **Curiosity** (discovery, attention, exploration, learning), **Other players** (social interaction, competition, co-operation) and **Narrative** (drama, role). In addition to these, wider concepts such as immersion (Brown & Cairns, 2004; Davidson, 2003; McMahan, 2003; Sweetser & Johnson, 2004), presence (McMahan, 2003; Nakatsu et al., 2005; Pinchbeck, 2005; Ryan et al., 2006; Takatalo et al., 2006b; Takatalo et al., 2004), flow (Nakatsu et al., 2005; Sweetser & Wyeth, 2005; Takatalo et al., Submitted), involvement (Davidson, 2003; Takatalo et al., 2006b), engrossment and engagement (Brown & Cairns, 2004) are often used to explain fun and UX in games. But, immersion, for example, is such a wide concept that it includes all the above experiential aspects. So, it would be more useful to measure its psychological components in order to recognize its basic elements. Because of this kind of psychological challenges in measuring UX, we have developed a Presence-involvement-Flow –Framework (PIFF^2) that aims at measuring essential psychological components of UX in games. PIFF^2 can be seen as a game specified UX framework that is in accordance with the presented general psychological framework and the experiential cycle. It just goes deeper in order to provide a holistic understanding of the UX in digital games and explaining how UX gets its quality, intensity, meaning, value, and extensity.

4. Presence, Involvement and Flow in Games

4.1 Involvement and Presence

Issues related to need & motivation and adaptation in our experiential cycle are dealt in PIFF^2 with more game-related concepts: involvement and presence. First of all, players must invest time, effort and attention into a game in order to get any relevant experience from it (Brown & Cairns, 2004; Davidson, 2003). This can be measured with an involvement construct (Zaichkowsky, 1985), which is defined as a continuum of unobservable state of motivation, arousal or interest towards a particular situation or stimulus (Rothschild, 1984). The involvement construct and its two distinct but correlated dimensions importance and interest (McQuarrie & Munson, 1992) together assess the psychological depth and quality of the player-game relationship. The psychological nature of importance is dominantly cognitive and it concerns the meaning and relevance of the stimulus, e.g., what matters to the player, whereas interest measures emotional and value-related valences, with response items such as “it was exciting” (Schiefele, 1991).

Presence has been studied in variety of media, for example, virtual environments, television, movies and digital games (Lombard & Ditton, 1997). The existing studies indicate that the experienced presence can significantly vary with the technology used. Lombard and Ditton

(Lombard & Ditton, 1997) conceptualized the presence in mediated environments as a combination of physical and social presence. They differentiated three components in the physical presence: attention (psychological immersion), perceptual realness (naturalness) and spatial awareness (transportation). This threefold construct has also been confirmed in previous factor analytical studies (Lessiter et al., 2001; Schubert et al., 2001). In addition, the range and consistency of the physical interaction is considered an integral part of the sense of presence (Lombard & Ditton, 1997). Some authors even see it as the only determinant of the presence experience (Zahorik & Jenison, 1998). Thus it is also valuable and even necessary to be considered in interactive digital environments such as game worlds (Davidson, 2003; Sweetser & Johnson, 2004).

The sense of presence is not related to physical aspects alone but to the social scope of the technology environment. This type of presence experience needs to be considered especially in mediated environment with social content. Such environments are likely to elicit the sense of social presence. In the Lombard and Ditton's (Lombard & Ditton, 1997) explication, social presence was composed of social richness (intimacy-immediacy), social realism and co-presence (shared space). Social richness is "the extent to which a medium is perceived as sociable, warm, sensitive, personal or intimate" (Lombard & Ditton, 1997). Social realism refers to the sense of similarity between real and game-world objects, people and events. In gaming the perceived drama and plot and engagement to own role in the storyline fit well in to this aspect of social presence. Co-presence is the feeling of being and acting there in a game-world together with other agents. Often such a social impact is strongly demonstrated in situations where participating agents have the same object of interest as the player has.

Together involvement and presence describe how players voluntarily form a relationship with a physical and social aspects of a digital game, that is, adapt to it (Takatalo et al., 2006b; Takatalo et al., 2006c). Taken together these two distinct dimensions form our adaptation measurement model. Psychologically, it describes the perceptual-attentive, motivational and cognitive aspects of the UX in games. In addition, arousal regulation is intimately linked with the human attentive system: high level of emotional arousal enables greater allocation of attentive resources into a particular event or stimuli (Kahneman, 1973). Together these psychological components describe the meaning and value as well as the intensity and extensity (i.e., voluminous or vastness, a spatial attribute) of the UX in games.

4.2 Flow

Appraisal part of the experiential cycle is approached with the theory of flow (Csikszentmihalyi, 1975). Csikszentmihalyi (Csikszentmihalyi, 1975) defines flow as a positive and enjoyable experience stemming from interesting activity that is considered worth doing for its own sake. In a state of such an optimal experience, individuals tend to be playful (cognitively spontaneous, inventive, and imaginative) (Ghani & Deshpande, 1994; Novak et al., 2000; Webster & Martocchio, 1992), self-consciousness is lost, action and awareness merge, and time passes more rapidly (Csikszentmihalyi, 1975; Csikszentmihalyi, 1990; Ghani & Deshpande, 1994). In addition, concentration, clear goals, instant feedback, and a sense of control are considered to contribute to flow (Csikszentmihalyi, 1990).

However, the right number and the relevance of flow factors are not clear (Finneran & Zhang, 2002). For example, losing self-consciousness and the merging of action and awareness have been found to be difficult for respondents to recognize (Rettie, 2001). On the other hand, in almost every study related to flow, the two cognitive key antecedents – evaluations of the challenges provided by the activity and the skills possessed by the respondents – are included.

Every time people engage in a meaningful activity, a mental process is activated where the evaluation of its challenges and the skills it requires occurs (Csikszentmihalyi, 1975). Different ratios between these two are likely to lead to different emotional outcomes. There are a few different flow-channel models that share this basic idea of having various emotions due to an evaluation process that concerns human-environmental interaction. For example, the eight-channel model (Massimini & Carli, 1988) includes eight different emotional outcomes and two different cognitive evaluations. A positive state of flow evolves through a process in which both the skills and the challenges are evaluated as being high and in balance. Psychologically, the core idea of the flow theory (Csikszentmihalyi, 1975) is similar to cognitive theories of emotions (Ellsworth & Smith, 1988; Frijda, 1987; Lazarus, 1991b). These theories suggest that cognitive interpretations and appraisals of events in the world are necessary parts of emotions. Some neuroscientific data seems to support these findings (Roseman & Evdokas, 2004). There are various appraisal features and components, such as the effort anticipated in a situation, perceived obstacles, and the sense of control, all of which shape the emotions attached to these events (Ellsworth & Smith, 1988).

Psychologically, the flow part of the PIFF^2 describes the cognition, emotion and feeling. Also memory and previous experiences have an affect on the cognitive appraisal process and the forming of emotions and feelings depicted here. The cognitive and emotional profiles provide by the flow model give insight to the both qualitative and intensity attributes of the experience and their cognitive predecessors. Combining adaptation and flow measurement models into one framework has a strong theoretical foundation. Together the two models give us a holistic profile of the content of the UX in games. These profiles are based on players' subjective interpretations of the game event, made within the pre-set psychological boundaries.

4.3 PIFF^2: Methodological Issues

We emphasize the importance of evaluating the conscious top-down experiences. We have used subjective methods (e.g., interviews and questionnaires) that allow users reflect their own experiences. Such subjective analysis methods have a long tradition in the field of psychometrics and behavioral sciences to assess, for instance, attitudes, aptitudes, interests and personality. However, there has been a debate against the use of subjective methods in analyzing subjective experiences such as the sense of presence (Slater, 2004). Naturally, all methods have their pros and cons. Critics are appropriate towards questionnaires including question items with unfamiliar or many-faced constructs. Sometimes researcher may cut corners and develop single questions or simple scales to study multidimensional phenomenon such as presence. This is a major fault methodologically leading to unreliable scales (Cronbach, 1990) and poor construct validity of the measurement model. In our previous study we have shown what will happen if an oversimplified measure of presence is used to analyze difference between four different PC-games (Takatalo et al., 2006c). In our study the measured experience of presence digital games context included five dimensions. When these dimensions were grouped into one “meta-presence” dimension, all the games scored high in presence. As the games were studied with separate presence dimensions, clear differences between games were found. This example clearly demonstrated the multidimensionality of the presence construct. It also shows that presence is a latent psychological construct (Tarkkonen & Vehkalahti, 2005) that cannot be directly reached. Compressing five distinct presence dimensions into one “meta-presence” dimension is analogous to the simple presence scale measuring only one aspect of presence. However, in digital games context the sense of being there is composed of sub-scales such as attention and role engagement (Takatalo et al., 2006b). An extra positive outcome of such an approach is that the participants of the studies have easier to understand what is meant by subjectively more observable components. Thus, the subjective phenomena become easier to assess.

To analyze complex issues such as UX in the field of HCI, any reasonable method should not be overlooked. When subjective methods are used, special care must be taken in considering both reliability and validity of the measurements. It is not enough to concentrate on the statistical side of the scale construction (how to measure) and ignore the theoretical issue of what is measured. If the scales are made up without any theoretical background, they can measure totally different things than they were supposed to do, despite the high levels of internal consistency of the scales (Slater, 2004). When human mind and subjective experiences are concerned, the scales measuring them should have some relationship to basic psychology otherwise the H - with all its history and theoretical developments in psychology - in HCI will forgotten. Similarly, if psycho-physiological methods (e.g., electromyography) are used without proper knowledge of human anatomy the results may be reliable but not valid in any way. For example, if electromyography (contraction of facial muscles) electrodes are attached to one of the thigh quadriceps before playing a game, a consistent and accurate response graph will be obtained but it has no relevant meaning. The above theoretical and methodological issues in mind we have collected empirical data on digital games context and statistically extracted measurement scales forming PIFF[2].

5. The Data to Form Measurement Scales

5.1 Origin and Collection of the Dataset

The data have been collected from both laboratory experiments and an Internet survey using the EVE-Experience Questionnaire (EVEQ-GP) (Takatalo, 2002; Takatalo et al., 2004). In the field of behavioral sciences the use of questionnaires has proven to be a valid way of assessing an extensive number of mental phenomena (Breakwell, 2006; Couper, 2000; Labaw, 1980; Rust & Golombok, 1999). Both the paper and pencil and the online versions of the EVEQ-GP are composed of 180 items (1-7 Likert-scale and semantic differentials) measuring different experiential aspects obtained from being, performing and experiencing the game world. When the data was collected outside laboratory, the instructions for completing the EVEQ-GP encouraged participants to reflect on their subjective gaming experience of one particular gaming session that they have just finished. The instructions emphasized that the questionnaire was to be completed immediately after a gaming session. In the laboratory, the situation was straightforward as the players just described the game they had just played there. Thus, the gaming experience was operationalized as a situated experience stemming from one particular game. The method used enabled the player to report, within pre-set multidimensional boundaries, how it felt to interact with a specific digital game. Also included were 27 background questions.

An online version of the EVEQ-GP (VK2) was used to collect data from the Internet. Participants were asked to focus on one particular gaming session and fill in the questionnaire while keeping that session in mind. The instructions recommend filling in the questionnaire right after a playing session. The application development software used to create VK2 was Lotus Domino Designer 6.5. Domino Server ran on HP Proliant DL380. The questionnaire was online for one month on the home page of the Pelit [Games] magazine (www.pelit.fi). Pelit is a leading PC gaming magazine in Finland, with a circulation of approximately 38 300 and registered online users approximately 27 000. During the first week VK2 was on the main page; for the remaining three weeks it was linked to a short news story, located in the news section. One month on the Internet resulted in 1912 properly filled-in questionnaires.

In addition, two distinct laboratory experiments were conducted. In the first, 240 university students (120 males, 120 females) played two different driving games with two different displays. In this experiment a 2x2 between subject-design was used. Each participant played for 40 minutes, after which they were asked to fill in the EVEQ-GP. In the second laboratory experiment, 30 university male students played Halo: Combat Evolved for two consecutive sessions. After the second session they filled in the EVEQ-GP. Results from these particular studies have been reported elsewhere (Takatalo et al., 2004).

5.2 Description of the Dataset

The data set consist of 2182 subjects (1972 males, 210 females), who filled in the questionnaire. The mean age of the respondents was 21.5 years (SD=6.0). The average time of playing was 127 minutes (SD=111) and the average size of the display used was 19.2" (SD=4.4). 33% of the respondents played daily, 29.6% played at least every other day, and 24.5% played often but not as often as every other day.

Most of the games played (31.5%) before the questionnaire was filled in were first-person shooters (FPS) either online (15.0%) or offline (16.5%). The second most popular genre (15.0%) was massive multiplayer online role-playing games (MMORPG), and the third (13.1%) was single role-playing games (RPG) (13.1%). The most popular single game played was World of Warcraft (n=265), which is a MMORPG. Altogether the data included approximately 320 different games, giving a broad scope to the psychology of digital games.

Since Pelit magazine focuses on PC games, 85.2% of the games were played with a PC and 14.8% with a console.

6. Measurement Scales

The first version of the Presence–Involvement-Flow –framework (PIFF) was based on two, earlier collected and smaller datasets (n=68 & n=164). It included 23 scales measuring UX in games (Takatalo et al., 2004). After increasing the total sample size into 2182 participants, a factor analysis was conducted to the sample and 15 measurement scales composed of 139 individual variables were extracted (Takatalo et al., 2006b; Takatalo et al., Submitted). The resulting framework was thus named as PIFF^2 and it integrates two separate measurement models that assess presence and involvement (Takatalo et al., 2006b) as well as flow (Takatalo et al., Submitted). In addition to statistical psychometric validation the framework was also grounded to previous gaming studies, studies concerning both presence and flow and relevant psychological concepts.

6.1 Involvement and Presence: Adaptation

Of the total 180 EVEQ-GP items, 83 items were used to form an **adaptation** measurement model (Table 1.). A factor analysis (oblimin rotation) allowed us to extract 8 experiential, underlying dimensions. Together they describe how players voluntarily form a relationship with physical and social aspects of a digital game, that is, how they adapt themselves into a game-world (Takatalo et al., 2006b). The extracted scales have already been applied to analyze adaptation in different games (Takatalo et al., 2006c) and display types (Takatalo et al., 2006a). To learn more about the origin and previous use of the items the reader is referred to (Takatalo, 2002; Takatalo et al., 2004).

Of the 180 items, **flow** is measured with 56 items (Takatalo et al., Submitted). It is composed of two cognitive and five emotional dimensions. Together they depict both the cognitive evaluation and emotional consequences of a gaming session. They also show that different emotions result from a particular combination of cognitive evaluation and that clearly more complex feelings are related to gaming than just simply fun. Although the cognitive evaluation of the Interaction scale was extracted in the adaptation measurement model, it is dealt with two flow-related cognitive evaluations of competence and challenge in the results. All the statistical analysis was conducted with SPSS 13.0 statistical program.

7. Case: Display and Gender in UX

In this general case example we will show how PIFF^2 works in practise. The aim is to show, how our multidimensional framework of UX in gaming is affected by the background of the user as well as the technological context, in this case, simply the form of the display. The results are shortly discussed and referred to experiential cycle as well as the experiential attributes. We do not go into technical or methodological details in this case. More trough analysis is presented elsewhere (Takatalo et al., 2006a). The case is restricted to one driving game (Need For Speed Underground), which was played for 40 minutes by 148 participants

(75 women and 73 men) in a lab. After the gaming session, they filled in the EVEQ-GP questionnaire. The game was played with five different displays (CRT_1, CRT_2, stereo_1, stereo_2 and HMD) in a between-subjects design. The following background variables were used as covariates in an ANCOVA analysis in this example: experience with computers, skill in driving games, digital gaming frequency, driving gaming frequency, attitudes towards computers, attitudes towards driving games, average computer gaming time, and computer usages in a week.
Table 1. Name, number of items, Cronbach's alpha coefficient and short description of the scales forming the Presence-Involvement-Flow framework.

No gender differences were found in any adaptation scales. In flow scales, significant independent of covariates effects were evident in competence and challenge. Males

Scales	Name & nro.of items		α	Description
ADAPTATION				
1	Role Engagement	12	.87	Captivated and enclosed into the role and place provided by the story
2	Attention	12	.89	Time distortion. focus on the game world instead of the real world
3	Interest	6	.80	The game was interesting, exciting as well as lively
4	Importance	8	.90	The meaning of the game, game was relevant, close, personal and sensitive
5	Co-Presence	14	.89	Feeling of sharing a place with others, being active in there
6	Interaction	9	.74	Speed, range, mapping, exploration, predictability of own actions
7	Arousal	5	.64	Active, stimulated vs. passive, unaroused
8	Physical Presence	17	.88	Feeling of being transported into a real, live and vivid place
FLOW				
9	Valence	10	.86	Positive valence, happy, not bored or anxious
10	Impressiveness	9	.75	Amazed and impressed by the game-world, the game elicited real feelings
11	Competence	11	.87	Skilled, competent, enjoying using the skills, clear goals
12	Challenge	5	.69	Game was challenging, game required the use of my abilities
13	Enjoyment	7	.83	Playing was pleasant, enjoying and exciting, I'll recomend it to my friends
14	Playfulness	9	.78	Ease of doing, creative, live and vivid, not unimaginative
15	Control	5	.71	Feeling of being in control and independent

Table 1. Name, number of items, Cronbach's alpha coefficient and short description of the scales forming the Presence-Involvement-Flow framework[2].

evaluated their competences higher and females the challenges provided by the game higher. After controlling for the attitudes toward driving games and the duration of a typical gaming session, females considered gaming more playful. Females considered gaming also more enjoying, when both the attitudes towards new technology and towards driving games were controlled for. Also non-significant independent tendencies were found, males with equal previous experience with car games considered to be marginally more in control whereas females considered the game more meaningful and personally relevant (Fig. 2.).

To sum up, gender, independent from background of the users, makes participants evaluate themselves differently in the gaming situation. The difference in competence-challenge ratio indicates different cognitive abilities between genders. Fast paced, 1st person driving game requires skills such as 3D rotation and field independent spatial perception, in which males have advantage (Sherry, 2004). Despite this, females with equal previous experience and attitudes had qualitatively richer UX than males. This shows the complex nature of the UX: being competent and in control is not always enough for a qualitatively rich UX. Personal relevance and meaning have also impact on the quality of UX. It shows that males tend to

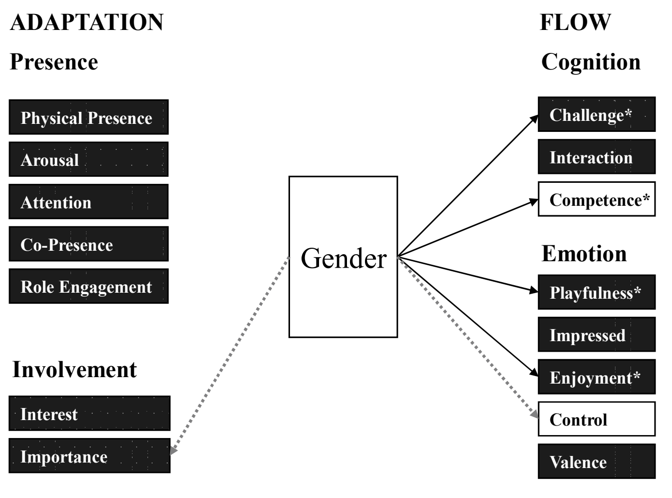


Fig. 2. Independent effects of on the PIFF² factors. Lighter boxes indicate higher scores for men, darker boxes higher scores for women. * indicates $p<0.05$, dashed lines indicate marginally significant effects.

evaluate games more cognitively and competitively. They also seem to feel that they are more competent, whereas females seem to attach emotional values more easily to the games. The type of the display did not affect the appraisal of the game event. The display type affected only on adaptation, especially on three out of five presence dimensions. Figure 3. presents the independent effects of the display type to the PIFF² factors, after controlling for gender and the other background factors. Highest in Physical presence and Role

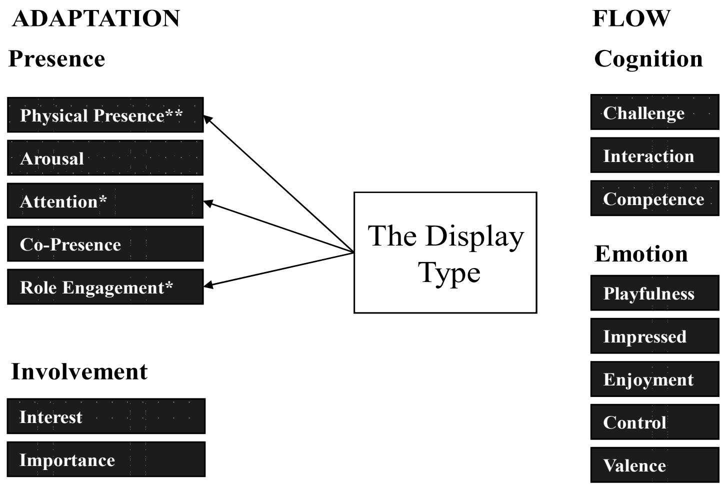


Fig. 3. Independent effects of on the display type on PIFF² factors. ** indicates $p<0.01$ and * indicates $p<0.05$

Engagement were stereo_1 and HMD conditions. HMD was also highest in Attention. These results indicate how display parameters shape our adaptation by affecting the sense of being in the game world. These changes in presence affect the extensity (e.g., voluminous) and intensity of the UX. Taken these two simple results together, we have demonstrated how to make UX in games measurable. With PIFF² analysis we have disclosed two factors that affect different parts of the experiential framework, and thus different experiential attributes. Linking PIFF² factors with game characters, such as mechanics and dynamics will tell game designers how a particular game feature and any change of it will contribute on the UX in games in any given user groups.

8. Conclusion

In the near future, the field of UX research will become one of the core fields of the psychological science. This will happen when the field adopts systematic and methodologically solid approaches. At the moment, the drivers for the interest in, e.g., game research are both technological and business related. The aim of this chapter is to introduce, how to approach UX in a psychologically founded way. We have presented general mental components that can be used to assess and evaluate any kinds of human experiences. Then we have formed an experiential cycle including higher level constructs of need & motivation, adaptation and appraisal to show how psychology can be used to evaluate UX in cellular phone and in digital games. Emphasis is on the understanding of what is measured and how. Experiential attributes are presented that can be regarded as measurement goals. More specific way to analyze gaming experience is then presented in our third framework, PIFF². The theoretical and methodological background of the PIFF² is explained and concrete example of it presented. In an empirical case study, it was showed that the display type had an effect on the adaptation, extensity and intensity of the UX, whereas gender had more effect on appraisal, quality and meaning of the experience.

Although gaming is getting more and more attention among the researchers, there are still only few attempts to study UX in games holistically. Although holistic approach is more demanding compare to focusing on one part of the UX, it enables evaluating complex mental relationships. Considering rich top-down processes in gaming will increase the understanding of the nature of today's games, development of better UX in new games, and future interactive interfaces that adapt users to whatever interactions. In addition, today's audio-visually rich and highly interactive games provide magnificent platforms to study human experiences. Increasing understanding about our mental processes will increase understanding of the H in HCI and eventually human life more generally.

9. References

Angell, J. R. (1907). The province of functional psychology. *Psychological Review,* 14, 2, 61-91, 0033-295X

Barr, P., Noble, J. & Biddle, R. (2007). Video game values: Human–computer interaction and games. *Interacting with Computers,* 19, 2, 180-195,

Breakwell, G. M. (2006). *Research Methods in Psychology,* Sage Publications

Brown, E. & Cairns, P. (2004). A grounded investigation of game immersion, *Proceedings of CHI2004,* pp. 1297-1300, ACM

Couper, M. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly,* 64, 4, 464-494, 0033-362X

Cronbach, L. J. (1990). *Essentials of Psychological Testing,* HarperCollins, New York

Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety,* Jossey-Bass Publishers, San Francisco

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience,* Harper & Row, New York

Csikszentmihalyi, M. & Csikszentmihalyi, I. S. (Ed). (1988). *Optimal experience: Psychological studies of flow in consciousness,* Cambridge University Press

Damasio, A. R. (1994). *Descarte's Error: Emotion, Reason, and the Human Brain,* Grosset/Putnam, New York, NY

Davidson, D. (2003). Interactivity in ico: Initial involvement, immersion, investment. *Proceedings of the Second International Conference on Entertainment Computing,* 1-21

Desurvire, H., Caplan, M. & Toth, J. A. (2004). Using heuristics to evaluate the playability of games, *Proceedings of CHI 2004,* pp. 1509-1512, ACM Press, New York, NY, USA

Dewey, J. (1934). *Art as Experience,* Capricorn Books, New York

Ellsworth, P. C. & Smith, C. A. (1988). From appraisal to emotion: Differences among unpleasant feelings. *Motivation and Emotion,* 12, 3, 271-302

Fabricatore, C., Nussbaum, M. & Rosas, R. (2002). Playability in action videogames: A qualitative design model. *Human-Computer Interaction,* 17, 4, 311-368

Finneran, C. M. & Zhang, P. (2002). The challenges of studying flow within a computer-mediated environment, *Proceedings of Eighth American Conference in Information Systems,* pp. 1047–1054, Dallas, TX

Forlizzi, J. & Ford, S. (2000). The building blocks of experience: An early framework for interaction designers, *Proceedings of in Proc. DIS-2000,* pp. 419-423, ACM Press New York, NY, USA

Frijda, N. H. (1987). *The Emotions,* Cambridge University Press, Cambridge

Ghani, J. A. & Deshpande, S. P. (1994). Task characteristics and the experience of optimal flow in human-computer interaction. *The Journal of Psychology,* 128, 4, 381-391

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences,* 20, 1-19

Hassenzahl, M. & Tractinsky, N. (2006). User experience–a research agenda. *Behaviour & Information Technology,* 25, 2, 91-97

Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences,* 16, 2, 107-117, 0022-5061

Hunicke, R., LeBlanc, M. & Zubek, R. (2004). MDA: A formal approach to game design and game research, *Proceedings of AAAI Workshop on Challenges in Game AI,* p. 4

James, W. (1890). *The Principles of Psychology,* H. Holt and company

Kahneman, D. (1973). *Attention and Effort,* Prentice-Hall, Englewood Cliffs, NJ

Labaw, P. J. (1980). *Advanced Questionnaire Design,* Abt Books

Lazarus, R. S. (1991a). Cognition and motivation in emotion. *The American Psychologist,* 46, 4, 352-367, 0003-066X

Lazarus, R. S. (1991b). Progress on a cognitive-motivational-relational theory of emotion. *The American Psychologist,* 46, 8, 819-834, 0003-066X

Lazzaro, N. (2004). *Why we play games: Four keys to more emotion without story.* Retrieved November, 2007, from http://www.xeodesign.com/whyweplaygames/xeodesign_whyweplaygames.pdf

Lessiter, J., Freeman, J., Keogh, E. & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence: Teleoperators & Virtual Environments,* 10, 3, 282-297, 10547460

Lombard, M. & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication,* 3, 2, 20

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science,* 4, 13, 333-369

Massimini, F. & Carli, M. (1988). The systematic assessement of flow in daily experience. In *Optimal experience: Psychological studies of flow in consciousness,* I. Csikszentmihalyi, & M. Csikszentmihalyi (Ed.), 266-287, Cambridge: Cambridge University Press

Mayer, J. D. (2001). Primary divisions of personality and their scientific contributions: From the trilogy-of-mind to the systems set. *Journal for the Theory of Social Behaviour,* 31, 4, 449-477

McMahan, A. (2003). Immersion, engagement and presence: A method for analyzing 3-D video games. In *The video game theory reader,* M. J. P. Wolf, & B. Perron (Ed.), 67-86, New York: Routledge

McQuarrie, E. F. & Munson, J. M. (1992). A revised product involvement inventory: Improved usability and validity. *Advances in Consumer Research,* 19, 108-115

Nakatsu, R., Rauterberg, M. & Vorderer, P. (2005). A new framework for entertainment computing: From passive to active experience, *Proceedings of ICEC 2005,* pp. 1–12, Kwansei Gakuin University, Kobe Sanda Campus, Sanda, Japan, IFIP

Novak, T. P., Hoffman, D. L. & Yung, Y. F. (2000). Measuring the customer experience in online environments: A structural modeling approach. *Marketing Science,* 19, 1, 22-42,

Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. & Fuller, T. (2003). User-centered design in games. In *Handbook for human-computer interaction in interactive systems,* J. Jacko, & A. Sears (Ed.), 883-906, Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Pinchbeck, D. (2005). Is presence a relevant or useful construct in designing game environments? *Proceedings of Computer Game Design and Technology,* Liverpool, UK. 8-9th November 2005

Pope, K. S. & Singer, J. L. (1978). *The Stream of Consciousness: Scientific Investigations into the Flow of Human Experience,* New York: Plenum

Rettie, R. (2001). An exploration of flow during internet use. *Internet Research: Electronic Networking Applications and Policy,* 11, 2, 103-113

Roseman, I. J. & Evdokas, A. (2004). Appraisals cause experienced emotions: Experimental evidence. *Cognition & Emotion,* 18, 1, 1-28

Rothschild, M. L. (1984). Perspectives on involvement: Current problems and future directions. *Advances in Consumer Research,* 11, 216-217

Rust, J. & Golombok, S. (1999). *Modern Psychometrics: The Science of Psychological Assessment,* Routledge

Ryan, R., Rigby, C. & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation & Emotion,* 30, 344-360, 01467239

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist,* 26, 3, 299–323

Schubert, T., Friedmann, F. & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments,* 10, 3, 266-281, 1054-7460

Sherry, J. L. (2004). Flow and media enjoyment. *Communication Theory,* 14, 4, 328-347

Sherry, J. L., Lucas, K., Greenberg, B. S. & Lachlan, K. (2006). Video game uses and gratifications as predictors of use and game preference. In *Playing video games: Motives, responses, and consequences,* P. Vorderer, & J. Bryant (Ed., 213-224, Mahawa, New Jersey: Lawrence Erlbaum Associates

Slater, M. (2004). How colorful was your day? why questionnaires cannot assess presence in virtual environments. *Presence: Teleoperators & Virtual Environments,* 13, 4, 484-493, 1054-7460

Sweetser, P. & Johnson, D. (2004). Player-centred game environments: Assessing player opinions, experiences and issues, *Proceedings of ICEC 2004,* pp. 321-332, Springer

Sweetser, P. & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *Computers in Entertainment,* 3, 3, Article 3a

Takatalo, J., Häkkinen, J., Särkelä, H., Komulainen, J. & Nyman, G. (2006b). Involvement and presence in digital gaming, *Proceedings of NordiCHI 2006,* pp. 393-396, ACM Press, Norway

Takatalo, J. (2002). Presence and flow in virtual environments: An explorative study. Master's thesis, University of Helsinki.

Takatalo, J., Häkkinen, J., Kaistinen, J., Komulainen, J., Särkelä, H. & Nyman, G. (2006c). Adaptation into a game: Involvement and presence in four different PC-games, *Proceedings of FuturePlay 2006,* London, Ontario, Canada, 10 - 12 October 2006

Takatalo, J., Häkkinen, J., Kaistinen, J. & Nyman, G. (2007). Measuring user experience in digital gaming: Theoretical and methodological issues, *Proceedings of IS&T/SPIE Symposium on Electronic Imaging,* pp. 1-13, San Jose, California, USA, IS&T/SPIE

Takatalo, J., Häkkinen, J., Komulainen, J., Särkelä, H. & Nyman, G. (2006a). The impact of the display type and content to a game adaptation, *Proceedings of MobileHCI 06,* pp. 17-20, Espoo, Finland, 12 - 15 September 2006, ACM Press New York, NY, USA

Takatalo, J., Häkkinen, J., Lehtonen, M., Komulainen, J., Kaistinen, J. & Nyman, G. (Submitted). Quality of the user experience in digital gaming

Takatalo, J., Häkkinen, J., Särkelä, H., Komulainen, J. & Nyman, G. (2004). The experiential dimensions of two different digital games, *Proceedings of PRESENCE 2004,* pp. 274-278, Valencia, Spain, October 12-15, UPV

Tarkkonen, L. & Vehkalahti, K. (2005). Measurement errors in multivariate measurement scales. *Journal of Multivariate Analysis,* 96, 1, 172-189

Webster, J. & Martocchio, J. J. (1992). Microcomputer playfulness: Development of a measure with workplace implications. *MIS Quarterly,* 16, 2, 201-226

Wundt, W. M. (C.H. Judd, trans), (1897). *Outlines of Psychology,* Retrieved April, 2002 from http://psychclassics.yorku.ca/Wundt/Outlines/

Yee, N. (2005). *A model of player motivations.* Retrieved January, 2008, from http://www.nickyee.com/daedalus

Zahorik, P. & Jenison, R. L. (1998). Presence as being-in-the-world. *Presence: Teleoperators & Virtual Environments,* 7, 1, 78-89, 1054-7460

Zaichkowsky, J. L. (1985). Measuring the involvement construct. *Journal of Consumer Research,* 12, 3, 341-52, 00935301

20

Iterative user interaction design for wearable and mobile solutions to assess cardiovascular chronic diseases

Elena Villalba Mora, Dario Salvi, Manuel Ottaviano, Ignacio Peinado, Maria Teresa Arredondo Waldmeyer
Life Supporting Technologies, Technical University of Madrid
Spain

1. Introduction

Cardiovascular diseases (CVD) are the leading source of death in western world, causing 45% of all deceases. Besides, heart failure, considered the paradigm of CVD, affects mainly people older than 65. The proportion of old people (aged 65 or over) in the European Union is predicted to rise from 16.4% in 2004 to 29.9% in 2050 (Eurostat, 2005). This will increase the number of elderly suffering from chronic diseases and state a significant strain on personal heath care services with new technologies, such as wearable and mobile systems (Gatzoulis & Iakovidis, 2007) (Lamberis and Dittmar, 2007).
Ambient Intelligence (AmI) vision implies the creation of intelligent environments where users interact with the surrounding naturally without additional effort. The technology is integrated into the user's daily life. The environment is adapted to the users in a pro-active context (ISTAG; 2001).
This new social and technological paradigm (i.e. AmI) claims for a new way of designing user interaction and systems which enables services to be ubiquitous and adaptable to the user's particularities. Although personalization offers the possibility to adapt the system execution to the users' preferences, this new interaction model needs adaptation in real time to the user context.
The research reported in this book chapter is based on the iterative design, development and validation of a new model to improve the quality of life of people who live with chronic diseases following Ambient Intelligence principles. The model is validated in a solution to assess heart failure patients remotely. The interaction is natural and the system adapts to the patient dynamically by means of complex intelligent mechanisms. The chapter focuses on the patients' interaction.
The methodology used is ad-hoc adapted following the design principles of User Centred Design (ISO, 1999) and Goal Directed Design (Cooper, 2007) together with an Iterative Software Design (ISD) and Agile Modelling (Sotirovski, 2001) (Ambler, 2002).
The methodology has 3 iterative phases (modelling, implementing and deploying phases) which focus on observing and interviewing stakeholders such us medical experts and

patients in all stages of the global process. More than 80 people participated in an intensive validation along the three phases.
The proposed model of the patient interaction is applied to a use case: a solution to assess remotely heart failure based on daily monitoring of body signals and vital signs, both with wearable and mobile technologies.
This solution, called Heart Failure Management (HFM), makes use of the latest technologies for monitoring heart condition, with wearable garments (for measuring ECG, and respiration); and portable devices (such as weight scale and blood pressure cuff) with Bluetooth capabilities (see Fig. 1).
HFM aims to decrease the mortality and morbidity of the HF population. The system intends to improve the efficiency of the healthcare resources, maximizing the cost-benefit rate of the heart failure management.
The main users of the system are two: a) an HF chronic disease management service provider, with cardiologists and nurses; and b) patients with HF.

Fig. 1. Heart Failure Management solution.

The system consists of a user interaction platform and a professional interaction platform. The sensor used are a blood pressure, a weight scale, bed garments to monitor during night

and wearable garment such a vest or bra to monitor electrocardiogram, respiration and activity during exercising and resting daily.
The user platform groups all sensors and a personal digital assistant device (PDA) which receives data from the monitoring devices, processes it and encourages the patients in their daily healthcare. The professional platform includes the processing server to analyze all data, databases and a portal which provides ubiquitous access of the professionals.
The daily routine data are processed and evaluated for the detection of functional capacity, heart failure worsening and other complications. Motivation strategies must be taken into account in order to provide patients with relevant information, according to their physical and psychological status.

2. Background

In 1999, the IST Advisory Group (ISTAG, 2001) described the vision on Ambient Intelligent as the orientation for the workprogramme of 2000 and incoming years. ISTAG agreed on a single guiding vision wherein the citizen's everyday surroundings became the interface.
Gillian Crampton Smith, from the Interaction Design Institute Ivrea said, "In the same way that industrial designers have shaped our everyday life through objects that they design for our offices and for our homes, interaction design is shaping our life with interactive technologies – computers, telecommunications, mobile phones, and so on. If we were to sum up interaction design in a sentence, I would say that it's about shaping our everyday life through digital artefacts – for work, for play, and for entertainment" (Moggridge, 2007). This way she states what is interaction design. Besides, Michael Schrage, from MIT Media Lab, stated "Innovators don't change the world. The users of their innovations do".
These statements converge in one vision: the necessity of putting the customer, the patient, the person, the user in the centre of research and innovation. The person of the future is surrounded by advanced computing and networking technology that is aware of his presence, his personality, his needs and response intelligently.

2.1 HCI models

People have interacted with computers from the start, but it took time for human–computer interaction (HCI) to become a recognized field of research. Related journals, conferences, and professional associations appeared in the 1970s and 1980s. HCI is in the curricula of research universities, primarily in computer science, yet it has not coalesced into a single discipline. Fields with researchers who identify with HCI include human factors and ergonomics, information systems, cognitive science, information science, organizational psychology, industrial engineering, and computer engineering. (Grudin, 2005)
Human-computer interaction studies the interactions and the relationships between humans and computers. HCI is more than user interfaces; it is a multidisciplinary field covering many areas (Hewett, 1996). In the first ten to fifteen years of its history, HCI has focused on interfaces (particularly on the possibilities and design criteria for graphical user interfaces (GUIs) using windows, icons, menus, and pointing devices to create more usable systems. As interface problems were better understood, the primary HCI concerns started to shift beyond the interface (to respond to observations as articulated by D. Engelbart: "If ease of use was the only valid criterion, people would stick to tricycles and never try bicycles"). More recent HCI research objectives are concerned with tasks, with shared understanding,

and with explanations, justifications, and argumentation about actions, and not just with interfaces. The new essential challenges are improving the way people use computers to work, think, communicate, learn, critique, explain, argue, debate, observe, decide, calculate, simulate, and design. (Fischer, 2001)

Jonathan Grudin (Grudin, 2005) identifies three faces of human-computer interaction. First one extended human factors or engineering psychology to computing. Another developed when mainframes spawned business computing in the 1960s. The third, focused on individual use, arose with minicomputers and home computers and burgeoned with personal computing in the 1980s.

The first thread is the Human Factors and Ergonomics (HF&E). The second focuses on information systems (IS) management. The third one arose in the 1980s with personal computing, it is known as computer-human interaction (CHI) within this chapter.

Although they share some issues and methods, these research efforts have not converged. They emerged within different parent disciplines, at different times, and comprised different generations of researchers. Approaches, attitudes, and terminology differed. Two—computer operation and information systems (IS) management— embraced the journal-oriented scholarly tradition of the sciences; the third—comprising cognitive and computer scientists— has placed greater emphasis on conference publication. In addition, each thread initially emphasized a different aspect of computer use: mandatory hands-on use, hands-off managerial use, and discretionary hands-on use. Designing for a use that is a job requirement and designing for a use is discretionary can be very different activities.

Traditionally, computer usage was modelled as a human-computer dyad (see Fig. 1) in which the two were connected by a narrow explicit communication channel, such as text-based terminals in a time-sharing environment. The advent of more sophisticated interface techniques, such as windows, menus, pointing devices, colour, sound, and touch-screens have widened this explicit communication channel. In addition to exploring the possibilities of new design possibilities for the explicit communication channel, knowledge-based architectures for HCI have explored the possibility of an implicit communication channel (see Fig. 2). The implicit communication channel supports communication processes which require the computer to be provided with a considerable body of knowledge about problem domains, about communication processes, and about the agents involved.

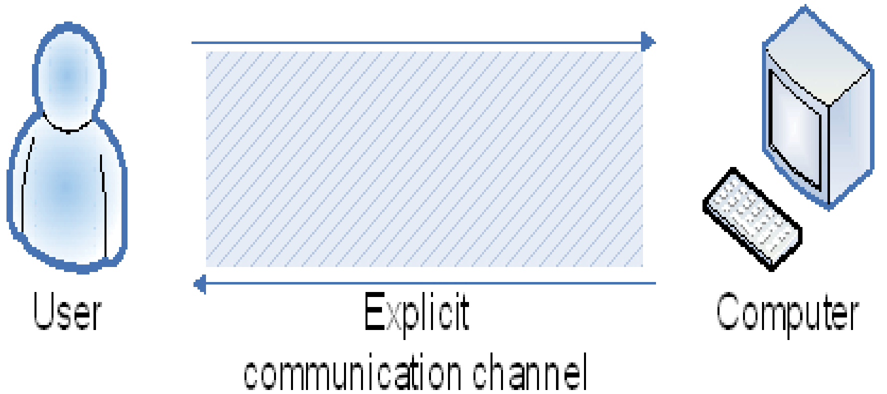


Fig. 2. Human-computer interaction dyad (From Fischer, 2001)

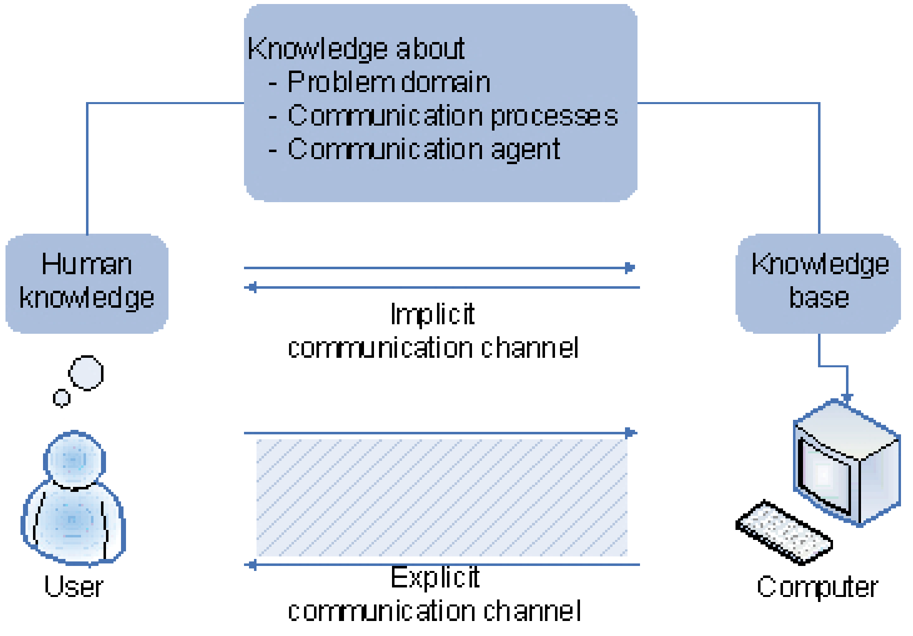


Fig. 3. Knowledge-based HCI (From Fischer, 2001)

As stated in Fig. 3 the knowledge comprises three different types of knowledge (Fischer, 2001):

- Knowledge about the problem domain: Shared knowledge builds upon large amounts of knowledge about specific domains. This knowledge constrains the number of possible actions and describes reasonable goals and operations in the domain of specific users, thereby supporting human problem-domain interaction and not just human-computer interaction.
- Knowledge about communication processes: The information structures that control communication should be accessible and changeable by the user. A knowledge-based HCI system should have knowledge about when and whether to assist the user, interrupt the user and volunteer information to the user contextualized to the task at hand.
- Knowledge about the communication agent: The "typical" user of a system does not exist; there are many different kinds of users, and the requirements of an individual user usually change with experience. Simple classification schemes based on stereotypes such as novice, intermediate, or expert users, are inadequate for complex knowledge-based systems because these attributes become dependent on a particular context rather than applying to users globally. One of the central objectives of user modelling in HCI is to address the problem that systems will be unable to interact with users cooperatively unless they have some means of finding out what the user really knows and does. Techniques to achieve this include: (1) being told by the users (e.g. by questionnaires, setting preferences, or specification components); (2) being able to infer it from the user's actions or usage data; and (3) communicating information about external events to the system.

There are several HCI models in which the interaction occurs always in an explicit way, that is to say, the user directly enter the data into the system and receives feedback. See Schomaker, 1995 for a basic HCI model.
Explicit interaction requires always a kind of dialog between the user and a particular system or computer the user is currently interacting with. This dialog brings the computer inevitably to the centre of the activity and the users focus is on the interface or on the interaction activity. This form of interaction is obviously in contrast to the visions of calm and Ubiquitous Computing. Also the idea of a disappearing computer [9] and ambient intelligence is hard to imagine with explicit interaction only. The realization of these visions can only be achieved when parts of the interaction between the computer and the human are transparent and not explicit, as stated above (Schmidt, 2005).
There are many things that influence the interaction between humans that are not contained in traditional "human computer interaction". The influence of situation, context, and environment offers a key to new ways of HCI. To come closer to the aim of creating interaction between humans and systems that is closer to natural interaction it becomes crucial to included implicit elements into the communication in addition to the explicit dialog that already used.
The following definition (see Table 1) characterizes the new paradigm of implicit human computer interaction (iHCI).

Implicit Human-Computer Interaction (iHCI)	iHCI is the interaction of a human with the environment and with artifacts which is aimed to accomplish a goal. Within this process the system acquires implicit input from the user and may present implicit output to the user.
Implicit Input	Implicit input are actions and behavior of humans, which are done to achieve a goal and are not primarily regarded as interaction with a computer, but captured, recognized and interpret by a computer system as input.
Implicit Output	Output of a computer that is not directly related to an explicit input and which is seamlessly integrated with the environment and the task of the user.

Table 1. Implicit Human Computer Interaction definitions (Schmidt, 2005)

The basic idea of implicit input is that the system can perceive the users interaction with the physical environment and also the overall situation in which an action takes place. Based on the perception the system can anticipate the goals of the user to some extent and hence it may become possible to provide better support for the task the user is doing. The basic claim is that Implicit Human Computer Interaction (iHCI) allows transparent usage of computer systems. This enables the user to concentrate on the task and allows centring the interaction in the physical environment rather than with the computer system.

To support the creation of systems that use implicit interaction it is important to provide a simple model that reflects this interaction paradigm. In Fig. 4 an abstract model of implicit interaction is shown.

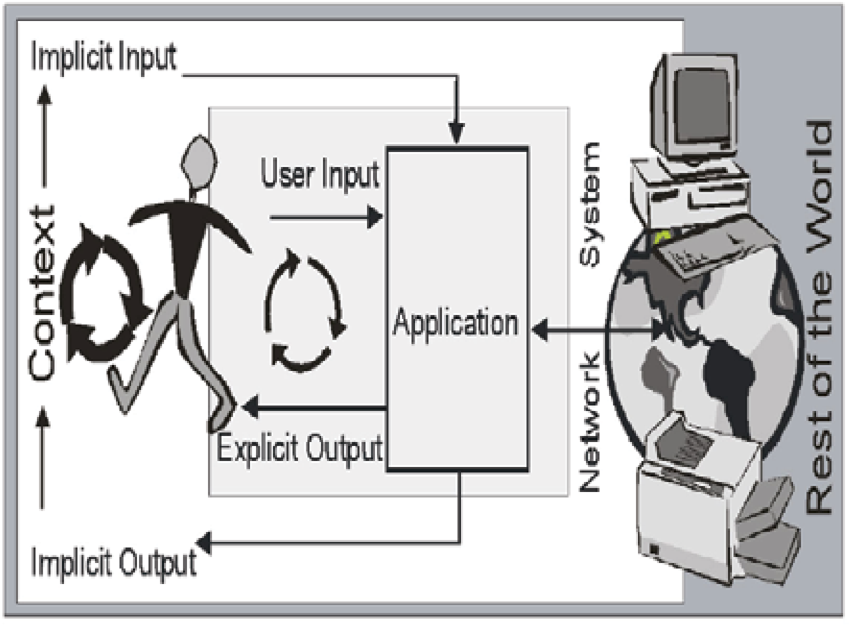


Fig. 4. Implicit Human Interaction Model (Schmidt, 2005)

All actions carried out by a human are taking place in context – in a certain situation. Usually interaction with our immediate environment is very intense (e.g. sitting on a chair, feet on the ground, garment on the body, moving books on the table, drinking from a glass, etc.) even if we don't recognized it to a great extent.

All contexts and situations are embedded in the world, but the perception of the world is dictated by the immediate context someone is in. Explicit user interaction with an application is embedded into the context of the user and is also a way of extending the context of the user, e.g. by having access to the network.

Applications that make use of iHCI take the context into account as implicit input and also have an influence on the environment by implicit output. The proposed model is centred on the standard model in HCI where the user is engaged with an application by a recurrent process of input and output. In the iHCI model the user's centre of attention is the context – the physical environment where the task is performed. The interaction with the physical environment is also used to acquire implicit input. The environment of the user can be changed and influenced by the iHCI application.

The system and also the network are to some extent part of the context but are also accessible by the application directly.

The Holistic Patient Interaction Model is based on the iHCI. Following sections explains the methodology and all models that form the hPIM.

3. Methodology

Since there were no previous documented work on this topic, the methodology used needed to be really user-centered. Besides, stakeholders were involved in all stages, from the early conceptualization till the final validation.

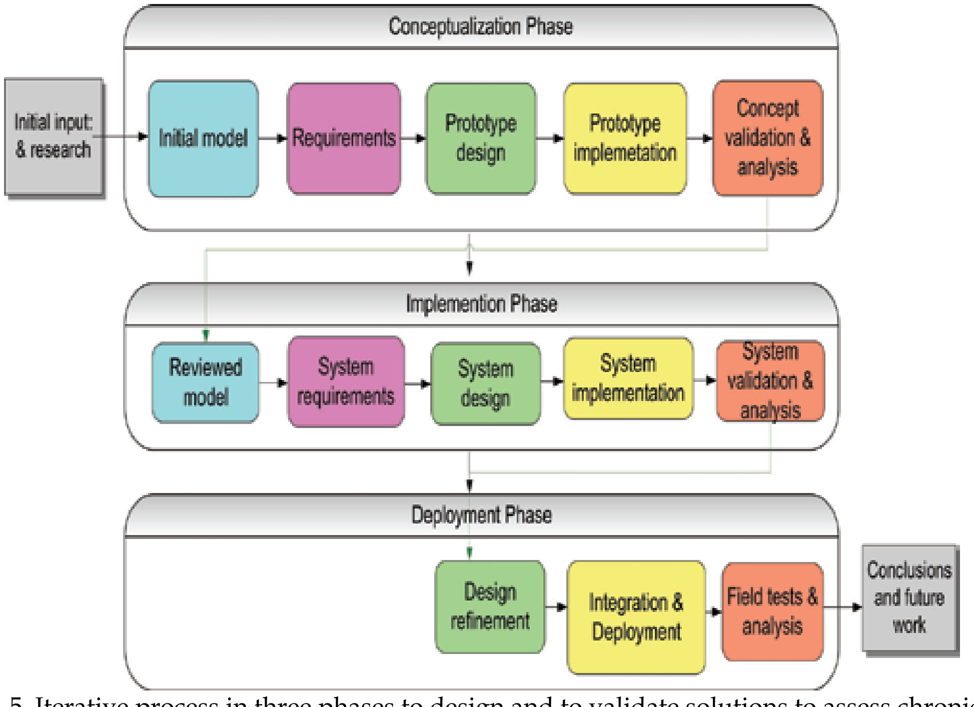


Fig. 5. Iterative process in three phases to design and to validate solutions to assess chronic conditions in AmI.

The whole process (see Fig. 5) is divided into three iterative phases: Conceptualization Phase, Implementation Phase and Deployment Phase.

Each phase follows the design principles of User Centered Design and Goal Directed Design (GOD) (Cooper, 2007). This last design is divided into Research, Modeling, Requirements, Framework and Design and is based on ethnographic techniques to people research.

The requirements, the definition of persona, scenarios and key paths are based on GOD techniques.

4. Holistic Patient Interaction Model

To assure the success of solutions to self-manage the healthcare by chronic patients at their own homes, we need to understand all actors and interactions that occur in a holistic framework. The Holistic Interaction Model comprises three contexts that round the patient. Fig. 6 illustrates the conceptual model.

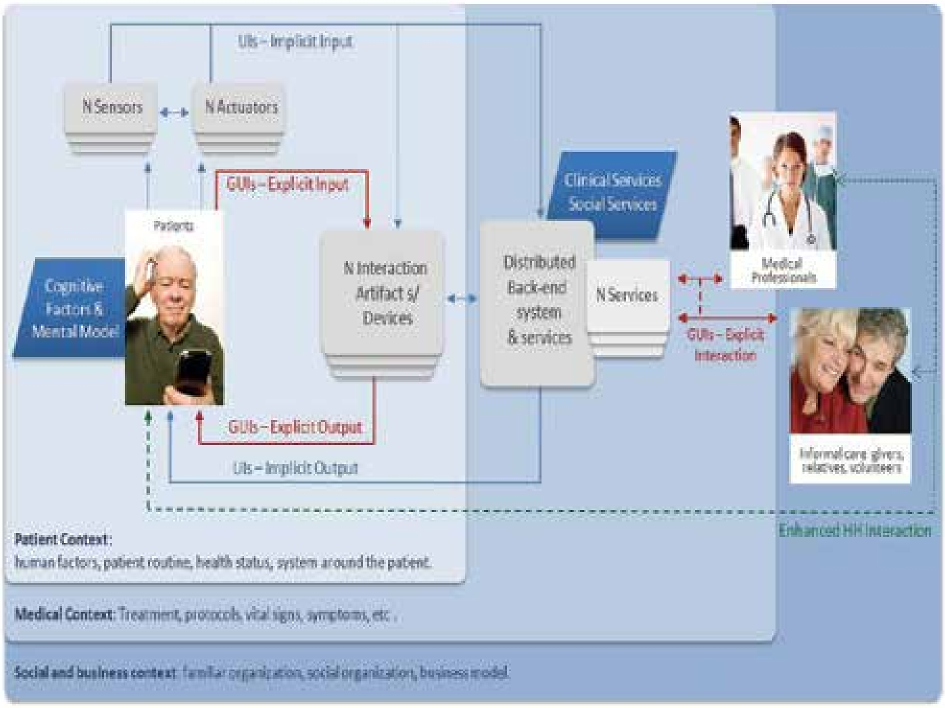


Fig. 6. Conceptual view of the Holistic Patient Interaction Model

In a generic scope, the interaction is modelled as several contexts and several interaction loops. The holistic Patient Interaction Model is organized in three contexts, the Patient Context which defines the human factors or the patient personal routine. Sensors around the patient play important role in their interaction since allows implicit interaction without specific patient input. Around this first context, the Medical Context comprises also the services which provide the patients with a remote monitoring assessment. This context groups all medical professionals. Social and Business Context appears around all. This context states the social and clinical rules that must be taken into account. With this holistic approach all actors are studied enhancing also the human-human interaction.

Adaptation to personal routines is the most important user requirement. Specifically, each user will have a different daily health schedule according to particular health status, preferences, mental status and recommended medical protocol.

Adaptability to user preferences and routines is achieved via dynamic workflow execution (which depends on the context information). First, we defined taxonomy: a session is a day using the system; a day is divided in contexts (e.g. morning, exercise, evening, and night). Each context comprises a set of activities requiring user participation at the same temporary term (i.e. a task or activity is the measurement of blood pressure). It is done while measuring the subject's weight and the morning questionnaire. Thus, we have the morning context. Each context is defined by a beginning and ending time that sets the validity period, restricts the context execution, sets activity performances, and describes the current context state. The state can be inactive, active, performed, aborted and incomplete. Following this scheme, after the application is turned off, all contexts are inactive. Fig. 7 shows the flow interaction model based on contexts.

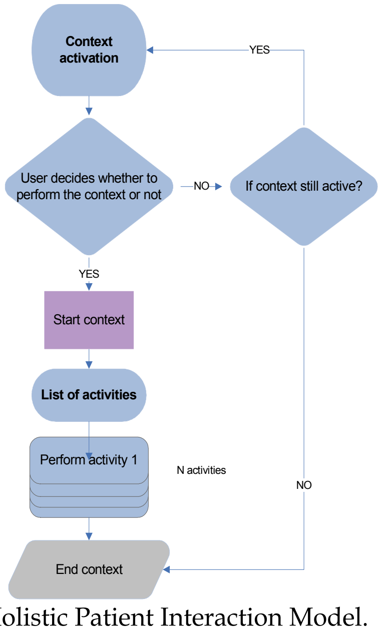


Fig. 7. Dynamic flow of the Holistic Patient Interaction Model.

Different rules govern context activation (i.e. they are performed between starting and ending time). They also have restrictions in the medical protocol (i.e. exercise must be finished two hours after medication).

Following context activation, the alarm is launched and executed with the user acknowledgement. The context execution equates the activity execution in the list. Each manager controls each activity, thus allowing modularity. The manager invokes all necessary SW and HW modules (e.g. sensors).

Restrictions are events-based occurring in real time. They are triggered either by the user via direct interaction (user voluntarily stops an activity) or by the system. The system launches events coupled to the health status from the study of vital information or the environmental status (patient location, etc.).

The activities execution dynamically occurs depending on the current context: period status, user dialogue, and the health status together with other contextual information.

All gathered data, raw processed raw signals, notifications are sent to the Back-end for further processing and management.

The professionals access all data via the portal. The first available information is an outline of every user highlighting crucial events. The professionals can also consult and edit the information for specific uses. Thus, professionals close the loop and enhance their relationship and interaction with patients.

In order to complete the Holistic Interaction Model, Personas and scenarios are defined following GOD principles. To conclude, the interaction is modelled as key-paths and variations of key-paths or workflows. They represent all interaction flows that occur in a real situation based on context activation (Cooper, 2007).

4.1 Persona for Heart Failure Management

The generic user ("persona") of HFM is Carlos Gómez, 72 years old. He is retired and has heart failure. His awareness of his heart condition leads him to be proactive in his health. He can use an electronic device following an intuitive system. He requires no special needs regarding accessibility (e.g. blind people).

His chief goals are self-assurance and self-confidence when performing his daily routine. He must feel unperturbed and lose his fear of a sudden death. He also aims to control his own health evolution by self-managing his health.

He wishes to live normally, thus making it crucial to give him a system that is non-intrusive that invisible from public view while under treatment. Namely, the system must adapt to his daily routine.

4.2 Scenarios and user requirements within Heart Failure Management

The team creates stories about ideal user experiences, describes how the system fits into the personas' life and their environment, and helps them achieve their goals. We based the method on the description of scenarios or contexts of daily use.

In the HFM context, the end users are prompted to follow a daily routine consisting of a set of activities (i.e. symptoms questionnaires, measurements using wearable garments and portable devices). The vital signs assessed are ECG, heart rate, and respiration. The portable devices are a blood pressure cuff for systolic and diastolic blood pressure and weight scale. All devices and garments have communication capability (i.e. Bluetooth). Moreover, the user can perform a light exercise of 5-6 minutes, several days a week to improve their health. This routine varies for every patient but must follow some rules for medical reason (e.g. blood pressure must be taken every morning). The routine can be personalized for each patient despite the light constraints.

There were two scenarios detected within the system: indoors and outdoors. The former contains a set of measurements, using the wearable garments and portable devices at home. The user answers two questionnaires defined by the medical team. The later contains an exercise scenario (e.g. a short walk) that promotes a healthy lifestyle and improves cardiovascular capability. The professional checks the status of all patients via portal.

Adaptation to personal routines is the most important user requirement. Specifically, each user will have a different daily health schedule according to particular health status, preferences, mental status and recommended medical protocol. Furthermore, the user application must be intuitive, user-friendly, and must allow natural interaction. The PDA with a touch-screen allows these requirements.

4.3. Graphical user interface as interaction driven within Heart Failure Management

According to the results of the interviews during the validation phase in Eindhoven, using metaphors such as arrows to represent backwards and forward or "X" to skip was generally confusing. The users suggested that buttons had to be button-shaped and labeled with the action that would take place when pushed. It was decided to implement a color code that makes use of the average user mental model.

Table 2. describes the aim of each functional block in the graphical user interface.

Header area	The header area includes information about the current context and activity, the application logo and a "Skip" button that will let the users exit the fore mentioned activity.
Information Area	The most important portion of the screen, the "Information Area" supports the interaction through messages and instructions.
Navigation area	The navigation area comprises the navigation buttons that will help the user navigate through the screens.

Table 2. Functional groups of the GUI

Table 3. describes the three types of buttons that are used.

Green buttons	The green buttons are associated to the key path.
Gray buttons	Gray buttons represent variants of the key path.
Red buttons	Red buttons are intended for exceptional situations, such as stopping the application.

Table 3. Buttons' types of the graphical user interface

With all this design, the resulted GUI was text based, with no extra pictures or methaphores and based on haptical input (see Fig. 8).

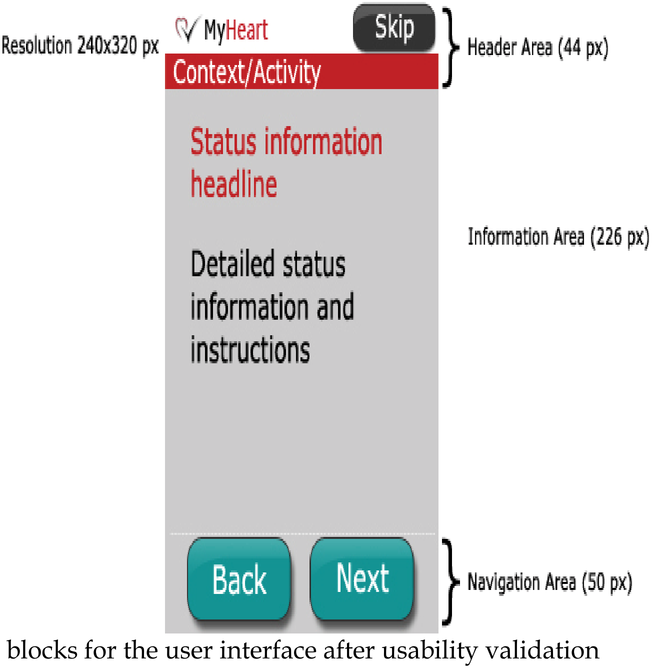


Fig. 8. Functional blocks for the user interface after usability validation

To finalize with the modelling, the patient interaction flow is described in a set of screen and the transition among them in a key path and complete paths (Cooper, 2007). See Fig. 9.

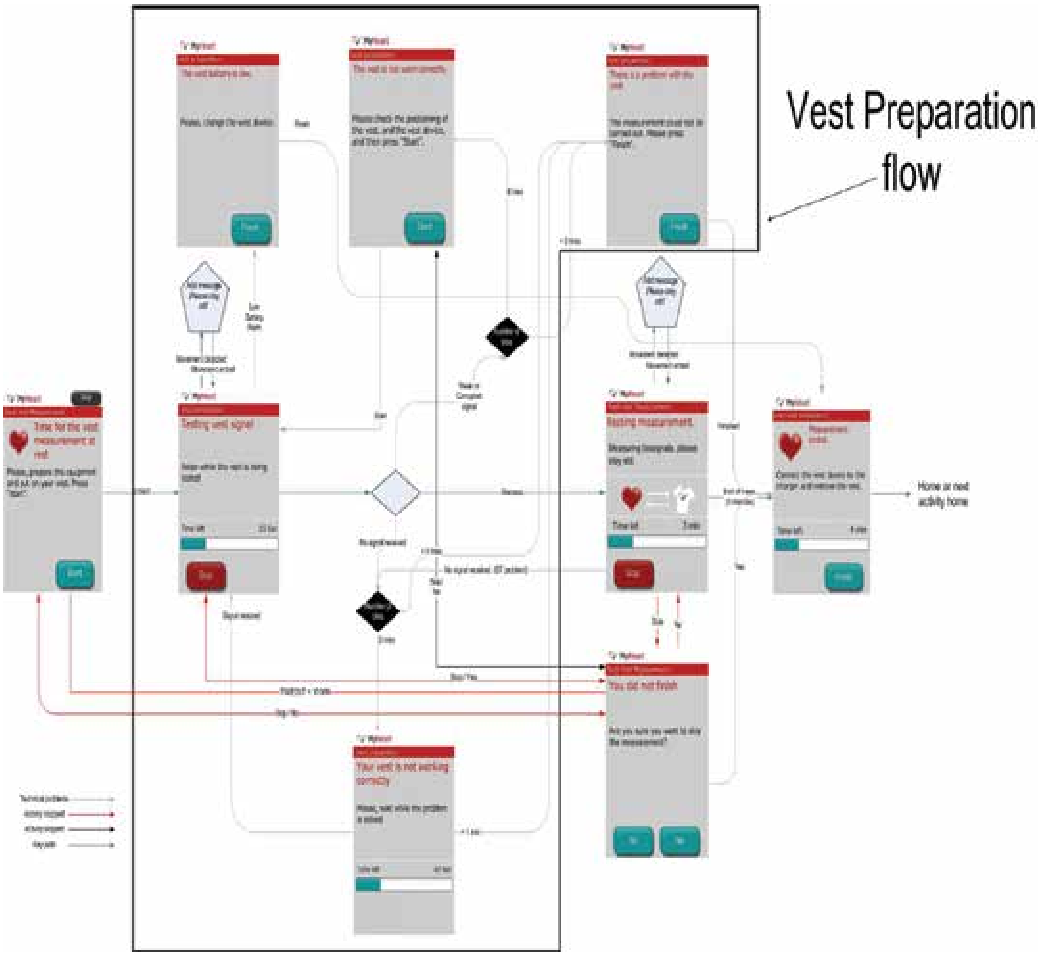


Fig. 9. Complete path to describe the user interaction in the vest activity which permits the patient to measure the heart condition thought the wearable garment implemented in a vest.

5. Resulted patient interaction system

Taking as input all information from the models, we implemented a solution based on a context workflow engine.

The resulted distributed system is composed of a patient interaction which comprises all necessary sensors and interaction device (i.e. a PDA) and the Back-end with all servers, databases and a portal to allow the professional interaction.

The final user platform is a modular architecture running on Microsoft's .NET Framework (Rubin & Yates, 2003). The Session Context Engine is the core element that provides flexibility in the protocol that users follow. We based it on the workflow engine that invokes the tasks the user performs according to a workflow that varies with every patient. The varied tasks are organized in contexts.

Fig.10 illustrates the user platform's architecture whose modules are further explained below.

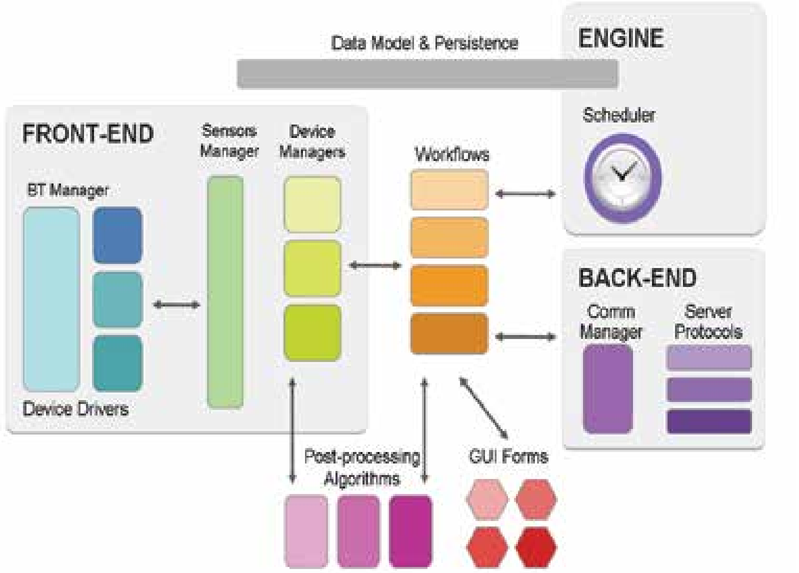


Fig. 10. Modular architecture of the user platform

The architecture is divided into 4 main areas: a) the Front-end module, b) the Workflows, c) the Engine and d) the Back-end module. The PDA is also composed of Post-Processing Algorithms, a set of programming Tools and GUI Forms that allow user interaction.
The Front End module provides a standard interface to the sensors. This module isolates the complexity of the communication protocols towards the different sensors. The sensors are a blood pressure monitor, a weight scale and wearable on body sensors to take information from electrocardiogram, respiration and activity.
Workflows define the user interaction via displayed messages and actions taken when the user presses the button. Events from the scheduler trigger the workflows when a context requires performance; the related workflow starts showing the correct user interface. Activity workflow communicates with the device manager and receives the measured data plus events relating to activity status during the activity's execution. The workflows use the graphical user interface via opening and closing formularies, and setting their text fields. The workflows use a Localization module that gathers the local language's textual information from an .xml file to provide internationalization into the application.
The engine module covers a) the application data meaning and storage, b) a scheduler, c) some modules for the application configuration and d) an error management module. This engine implements and manages the Session/Context/Activity model. The application runtime information is stored on Tabloid, a data structure. A tabloid implements the table containing the list of actions the application performed and will perform during the day.
A scheduler persistently updates the tabloid. The scheduler checks the system clock and decides the actions the user performs. It manages contexts, sends information, and launches events warning the user of pending action. Fig. 11 sketches the scheduler.
Thus, two programming threads work within the scheduler: the scheduler thread updates tabloids and main forms thread handles the events that activate the GUI.

Context execution is managed on the tabloid. Each context is defined via starting and ending time that sets the validity period, a set of restrictions constraining the context situation, a set of activities the user will be performing, and a variable describing the current context state.

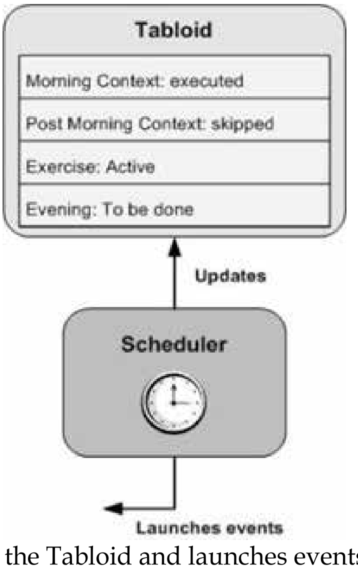


Fig. 11. The Scheduler updates the Tabloid and launches events

Some possible states are described in Table 4:

State	Description
INACTIVE	when a context has not begun yet
ACTIVE	when a context can be performed and is waiting for an event to be started
PERFORMED	when a context has been completed correctly
ABORTED	if the context execution has been aborted by the user
UNCOMPLETE	when the context has been executed, but not finished

Table 4. Possible states of a context.

Following this scheme, when the users' application sets off, all contexts are inactive. Context activation follows the following rules:

1. Each context can be performed between its starting and an ending time.

2. If a context has no restrictions, only the time is checked to know when to activate the context.
3. If the context has restrictions, all of them must be matched.

An alarm is launched and executed via user acknowledgement once a context is activated. The context execution means the execution of activities included in the list. Activity managers control each activity: a questionnaire activity manager supervises the questionnaire's extraction from a configuration file, fills the questionnaire data structure and manages the answers.

Graphical user interaction is provided as .Net formularies (forms). Each form represents an application screen, with static, graphic elements (text/pictures) and dynamic and interactive elements like buttons and slides. User interaction and workflows manage the forms via events.

6. Validation along the process

According to the philosophy of participatory design and involvement of users in all stages of the process life cycle, the validation was performed along the three phases (ISTAG; 2004). Table 5 summarizes all evaluations and testing.

Process phase	Validation	Place	N patients	N non-patients
Conceptualization	Conceptual validation	Spain	10	16
Implementation	Experts – Heuristics evaluation	Eindhoven, The Netherlands		4
	Usability tests with final users	Eindhoven, The Netherlands	5	
	Field tests with final users	Basel, Switzerland	4	
Deployment	Field tests with final patients	Madrid, Spain	37 (31 + 6)	
	Field tests with final non- patients users	Aachen, Germany		6
TOTAL			82	

Table 5. Iterative validation along the whole process

The conceptual validation consisted of a set of personal interviews to respondents of identified stakeholders. These interviews lasted between one hour and two hours and had several sections. This evaluation was done during first half of 2005 (Villalba, 2007).
During the Implementation Phase, firstly a Heuristic evaluation (Nielsen, 1994) with experts was performed in October 2006. Direct afterwards, the interfaces were confronted to real old patients, in total 5. Both testing were done in Eindhoven, The Netherlands.
The interaction was then revised and re-implemented in the next prototype. The new system was validated with 4 patients in a field test, where the patients used the solution at their own homes during one week. This validation occurred in July 2007 in Basel, Switzerland.
We held the field test during the implementation stage with Medgate in Basel throughout July 2007 with 4 old ladies. The study’s primary objective was to validate the global system for Heart Failure’s assessment in a real environment with real users who will use the system in their homes. We tested version 1 of the system (Villalba, 2007).
Patients performed all tasks until technical problems occurred. Only one patient, whose PDA malfunctioned from the third day, didn’t complain.
We started the field text at home and the day they returned all devices. We prompted all patients to fill in a user-experience questionnaire.
The patients initially felt intimidated by the device. After the pilot, they felt more positive although they were still unsure how it will fit into their daily routines.
In the Deployment Phase, the system was again validated with patients and non-patients users. This last validation phase took place in Aachen, Germany and Madrid, Spain from January to March 2008.
In total 37 people were interviewed, 31 at the Hospital Clínico de Madrid (HCM) and 6 in Getafe, Madrid at patients’ homes. All of them were cardiovascular patients. People in Getafe were post-event patients, both angina and myocardial infarction. Patients from HCM were heart failure patients. Tests at HCM were performed under the supervision of technical people, user interaction experts and medical professionals. The test occurred during January 2008.
Fig. 12 shows the distribution of the sample of population involved in the test of Getafe, Madrid (left) and HCM, Madrid (right). The mean age is 60,3 years and the standard deviation is 5,75. The average age at HCM is 67,13 years and the standard deviation is 12,32.

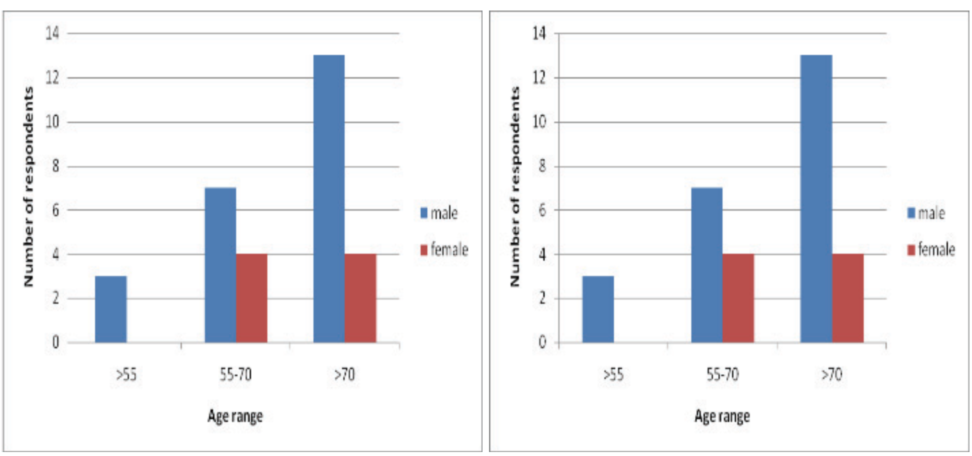
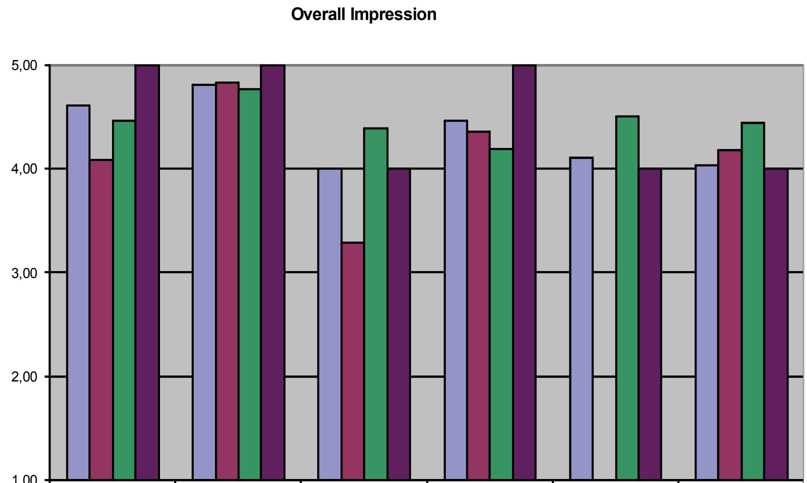


Fig. 12. Population sample of the field test in Getafe (left) and Madrid (right)

The interviews lasted 60 minutes in which the patient had to use the system for 15 minutes and to fill in a final questionnaire to extract objective data.
The results of this test, specially the 31 patients had very positive results. All usability and acceptability problems seem to be solved after all iterations with stakeholders and re-design (see Fig 13 and Fig. 14).



	Interesting	Not scary	Pretty	Fun	Simple	Easy
General concept	4,61	4,81	4,00	4,46	4,11	4,04
Vest	4,08	4,83	3,29	4,36		4,18
PDA	4,46	4,77	4,38	4,19	4,50	4,44
BP	5,00	5,00	4,00	5,00	4,00	4,00

Fig. 13. Overall impression of the testing at HCM.

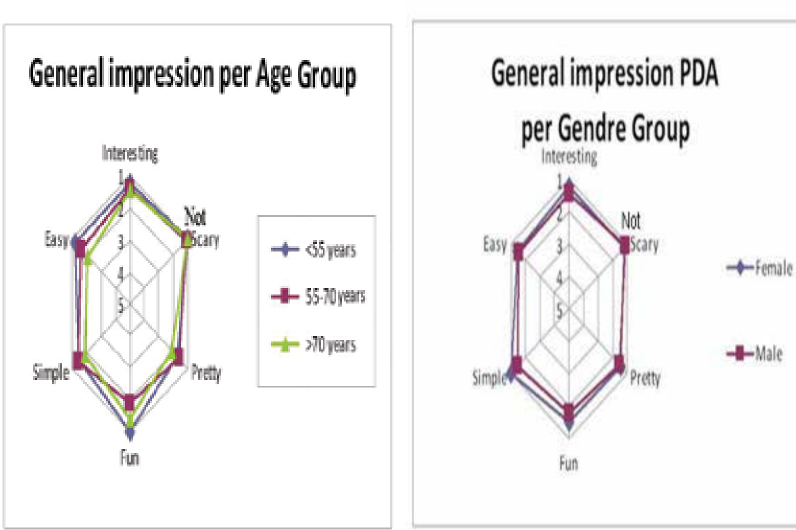


Fig. 14. Representation of the general impression of the overall system and the PDA in a spider graph. We can notice that the results are pretty positive.

Regarding the ease of use, there were some concerns about the use of the wearable garment or vest. However, in general all results are over 3,70 over 5. That is to say, the resulted final solution is suitable to be use by heart failure patients.
Each testing iteration shows a better understanding and better acceptance of the use of the technology by heart failure patients.

7. Conclusion

The research presented within this book chapter has modelled the patient interaction in an Ambient Intelligence environment. This interaction includes all actors, stakeholders and distributed systems that are needed to assess chronic conditions. In the current point of the way, the success of these systems relies mainly in the user acceptance. This acceptance will occur only when the technology, the systems and the interaction are design for and with the real patients and other stakeholders.
The current work was focused on the improvement of the usability, minimizing the interaction requirements and increasing the contextual awareness.
Development of user-friendly interfaces will help users get familiar with the latest technology available in the market. Ultimately the industry will benefit from this as well, as it will boost purchase and usage behaviour. Development of user-friendly systems requires knowledge about factors that influence the user's learning behaviour while handling new systems.
The results are very promising in terms of the interaction modality implemented. A detailed analysis to enhance individuals experience incorporating this system into the daily routine is still lacking. For this reason, an "in depth" study regarding behaviour components towards e-health to create a tailored communication framework, as well as to increase the patient's interest to use such systems is being performed. A framework to be followed considers the analysis of different variables.
Finally, these systems will play an important role on the daily life of chronic patients, supporting a better quality of life and helping to prevent and to treat chronic conditions in a more effective way.
Moreover, there is also the need of involving public authorities in this new technological and social wave. Without their collaboration in the global policies, the success of theses systems will be drastically reduced.

8. References

Ambler S. W. Jeffries R (2002). *Agile Modeling: Effective Practices for Extreme Programming and the Unified Process*. John Wiley & Sons. 1st edition (March 22, 2002).
C. Zwick, B. Schmitz, K. Kuehl, (2005) *Designing for Small Screens*. Edited by AVA publishing SA, 2005.
Cooper A, (2007) *About Face 3.0 The Essentials of Interaction Design*. Edited by Wiley Publishing, Inc. 2007.
Eurostat (2005) *Population projections 2004-2050*.
Fischer, G. (2001). User Modeling in Human-Computer Interaction. *User Modeling and User-Adapted Interaction*, 11, 65-86. Kluwer Academic Publishers. The Netherlands.

Gatzoulis, L. & Iakovidis, I. Wearable and Portable eHealth Systems. *Engineering in Medicine and Biology Magazine, IEEE* Volume 26, Issue 5, Sept.-Oct. 2007 Page(s):51 – 56.

Grudin J (2005). Three Faces of Human-Computer Interaction. *IEEE Annals of The History of Computing*. 2005. Pages 46-62.

Hewett, Baecker, Card, Carey, Gasen, Mantei, Perlman, Strong and Verplank (1996). Chapter 2: *Human-Computer Interaction*. ACM SIGCHI Curricula for Human-Computer Interaction. Association for Computing Machinery, Inc.

ISO 13407. (1999) *Human-Centred Design Processes for Interactive Systems*, Geneva, (Switzerland).

ISTAG (2001) *Scenarios for Ambient Intelligence in 2010*. Edited by European Communities, 2001.

ISTAG (2004) *Report on Experience & Application Research*. Involving users in the Development of Ambient Intelligent. Edited by European Communities, 2004.

Lamberis, A. & Dittmar, A. Advanced Wearable Health Systems and Applications - Research and Development Efforts in the European Union. *Engineering in Medicine and Biology Magazine, IEEE* Volume 26, Issue 3, May-June 2007 Page(s):29 – 33.

Moggridge B, *Designing Interactions*. The MIT Press, Cambridge, Massachusetts 2007.

Nielsen J. (1994), Heuristic Evaluation. In Nielsen, J and Mack, R.L., editors, *Usability Inspection Methods*. John Wiley and Sons, Inc. 1994.

Rubin, E. & Yates R. (2003). *Microsoft .NET Compact Framework*. Edited by Sams Publising, 2003.

Schmidt A. (2005) Interactive Context-Aware Systems Interacting with Ambient Intelligence. *Ambient Intelligence*. G. Riva, F. Vatalaro, F. Davide, M. Alcañiz (Eds.). IOS Press, pp. 159-178, 2005.

Schomaker et al., (1995) Schomaker, L., Munch, S., Hartung, K., 1995. A taxonomy of multimodal interaction in the human information processing system. *A Report of ESPRIT Project 8579 MIAMI, Technical Report*, Nijmegen Institute of Cognition and Information, The Netherlands. February 1995.

Sotirovski, D. Heuristics for Iterative Software Development. *Software, IEEE*. Volume 18, Issue 3, May 2001 Page(s):66 – 73.

Villalba E, Peinado I, Arredondo MT. (2007) User Interaction Design for a Wearable and IT based Heart Failure System. *Lecture notes in Computer Science* Volume 4551/2007.

21

Strategic Negotiations in Tabletop

Tokuo Yamaguchi[1], Sriram Subramanian[2],
Yoshifumi Kitamura[1] and Fumio Kishino[1]
[1]Osaka University, [2]University of Bristol
[1]Japan, [2]United Kingdom

1. Introduction

When working in groups, we often conduct face-to-face meetings to accelerate the exchange of ideas or opinions, or to complete a cooperative task. Many studies and systems have been proposed to facilitate face-to-face collaboration during group meetings or discussions.

Collaboration may be sought personally or imposed managerially. A variety of collaborative situations arises in everyday situations and has been analyzed by many researchers (Lim, 2003). Some types of collaborative tasks, especially in business situations, fundamentally include both competitive and collaborative aspects. Some examples can be found in playing card games or board games with friends or family. Members are competing with each other but the ultimate goal is for everyone to enjoy the game. Other examples can be found in various trading floor-like and auction scenarios.

In a collaborative task that includes both competitive and cooperative aspects, one of the fundamental actions of participants is to achieve their desired results through negotiation with other participants. A participant has to observe the transition of the task and find the ripest timing and best partner with whom to negotiate. Through the negotiation with a particular partner, each participant attempts to increase his individual benefit and this leads to an increase in the group's benefit. This process can be called "strategic negotiation" and is complex because conflicts between the personal and group priorities frequently arise. While it has been discussed in business situations (e.g. Dietmeyer & Kaplan, 2004), there is no investigation of this aspect with respect to designing tabletop systems.

On the other hand, digital tabletops in particular are increasingly being situated in a variety of work and public spaces. Collaborative table displays are designed to enhance functions of ordinary meeting tables to support small-group collaborative activities and have lately considerable attention (Scott et al., 2003). One of the most important research topics of collaborative table displays is to develop techniques to support transitions between personal and group work. For example, many systems typically use multiple displays to share information among participants in a meeting, and use one or several large wall displays to show public information. Strategic negotiations also require users to support transitions between personal and group work, and when we negotiate with others while sharing information, we often have to be careful about which parts of the information can and cannot be shared (Elwart-Keys et al., 2006). It offers them the opportunity to consult hidden information to make an informed decision or to present information at the most appropriate time to maximize its impact and increase its value to the presenter. However, digital tabletops today are either clumsy or not capable of effectively handling strategic negotiations. A special framework is often required to allow multiple users to deal with private as well as public information.

As a display device that suited for strategic cooperative task, we have proposed a collaborative, public and private, interactive display, called "SharedWell" (Kitamura et al., 2005). With this table users can create, manage and share information intuitively, strategically and cooperatively by naturally moving around the display. Users can interactively control private and public information space seamlessly according to their spatial location and motion. It enables users to dynamically choose negotiation partners, create cooperative relationships and strategically control the information they share and conceal. Furthermore to understand the low-level dynamics of user actions that accompany strategic negotiations, we have conducted an observational study to investigate strategic negotiations in various digital tabletop settings (Yamaguchi et al., 2007). Our results show that in the real world, strategic negotiation involves three phases: identifying the right timing, using epistemic actions to draw attention and evaluating the value of the negotiation. We repeated the real-world experiments in different digital tabletops and found several differences in the way users initiate and perform strategic negotiations. We identify many implications for the design of digital tabletops that arise from our findings.

2. Related Work

Recently, there has been a proliferation of systems and techniques that support digital tabletop interactions. Here we review the literature in two related areas – we present previous efforts in prototyping novel tabletop systems and investigations into the dynamics of face-to-face collaboration and various tabletop designs for managing user privacy in public information spaces.

2.1 Tabletop Systems

From time immemorial, tables have been used to discuss and make important decisions by a group of co-located people. The example of King Arthur's fabled Round Table still persists in the popular imagination. Today, we often have discussions while standing or sitting around a table to accelerate the exchange of ideas with multiple persons. Focusing on this type of interaction, there is much literature devoted to interactive tabletop displays to support face-to-face cooperative works. For example, InteracTable allows a group to annotate digital content on a computationally-enhanced table (Streitz et al., 1999), and DiamondTouch is a touch-sensitive tabletop display for multiple users (Dietz & Leigh, 2001). ConnecTables allows users of combined mobile desks to create a larger horizontal workspace and share and exchange documents, and a rapid sub-grouping in an office environment can be elegantly achieved (Tandler et al., 2001). An approach to tangible interface that uses phicons and phandles on the tabletop can be found in metaDESK (Ullmer & Ishii, 1997) and Sensetable (Patten et al., 2001). Augmented Surfaces (Rekimoto & Saitoh, 1999) is an example of a shared continuous workspace that combines walls, tabletops and laptops. Other tabletop displays are surveyed in (Scott et al., 2003).

Many researchers have investigated tabletop collaboration and proposed some characteristics as foundations for the design of interaction techniques. Pinelle et al. propose the *mechanics of collaboration* as a set of low-level actions and interactions that must be supported if team members are to accomplish a task in a collaborative fashion (Pinelle et al., 2003). Basic actions include communication, coordination, planning, monitoring and protection. Kruger et al. studied the *role of spatial orientation* on communication and collaboration, through observational studies of collaborative activity at a traditional table (Kruger et al., 2004). They found that orientation is important in establishing personal and group spaces and in signalling ownership of objects. Ryall et al. explored the effect of table size and number of collaborators on collaboration (Ryall et al., 2004). They found that even larger groups were successfully able to manage work at a small table. In order to avoid interference, collaborators usually separated the workspaces based on their seating positions and the task semantics (Tse et al., 2004). Scott et al. took a closer look at how *territoriality affects* collaboration in tabletop workspaces (Scott et al., 2004, Scott et al., 2005). They found that three types of territories were common – personal, group, and storage territories – and that these spatial divisions helped coordinate people's activities in shared tasks. Recently, Tang et al. investigated various forms of collaborative coupling, the manner in which collaborators are involved and occupied with each other (Tang et al., 2006). They identified six distinct styles of coupling based on whether users work in the same area of the shared workspace or if they perform similar tasks.

Most of the above research focused on designing systems for face-to-face cooperation and understanding users' pragmatic actions in cooperative settings. Most of the collaboration characteristics describe physical actions that allow users to cooperate better. It is not clear if these findings transfer directly to strategic collaborations. There is no reported investigation on the dynamics or characteristics of managing the various spaces during a strategic negotiation.

2.2 Privacy in Public Information Spaces

The space between users can be broadly divided into private, personal, shared and public spaces. Private space can be defined by the area where the owner can see and manipulate data but others cannot see the data or observe the owner's detailed manipulations of the data; personal space is the area where the owner can see and manipulate data while at the same time other users can observe the owner's actions in that area (without being able to observe the details of the data); a public space is defined by the area that allows all users to see and manipulate all data in it; a shared space is a form of public space that is created for a specific subgroup of users.

Because all the pieces of information displayed on the aforementioned tabletops can be easily observed by all participants equally, these systems do not provide appropriate support for strategic negotiations. For the purpose of strategic negotiations, a participant should be allowed to maintain parts of the information in a private space that is not readily observable by other participants until the opportunity is ripe for sharing all or part of that private information. The next paragraph describes digital tabletop examples that protect private information in public spaces.

In Augmented Surfaces (Rekimoto & Saitoh, 1999), the authors support a private space that is integrated with the public space through an interaction technique called hyperdragging. Users can control which space to put the information in by using an amplified dragging gesture to move data between the private and public spaces. In RoomPlanner, users can create a private space on the table in front of them using a hand gesture that physically occludes the information from the other user's view (Wu & Balakrishnan, 2003). The Lumisight Table provides different images to different users around the tabletop and has private spaces for individual users but lacks a public space in which information can be shared by multiple users (Kakehi et al., 2005). UbiTable allows two users to transfer digital media from their laptops (private space) to a tabletop display (public space) where it can be shared and annotated (Shen et al., 2003). It also includes a personal space along the boundary between the private and public space where users can observe each other's actions. Storage Bins is a mobile storage mechanism that enables access to stored items anywhere on the tabletop (Scott et al., 2005). The electronic pile of stored items may be useful for hiding information from the other users and relies on social protocols to prevent others from manipulating them.

While system designers have built several tables that support private and public spaces, investigations into tabletop collaboration suggest the need for dynamic personal and public spaces. It's not clear from the literature what the interplay is between private, public and personal spaces with regards to strategic negotiations.

3. SharedWell

3.1 Principle

The SharedWell system consists of a normal display and a display mask, which has a hole in its center (Fig. 1(a)). The display mask is placed over the display surface at a suitable distance. Each user observes the display system through the hole. By controlling the position of the information display area for each user according to own viewpoint (measured by a suitable tracker), each user can always see own individual area of the display. Figure 1(b) shows an example in which three users simultaneously observe information on different display areas. The information display area P in Fig. 1(b), which is derived from user P's viewpoint, is an area for showing the information for user P alone. The other users (Q and R), standing at their positions in this figure, are unable to see the P's individual information display area because the display mask adequately occludes this area. Therefore, this area can be used to show the private information of P.

On the other hand, the Q and R's information display areas overlap; therefore this overlapping area can be used to show the public information shared by Q and R. The areas other than the overlapping area are areas where only one of them can observe information; as a result, these areas can be used to show the private information of Q or of R.

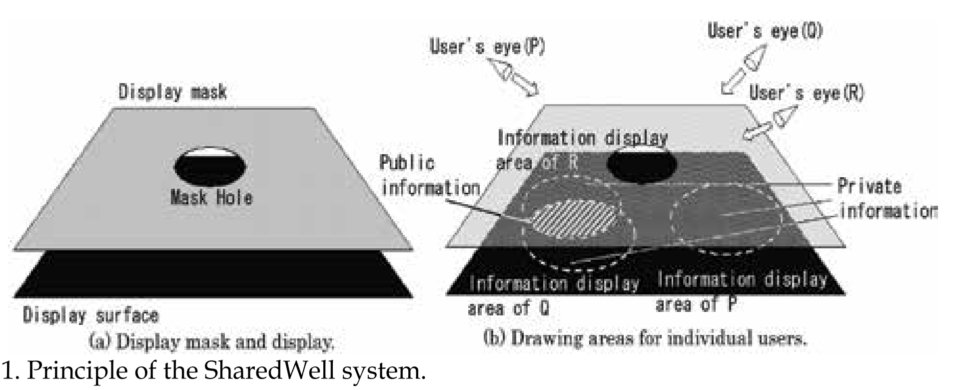


Fig. 1. Principle of the SharedWell system.

If a user moves own head, the information display area corresponding to the user also moves according to the motion of the viewpoint. This feature enables users to dynamically determine—by simply moving around the display—which display areas are used for private information and which display areas are used for public information that might be shared with a particular user.
The configuration of the display and the mask is along the same lines as the IllusionHole, a stereoscopic display system for multiple users (Kitamura et al., 2001), although stereoscopic images are not shown in this configuration. In IllusionHole the overlapping of image drawing areas among adjacent users was a weakness, which the design tried to avoid. In our current display system this weakness is turned into a key design feature.

3.2 System Configuration and Application Example

The proposed method is implemented by using a 68-inch conventional display system, i.e., BARON (Barco), to show the effectiveness of the method. The display system used in the trial system has a 1,360-mm-wide and 1,020-mm-deep display surface, and the height of the display surface from the floor is 1,000 mm. The display mask is placed over the display surface at a distance of 150 mm. The radius of the display mask hole is 200 mm. Each user's head position is tracked by an Intersense IS-600 Mark 2 SoniDisc, which is an acoustic 3D positional tracker. The system configuration is shown in Fig. 2. Here, a variety of interaction devices can be considered for manipulating objects. In this prototype, a game controller (SONY Dualshock2) is used for each user to manipulate objects.
Figure 3 shows an example of applying the SharedWell system to a goods-barter task. Three users (named Red, Green, and Blue, from left to right) are exchanging miscellaneous goods (such as used watches, televisions, jeans, and so on) by using their private and public display areas effectively. In this example, each user has his own miscellaneous goods to be exchanged in his private display area, and move the appropriate ones to the area shared with a particular person who is interested in the exchange. Each user can see the other participants' faces, and the negotiations progress through eye contact and hand/body gestures, which provides subtle communication cues. A user determines the best partner with whom to exchange information by changing his standing position and face-to-face communication using oral conversation. Figure 3 shows a the green user standing at the center, considering with which partner to exchange by comparing offered goods displayed on the overlapping areas with left user (Red) and right user (Blue). The top three images in Fig. 3 are user's views for Red, Green, and Blue (right, middle and left, respectively).

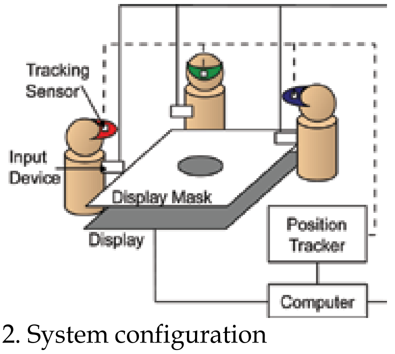


Fig. 2. System configuration


Fig. 3. A goods-barter application

4. Strategic Negotiations

4.1 Study in Real World

In order to better understand the characteristics and factors that influence strategic negotiations during face-to-face collaboration, we conducted an observational study in various collaborative settings at a real-world table. The results of this study were used to help focus our study of a digital table.

4.1.1 Task and Method

We designed three strategic tasks for groups of four to six people at a single table. The tasks were designed to leverage various types of digital tabletop settings. We were interested in three types of negotiations: public-space negotiations like on an auction floor, shared-space negotiations like on a trading floor, and shared-public negotiations like in a boardroom meeting.

Task 1: Here participants build a story based on a given theme on a large sheet by linking 10 images into a storyboard. Each participant was given 10 different images from which they had to select five to help build the story. After selecting their images, each participant had to convince the others to use as many of their five selected images as possible to create the story line. Since there were only 10 spots available for the storyboard, not all images from all participants made their way into the storyboard. Participants were instructed to build the storyboard with the intention of using as many of their images as possible. This task requires users to negotiate strategically in the public space where everyone could observe and interact with everyone else.

Task 2: The primary objective of this task was to examine strategic negotiations when users have to perform shared-public negotiations with several different partners at a table. The task is similar to the first one, but participants build two small stories and combine them into a big story by adding an image at the end. As with Task 1, after selecting their five images, participants are divided into two groups and asked to build two small stories. In this task, the story built by each group must have eight images, so each participant has to convince the others to use as many of their own five selected images as possible for the storyboard. Then, one image out of the remaining images has to be selected to serve as a link between the two stories. Participants try to convince each other that one of their own images best represents the two stories. This task requires the users to negotiate with partners in a dynamically created shared space followed by negotiations in the public space.

Task 3: The third task was a card game called "Pit" that uses 52 playing cards. Each participant starts with five cards and one card is placed face-up in the center of the table. Participants take turns exchanging cards to acquire a hand consisting of five cards of a similar suit (a flush in Poker). There are two ways to exchange cards - participants can either swap one of their own cards for the card in the center of the table or can choose another participant with whom to negotiate and trade. In the second case, either of the participants can reject the negotiation before the transaction is completed. This game examines strategic negotiations when partners use private spaces in conjunction with shared spaces to engage with negotiating partners.

Participants and Method: Six groups participated in this study in a between participants design. For Tasks 1 and 2, each group consisted of four participants. Task 3 consisted of one group of four and one group of six participants. All 26 participants were university students aged 24-30 and included both females and males of varying ethnic origins. Prior to the task,

each group received instructions on how to perform the task. Tasks 1 and 2 lasted 15 minutes each, whereas Task 3 lasted approximately 30 minutes. During each task, participants were comfortably seated around the table.
All sessions were videotaped and analyzed to compare and contrast the sharing of both the tabletop workspace and the objects on the table.

Fig. 4. Building a story in Task 1 (left), building stories in groups in Task 2 (middle), and playing a card game called "Pit" in Task 3 (right).

4.1.2 Results and Discussion

Figure 4 shows a snapshot of each task. Based on the analysis of the video, we observed that participants went through three successive stages when accomplishing strategic negotiation: Timing, Epistemic Action and Proposition Evaluation.
Timing: Capturing and retaining attention is important for strategic negotiations. In real-world setups, this is left to the charisma of the user. So users need to interrupt and gain others' attention to be able to efficiently negotiate. In Tasks 1 and 2, participants had to convince others to use as many of their own five selected images as possible. The center of the table (public space) and the spaces surrounding and between participants (shared space) were used to perform the main activities. During negotiation (when a participant suggested the use of his/her own image for the story), it was observed that the participant switched from his personal space to public space many times (81.6% of the total negotiations, and the negotiation with conversations was 34.9% of these negotiations). For example, when the participants discussed transportation in their story, one participant suggested using an image of ships instead of an image of airplanes. He tried to gain the others' attention and negotiated strategically by moving the images back-and-forth between the personal and public spaces.
In Task 3, it was difficult for participants to be time efficient, because of the turn-taking nature of the game. Participants could speculate on which suits were being collected by observing the card in the public space or the cards requested for exchange. Despite the limited value of timing in this task, we found that participants were conscious of the value of timing. Most often participants indicated their sense of urgency or preparedness by moving their next card (face-down) into a make-shift personal space.
Epistemic Actions: Kirsh and Maglio introduced the idea of epistemic actions to understand how users perform certain actions to improve their cognition of the world (Kirsh & Maglio, 1994). They argue that epistemic actions are physical actions that users perform to reduce the memory, number of steps or probability of error involved in a mental computation. In our task, we found that participants performed three types of epistemic actions: i) checking

the details of images or the card's suit repeatedly, ii) shuffling the image with no apparent reason or rearranging card positions in order of suit in preparation for the next negotiation, and iii) hesitating and shuffling the image and card repeatedly just before starting a negotiation. All participants performed these actions, which enabled them to create a strategy or a plan-of-attack in several places.

We found that participants frequently performed epistemic actions before initiating or terminating a negotiation. For example, in Tasks 1 and 2, participants often arranged the images with the storyline in their mind and tried to advocate use of their own image through repeated shuffling motions during discussion of the story. In Task 3, we observed that participants checked a card's suit endlessly while awaiting their turn and also rearranged the card's position when comparing their own cards and the card indicated for exchange. They also rearranged the card's position after exchange in preparation for the next negotiation. They were often observed to hesitate or reconsider their move when starting the negotiation.

Proposition Evaluation: We found that there were two steps in evaluating the value of a proposition to the current negotiation. The initial step is a coarse grain evaluation of the value of the negotiation. It means that a decision is reached without much consideration, and the object of negotiation is either deemed to be potentially valuable or not interesting at the current moment. The second step is accepting that the object could have potential value and performing a detailed examination of its value to the current negotiation.

For example, in building a story, when a participant suggested a different plot for the story by using his own images, he tried to indicate his image to the others via hand and body gestures that pointed to the images in his own personal space (23.1% in Task 1, and 88.9% in Task 2). Other participants recognized his actions and evaluated his proposal in a two-step process. If the initial examination suggests that the proposal is attractive, he is allowed the opportunity to place the images in the public space (Task 1) or the shared space (Task 2), and then present his case to further the discussion and evaluation. These actions were often observed and considered to be a more useful method for efficient collaboration. In Task 3, the cards introduced for transaction were carefully placed in front of the other participant by the giving participant (84.3% of the total negotiations). The exchanged card was usually quickly evaluated in the receiving participant's personal space (fast initial evaluation). If this card was rejected, it was moved back to the personal space of the giving participant. However, if the card was not immediately rejected, it was moved to the private space where it was evaluated again, this time more carefully in comparison with the other cards in the hand. Then if this card was not accepted for exchange, it was moved back to the personal space of the giving participant.

Summary: In all tasks, we observed the above sequence of actions for strategic negotiations. In general, users rely a lot on their personal space to effectively time the negotiation and use various forms of hand and body gestures to draw attention to their proposition or evaluate another user's proposition.

4.2 Study at Digital Table

In order to further explore the sequence of actions for strategic negotiations and examine similarities and differences between digital and real-world tabletop negotiations, we conducted a similar study using digital setups.

4.2.1 Task and Method

Our study in digital setups involved observing six groups (four participants each) performing in the same setting. We carried out our observations using three digital tabletops inspired by existing systems. All experimental setups used a table of size 1261×1530 sq mm and a horizontal display of size 635×1030 sq mm. Users could stand comfortably in front of the table. To keep the total experiment time to a reasonable amount, we restricted the task to Task 1 of the real-world setup. Similar to Task 1, 40 images were used for each task. Participants were only allowed to change image locations; no other image manipulations were allowed.

At the beginning of each session, the groups were given instructions indicating how to perform the task. Following the instructions, the participants had a three-minute practice session before beginning the experimental task on each setup. For each setup, the task duration was 15 minutes.

(a) UbiTable-Inspired (b) DiamondTouch-Inspired (c) SharedWell

Fig. 5. Three digital setups.

4.2.2 Tabletop Setups

Figure 5 shows the different experimental setups on the table. The three digital tabletop systems were inspired by UbiTable (Shen et al., 2003), DiamondTouch (Dietz & Leigh, 2001) and SharedWell. These systems were chosen for their different use of private, personal and shared spaces.

UbiTable-Inspired: Figure 5(a) shows the digital setup inspired by the UbiTable system. This system uses four small computers (two notebooks and two tablet computers) as the users' private space, and each user's screen top is connected to a large display. Users can handle their contents on their own screen as private space and transfer them to the large display as public space by moving them to the top of their own screen. Essentially, users can use the large display and their own small computers seamlessly, but cannot access each other's small computers. This system did not support personal or shared spaces.

DiamondTouch-Inspired: This digital setup is inspired by the DiamondTouch system in the sense that it's a large public space without any private, personal or shared spaces. Figure 5(b) shows the setup using a large single display. Participants can see all of the contents and can control their own cursors via an input device.

SharedWell: Figure 5(c) shows the users, whose head positions are tracked, looking through a hole in the table to view their digital contents. This hole allows users to maintain a private view of their contents even when they move around the table described in section 3. To show or share their own contents, users have to come close enough to each other so that their views through the hole overlap. The overlapping region creates a shared space for users to show or share contents. If a user wants to show or share contents with a particular partner, he puts the contents in the overlapping area by approaching the partner.

In all systems, participants were divided according to four colors (red, green, blue and yellow) and could recognize their own images and cursor by the color. A game controller was used as an input device for each participant to control the digital contents. All sessions were videotaped and analyzed to compare and contrast the sharing of both the tabletop workspace and the objects on the table.

4.2.3 Results and Discussion

Based on the analysis of the video and the image movements, we describe our observations on strategic negotiation in the different digital tables.

UbiTable-Inspired: Participants had 10 different images to view on their private screens (private spaces) at the beginning, where they often rearranged the images. These epistemic actions helped them create a storyline. Soon after determining the images to be used, some participants moved them to the public space as the selected images. However, we also observed that others kept relocating the images (including images to be used in their own storyline) in their own private spaces. As their recommendations, these participants gradually moved the images into the public space as the session progressed. We believe that they preserved these images in their own private spaces so as to strategize and wait for the right timing to present them. In addition, we also observed that when participants initiated a discussion or negotiation with the others, they pointed at their image in the public space by using their own cursor or hand. If the image was not accepted after the group evaluation (this is considered the second step of the evaluation), the owner typically moved it back to own private space.

DiamondTouch-Inspired: This setup allowed all participants to see and handle all images during the entire session. At the beginning, these 40 images were distributed to each participant and placed in front of her at the edge of the table. In order to create their own storylines, participants rearranged their images within the area where the images were delivered. After determining images to be used for their own storylines in their minds, they moved the unnecessary images to the corner of the table, whereas they left the necessary images stationary in front of them as the selected images. We noticed that all participants handled their images only in the public space. They did not wait for the appropriate timing to present a particular image to the others efficiently because they could see all the images all the time during the session. Here a participant had to move the images by a cursor, and this seemed to present cognitive uncertainty to the other participants about the operator's intentions and so on. The epistemic actions occurred in the form of redundant movements of images, however they were not effectively used among the participants because of the cognitive uncertainty of the cursor operation. From the observation, they seemed to be used only for the careful consideration of the participant. When a participant tried to propose a particular image to be used in the storyline shared by all participants, she moved the image from the area in front of her to the public space. If there were images forming a storyline in the public space, we often observed that a participant put his image on top of these images, occluding them. By using these actions he could initiate negotiation and force a group evaluation of the proposal. Evaluations were often done in a single step.

SharedWell: Because of the nature of this system, participants performed many more physical movements for initiating or starting negotiations than with any of the other systems investigated. At the beginning, we observed that participants positioned themselves away from the table, therefore, distances among participants were long enough to preserve their own private spaces, avoiding the overlapping of their spaces to create the shared spaces. When a participant tried to initiate negotiation or discussion, he moved himself closer to the other participant(s) to create a shared space. This space was used as the public space for the discussion and negotiation between the two (or sometimes more) participants. Actually, they initiated negotiation or discussion by moving images from their private spaces to the shared space when the timing was right. All of the actions mentioned here could be observed in all of the participants, and they demonstrated the epistemic actions as well. Soon after the images were moved to the shared space, others evaluated them quickly. This was the first step of evaluation and was repeated several times by changing partners. Through a sequence of negotiations, images were gradually collected in the public space to form a storyline. These images were finally re-evaluated carefully. This was the second step of the evaluation process.

5. Discussions

From the results of the experiment, it is clear that there are differences in how participants engage in strategic negotiations in real-world and digital setups.

5.1 Real-world Setups vs. Digital Tabletop Setups

Based on the observations of the real-world setups and digital setups, we found many problems with the current digital tabletops. These problems relate to differences in user actions at different stages of strategic negotiation.

Timing: For strategic negotiation, users typically use their private spaces to examine information that needs to be shown to the other participants at the most appropriate time to maximize its impact. In the UbiTable-inspired system, we found that the participants could not use their private spaces effectively. We observed that some participants moved their selected images from the private space to the public space directly soon after the session started. Here, they did not create their own personal space explicitly, and thus they missed a chance to strategize and present information at the most appropriate time. In the DiamondTouch-inspired system, participants could not determine the best timing to propose their images since all information was visible to all participants from the beginning of the session. On the other hand, in the SharedWell system, participants could determine the most appropriate timing by observing the other participants' explicit movements; however, the physical movements forced participants to miss some opportune moments and at the same time quickly fatigued them.

Epistemic Actions: In the real-world settings, all participants easily noticed all negotiations and actions, including epistemic actions performed by a participant. Through these actions, a participant could understand the status of the collaboration taking place on the table, e.g. who was negotiating with whom, and results of the negotiation. At first we thought that the DiamondTouch-inspired system was the most similar to the real-world settings except that it required indirect manipulation using a cursor. However, we noticed that participants had difficulties in recognizing epistemic actions made by others because of the cognitive uncertainty of the cursor operation. On the other hand, in the other two systems, participants could perform the epistemic actions. In the UbiTable-inspired system, participants had their own private spaces and showed epistemic actions by rearranging their images while they considered their own storylines. Similarly, in the SharedWell system, they also had their own private spaces, and through physical motions such as enlarging the private space to be shared with adjacent participants or moving images from their private spaces to the shared spaces, they showed epistemic actions.

Proposition Evaluation: We found that users often evaluated a proposition by transitioning attention between spaces, especially from the private space to the personal space and from the personal space to the shared/public space. Explicit transitions between spaces attract other persons' attention and help provide cues of the evaluation process to all participants. From observations in the real-world settings, we understand that the two-step evaluation is important for the strategic negotiations. Moreover, through the investigation of digital tables, we found that the facility of providing personal spaces is especially important for proposition evaluation. The SharedWell system was designed to transfer the information between spaces because of the nature of this system. Therefore, participants could often perform the two-step evaluation efficiently. In the DiamondTouch-inspired system, participants could see and manipulate all the images on the display; therefore, they had to transfer the images by using explicit actions such as hand gestures or utterances. On the other hand, the participants on the UbiTable-inspired system moved their images from the private space to the public space without paying special attention to the timing; therefore, they often missed the chance to do a first-evaluation at the most appropriate time.

5.2 Supporting Strategic Negotiations on Digital Table

Digital tables can be made more efficient for strategic negotiation by improving various aspects of timing, epistemic actions and proposition evaluation.

Value of Personal Space: One of the crucial elements of strategic negotiation in real-world collaboration was a users' ability to maintain a personal space. Users often moved valuable negotiation data to the personal space, which served two purposes; first it informed others that this user had something that could be perceived as useful to the negotiation without giving them insights to evaluate the value of the data. This gave the user the opportunity to initiate negotiation when the moment was ripe. The second benefit of having information in the personal space was that at the right moment the user could easily introduce data to the public space for negotiation, and because others were anticipating this, they were more willing to listen to the user's proposition and were not taken by surprise. We believe that digital tables should support both personal and private spaces for enhanced strategic negotiations. Scott et al. suggest that digital tables do not have to support personal spaces because users generate these spaces by themselves (Scott et al., 2004, Scott et al., 2005). However, as described in the case of the UbiTable-inspired and DiamondTouch-inspired systems, very few participants created personal spaces because they either failed to appreciate their value or had insufficient workspace area. Systems could provide a default personal space that can be fluidly and intuitively moved around. We agree with Scott et al. that it is important that users are able to flexibly and dynamically increase, decrease or relocate personal space within the workspace.

Sensitivity to User's Hand and Body Gestures: In the real-world task, users often created opportunities for negotiation by using various hand and body gestures. This rich communication language provided all users with awareness of each other's intentions allowing them to anticipate forthcoming actions. Hand and body gestures like rearranging cards within in the hand and drifting an image in-and-out of the public spaces provide rich awareness cues that users often pay attention to subconsciously. When a person tries to negotiate with others profitably, it is reasonable to expect this person to be aware of the other users' actions without compromising their privacy. It is this awareness of details that enables users to efficiently strategize negotiations in a group. Therefore, in order to support these negotiations, digital tables should be sensitive to users' actions related to body or hand gestures, and at the same time have the ability to keep private information private.

Interruptability and Epistemic Actions: We often observed that users interrupt each other with finesse to grab attention and propose an item for negotiation. While we did not explicitly examine interruption in our study we feel that digital systems should be proactive in providing support for interruptability. Systems could leverage a variety of multimodal information channels to further enhance strategic negotiation. As outlined in the results section, users rely on various epistemic actions to propose and evaluate negotiations. These could range from explicit transitions between spaces to attract other persons' attention to pondering and fiddling with the hands to indicate serious contemplation of the value of a proposition.

5.3 Implications for Design

The results of our investigation into strategic negotiations in digital tables have several implications for the design of future digital tabletop systems.

Support Creation of Personal Spaces: In our study of real-world strategic negotiations, we found that users often create personal spaces to negotiate efficiently. Tabletops that support strategic negotiations should not rely solely on private or public spaces and transfer information from private to public space directly. When designing strategic negotiations efficiently, an important implication is that the system must support creation of personal spaces.

Support Flexible and Fluid transition between Spaces: In all our digital systems, users repeatedly transferred information between private, personal and shared/public spaces. Fluid transition of information between these spaces is important for conducting strategic negotiations efficiently. Many researchers are exploring novel interaction techniques to support flexible and fluid transitions between different spaces. Our results reaffirm the need to do so.

Tabletop Systems should be Sensitive to Body and Hand Gestures: We observed that users relied on their hand and body gestures to negotiate in real-world setups. For strategic negotiation, these gestures also helped users to know the intentions of another other person's actions exactly. Therefore, we believe digital tables should be sensitive to these gestures while at the same time not demanding from users explicit gestures as with the SharedWell system. We believe that future systems must harness the body and hand gestures of users with greater finesse for strategic negotiations.

Provide Greater Support for Epistemic Actions: We repeatedly observed that users performed many epistemic actions during collaborations. While the digital tabletop systems did not explicitly factor in epistemic actions in their designs, the users were able to perform some of the epistemic actions observed in real-world settings. However, for digital tabletops to attain the flexibility and fluidity of real-world collaboration, we need to explicitly take into consideration typical epistemic actions when designing future tabletop systems.

6. Conclusion

In this article, we discussed the concept of the SharedWell, a collaborative, public and private, interactive display, and strategic negotiations that include both competitive and cooperative aspects. We investigated strategic negotiations in real-world face-to-face collaborations and compared the findings with three digital tabletop systems include the SharedWell system. We found that users strategize at multiple levels, preferring to use a personal space of dynamically re-changing size. We also identified several characteristics of group dynamics that can be valuable for designing next generation tabletop systems. Our results show that in the real-world, strategic negotiation involves three phases: identifying the right timing, using epistemic actions to draw attention and evaluating the value of the negotiation. We repeated the real-world experiments with different digital tabletops and found several differences in the way users initiate and perform strategic negotiations.

We believe that the collaborative interaction paradigm the SharedWell system suggests can be easily extended to support a wide range of activities. In the future we plan to look into studying strategic negotiations with a variety of personalities and leadership qualities to see if there are any differences. We are also exploring novel ways to extend tabletop systems like the SharedWell to capture the nuances of negotiation that are evident in the real world.

7. Acknowledgments

This study was supported in part by "Global COE (Centers of Excellence) Program" of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

8. References

Dietmeyer, B. & Kaplan, R. (2004). *Strategic negotiation: a breakthrough four-step process for effective business negotiation,* Kaplan Business, ISBN: 978-0-793-18304-3

Dietz, P. & Leigh, D. (2001). DiamondTouch: a multi-user touch technology, *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, pp. 219-226, ISBN: 1-58113-438-X, Orlando, Florida, United States, Nov. 2001, ACM Press

Elwart-Keys, M.; Halonen, D., Horton, M., Kass, R. & Scott, P. (1990). User interface requirement for face to face groupware, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* pp. 295-301, ISBN: 0-201-50932-6, Seattle, Washington, United States, Apr. 1990, ACM Press

Kakehi, Y.; Iida, M., Naemura, T., Shirai, Y., Matsushita, M. & Ohguro, T. (2005). Lumisight table: an interactive view-dependent tabletop display. *IEEE Computer Graphics and Applications,* Vol. 25, No. 1, (Jan.-Feb. 2005) 48-53, ISSN: 0272-1716

Kirsh, D. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science: A Multidisciplinary Journal,* Vol. 18, No. 4, (Jan. 1994) 513-549, ISSN: 0364-0213

Kitamura, Y.; Konishi, T., Yamamoto, S. & Kishino, F. (2001). Interactive stereoscopic display for three or more users, *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pp. 231-239, ISBN: 1-58113-374-X, Los Angeles, California, United States, Aug. 2001, ACM

Kitamura, Y.; Osawa, W., Yamaguchi, T., Takemura, H. & Kishino, F. (2005). A display table for strategic collaboration preserving private and public information, *Proceedings of IFIP 4th International Conference on Entertainment Computing*, pp. 167-179, ISBN: 978-3-540-29034-6, Sanda, Japan, Sep. 2005, Springer

Kruger, R.; Carpendale, S., Scott, S. D. & Greenberg, S. (2004). Roles of orientation in tabletop collaboration: comprehension, coordination and communication. *Computer Supported Cooperative Work*, Vol. 13, No. 5-6, (Dec. 2004) 501-537, ISSN: 0925-9724

Lim, J. (2003). A conceptual framework on the adoption of negotiation support systems. *Information and Software Technology*, Vol. 45, No. 8, (June 2003) 469–477, ISSN: 0950-5849

Patten, J.; Ishii, H., Hines, J. & Pangaro, G. (2001). Sensetable: a wireless object tracking platform for tangible user interfaces, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 253-260, ISBN: 1-58113-327-8, Seattle, Washington, United States, Mar.-Apr. 2001, ACM Press

Pinelle, D.; Gutwin, C. & Greenberg, S. (2003). Task analysis for groupware usability evaluation: modeling shared-workspace tasks with the mechanics of collaboration. *ACM Transactions on Computer-Human Interaction,* Vol. 10, No. 4, (Dec. 2003) 281-311, ISSN: 1073-0516

Rekimoto, J. & Saitoh, M. (1999). Augmented surfaces: a spatially continuous work space for hybrid computing environments, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 378-385, ISBN: 0-201-48559-1, Pittsburgh, Pennsylvania, United States, May 1999, ACM Press

Ryall, K.; Forlines, C., Shen, C. & Morris, M. R. (2004). Exploring the effects of group size and table size on interactions with tabletop shared-display groupware, *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, pp. 284-293, ISBN: 1-58113-810-5, Chicago, Illinois, United States, Nov. 2004, ACM Press

Scott, S. D.; Carpendale, S. T. & Inkpen, K. M. (2004). Territoriality in collaborative tabletop workspaces, *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, pp. 294-303, ISBN: 1-58113-810-5, Chicago, Illinois, United States, Nov. 2004, ACM Press

Scott, S. D.; Carpendale, S. T. & Habelski, S. (2005). Storage bins: mobile storage for collaborative tabletop displays. *IEEE Computer Graphics and Applications,* Vol. 25, No. 4, (July-Aug. 2005) 58-65, ISSN: 0272-1716

Scott, S. D.; Grant, K. D. & Mandryk, R. L. (2003). System guidelines for co-located, collaborative work on a tabletop display, *Proceedings of the eighth conference on European Conference on Computer Supported Cooperative Work*, pp. 159-178, ISBN: 978-1-402-01573-1, Helsinki, Finland, Sep. 2003, Springer

Shen, C.; Everitt, K. & Ryall, K. (2003). UbiTable: impromptu face-to-face collaboration on horizontal interactive surfaces, *Proceedings of the fifth International Conference on Ubiquitous Computing*, pp. 281-288, ISBN: 978-3-540-20301-8, Seattle, Washington, United States, Oct. 2003, Springer

Streitz, N.; Geibler, J., Holmer, T., Konomi, S., Muller-Tomfelde, C., Reischl, W., Rexroth, P., Seitz, P. & Steinmetz, R. (1999). i-LAND: an interactive landscape for creativity and innovation, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 120-127, ISBN: 0-201-48559-1, Pittsburgh, Pennsylvania, United States, May 1999, ACM Press

Tandler, P.; Prante, T., Müller-Tomfelde, C., Streitz, N. & Steinmetz, R. (2001). Connectables: dynamic coupling of displays for the flexible creation of shared workspaces, *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, pp. 11-20, ISBN: 1-58113-438-X, Orlando, Florida, United States, Nov. 2001, ACM Press

Tang, A.; Tory, M., Po, B., Neumann, P. & Carpendale, S. T. (2006). Collaborative coupling over tabletop displays, *Proceedings of the SIGCHI Conference on Human Factors in computing systems*, pp. 1181-1190, ISBN: 1-59593-372-7, Montréal, Québec, Canada, Apr. 2006, ACM Press

Tse, E.; Histon, J., Scott, S. D. & Greenberg, S. (2004). Avoiding interference: how people use spatial separation and partitioning in SDG workspaces, *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, pp. 252-261, ISBN: 1-58113-810-5, Chicago, Illinois, United States, Nov. 2004, ACM Press

Ullmer, B. & Ishii, H. (1997). The metaDESK: models and prototypes for tangible user interfaces, *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, pp. 223-232, ISBN: 0-89791-881-9, Banff, Alberta, Canada, Oct. 1997, ACM Press

Wu, M. & Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays, *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology*, pp. 193-202, ISBN: 1-58113-636-6, Vancouver, Canada, Nov. 2003, ACM Press

Yamaguchi, T.; Subramanian, S., Kitamura, Y. & Kishino, F. (2007). Strategic tabletop negotiations, *Proceedings of 11th IFIP TC 13 International Conference on Human-Computer Interaction*, LNCS 4663, Part II, pp. 169-182, ISBN: 978-3-540-74799-4, Rio de Janeiro, Brasil, Sep. 2007, Springer

Edited by Ioannis Pavlidis

This book includes 23 chapters introducing basic research, advanced developments and applications. The book covers topics such us modeling and practical realization of robotic control for different applications, researching of the problems of stability and robustness, automation in algorithm and program developments with application in speech signal processing and linguistic research, system's applied control, computations, and control theory application in mechanics and electronics.

Photo by Vladimir_Timofeev / iStock

IntechOpen

ISBN 978-953-51-5797-7