Context-based Innovative Mobile User Interfaces

Yun Zhou

To cite this version:

Yun Zhou. Context-based Innovative Mobile User Interfaces. Human-Computer Interaction [cs.HC]. Ecole Centrale de Lyon, 2012. English. NNT : . tel-00813532

HAL Id: tel-00813532
https://theses.hal.science/tel-00813532

Submitted on 15 Apr 2013

N° d'ordre 2012-29
Année 2012

Thèse

Context-based Innovative Mobile User Interfaces

Présentée devant
Ecole Centrale de Lyon

Pour obtenir
Le grade de docteur

Ecole Doctorale Informatique et Information pour la Société (EDIIS)

Par
Zhou Yun

Soutenue le 08 Octobre 2012 devant la Commission d'examen

Jury MM.

Christophe KOLSKI	Pr. (UVHC)	Rapporteur
José ROUILLARD	MdC HDR (Univ. Lille 1)	Rapporteur
Patrick GIRARD	Pr. (Univ. de Poitiers)	Examinateur
Franck TARPIN-BERNARD	Pr. (UJF Grenoble)	Examinateur
Bertrand DAVID	Pr. (ECL)	Directeur
René CHALON	MdC (ECL)	Co-directeur

Acknowledgements

During the three years I spent working on this thesis, many people most kindly helped me in my everyday life and in the success of my research work here in France.

First, I have been greatly fortunate to be a student working with my supervisor Professor Bertrand David, who always provided me with insightful comments, patiently answered all my questions, and guided me step by step in my research work. He also provided me with opportunities to get in touch with other researchers in international conferences, kindly supporting me in many ways during my study period. It would have been impossible for me to enjoy my interesting research and achieve the goal of my thesis without his help. I would also like to thank my co-supervisor René Chalon, who kindly answered all my questions concerning programming, research problems, and other technical problems, always providing me with useful solutions. I always obtained the inspiration and found the right solution after my discussions with him. My two supervisors are both excellent researchers in our field and I hope that my graduation will not mean the end of our collaboration and cooperation.

Second, I would like to heartily thank my colleagues Xu Tao, Florent Delomier, Cyril Bénazeth, Zhang Bingxue, and Jin Huiliang, who always exchanged their interesting research ideas with me, and generously gave me their time to help me prepare every evaluation in my thesis. I especially would like to thank Tao, my dear friend, who always encouraged me when I got frustrated. And I want to thank many Chinese students in ECL for their invaluable contribution in my evaluation: Zhu Chao, Li Huibin, Zhang Taiping, Feng Zhenfu, Zhu Nanhao, Ye Fanbin, Lv Ying, etc.

Third, I would like to specially thank Professor Christophe Kolski and associate professor José Rouillard, for having agreed to be reviewers of this thesis. They helped me improve the manuscript and gave me a lot of helpful comments. I also thank all other members of the jury, Professor Patrick Girard and Professor Franck Tarpin-Bernard, for having accepted to participate in the defense and evaluate work in the thesis.

Finally, I really want to thank my dear parents, who live in China but always gave me their selfless love, immeasurable sacrifices, and incessant encouragement. I also appreciate all the support and encouragement from my relatives, French teachers and friends in France, in China and all over the world. Without them, I could not progress as I have made.

Abstract

With the development of a wide variety of sensors and devices, computing is no longer limited to the desktop mode. However, the traditional user interface, used on the desktop computer, is no longer appropriate for ubiquitous computing. A sophisticated mobile environment requires dedicated design of interfaces, involving input and output techniques with new emerging features that go far beyond the capacities of traditional techniques. One of the solutions to enable ubiquitous interaction and end limitation of the desktop mode is nomadism, while another is mobility. We propose three interfaces related to these two solutions: In-environment interface (IEI), Environment Dependent Interface (EDI), and Environment Independent Interface (EII). We exclude IEI and mainly focus on wearable interaction.

This thesis aims to investigate research issues involved in the design, implementation and evaluation of EDI and EII. It presents our design approach to these three innovative interfaces (IEI, EDI and EII), their wearable configurations (camera-glasses device unit and camera-projector device unit), real examples of use (including the Research Team Interaction Scenario), and both the quantitative and qualitative user studies and evaluations to prove the feasibility and usability of our prototypes. Our work is a many-sided investigation on innovative wearable interfaces, as well as input and output techniques, which will pave the way for future research into wearable interfaces.

Table des matières

Sommaire

Table of Figures

1 Introduction

1.1 Introduction

In general, the research described in this thesis falls into the scope of Human-Computer Interaction (HCI), while also belonging to the domain of wearable computing. Tangible User Interfaces (TUIs), Augmented Reality (AR), Vision-based Interaction, and Projected User Interfaces have an impact on our research. We focus on the camera-glasses system and camera-projector system, as well as addressing the issues within the design, implementation and evaluation scope.

1.2 Motivation and Contributions

With the development of a wide variety of sensors and devices, computing is no longer limited to the desktop mode, but takes on a totally new look. At the same time, interaction modalities and interfaces have switched from WIMP to post-WIMP (Van Dam, 1997), and innovative inputs and techniques are being increasingly considered. These new interaction methods change people's lives and facilitate their tasks in everyday life and in the workplace, enabling people to access their personal data as well as public resources at any time and in any place. As technology progressively integrates every aspect of life, a greater requirement for innovative research into various aspects of ubiquitous computing has emerged. The issues related to ubiquitous computing and pervasive computing vary from the practical problems of user interaction modalities to the more ethical problems of privacy, data protection and social effect. We found that the traditional user interface, used on the desktop computer, is no longer appropriate for ubiquitous computing. Furthermore, it is insufficient and unable to satisfy the requirements of our daily tasks by simply emulating the existing WIMP modality. A sophisticated mobile environment requires dedicated design of interfaces involving input and output techniques with new emerging features offering far more than the capacities of traditional techniques.

One of the solutions available to enable ubiquitous interaction and end limitation of the desktop mode is nomadism, where the user is not equipped with any wearable or mobile devices. Another solution is mobility, where the user is equipped with wearable or mobile devices. Wearable devices can include the webcam, the pico-projector or other output displays. Mobile devices can include PDAs, smart mobile phones, etc. Classical portable devices such as laptops are excluded as mobile devices, since their size makes them unavailable and inconvenient to use when the user is walking or in other mobile settings. Also laptops are long to access input compared with mobile phones. However, the tablet or the special laptop could form one part of a wearable configuration, only contributing to the calculation function rather than other functions. To help the user interact all around and access information freely in the environment, we propose three interfaces (Zhou, David, & Chalon, 2011): In-environment interface (IEI), Environment Dependent Interface (EDI), and Environment Independent Interface (EII). We have chosen to focus on the last two interfaces: the environment dependent interface and the environment independent interface. EDI and EII are both based on wearable computing devices, allowing the user to interact in mobility. We aim to provide the user with information that is decided by the environment, i.e. the environment provides the users with information. In this way, the Environment Dependent Interface (EDI) refers to the strong relationship between the interface and the in-environment information. Going one step further, we propose the Environment Independent Interface (EII), which refers to the relationship between the interface and personal information. Contextualization is performed by the actual users by showing the webcam appropriate contextualized markers or menus. Users can contextualize their working environment by themselves.

System	Design of Interface			Continuum for Interface			Input and Output Devices				Selection Techniques					
	Nomadism	Mobility		Physical Interface	Mixed Interface	Digital Interface	Camera (Webcam)	Glasses	Pico-projector	Screen	ST 1	ST 2	ST 3	ST 4	ST 5	ST 6
	IEI	EDI	EII													
Preliminary System 1	×			×			×			×	×		×			
Preliminary System 2		×		×			×	×			×		×			
Preliminary System 3			×			×	×		×		×			×		
MobilePaperAccess		×	×	×			×	×				×	×	×		
PlayAllAround			×			×	×		×			×				
Bare-hand Interaction System			×			×	×		×							×
Mixed Interaction System		×	×		×		×		×			×			×	

ST1: Finger Entering Input (In Chapter 3)
ST2: Finger Hover Input (In Chapter 4, 5, 6, and 7)
ST3: Mask Input (In Chapter 3 and 4)
ST4: Book/Page Input (In Chapter 3 and 4)
ST5: Two-finger Pinch Input (In Chapter 6 and 7)
ST6: Fist-palm Input (In Chapter 6)

Figure 1.1 The design and development overview of our innovative user interface.

To concretize the concepts of EDI and EII, we developed different interaction techniques and systems. The summary of the actual interfaces, configurations and selection techniques related to the systems is listed as shown in Figure 1.1. Our work is mainly based on the basis of the wearable camera-glasses system and camera-projector system, and the contributions are described briefly as follows:

- First, we developed several preliminary systems using the finger entering gesture and tangible interaction techniques such as the mask selection technique for selecting interactive items in the paper-based interface and projected interface. The feasibility of our systems and interaction techniques are evaluated preliminarily.
- We improved our interaction techniques and then developed the MobilePaperAccess system which is a wearable camera-glasses system with a tangible user interface allowing mobile interaction. We organized an evaluation to compare techniques and obtain user comments and preferences for these techniques.
- We further investigated the hover gesture in-depth and explored the scalability of the projected interface. Based on the concept of EII, the

PlayAllAround system aims to provide both the nearer small-size interface and the farther large-size interface supporting private and public use.

- In addition, we compared the hover gesture and pinch gesture in mobile settings through the evaluation and discussed the bare-hand interaction techniques.
- Finally, we designed and developed the physical-digital mixed interface for EDI and EII. The mixed interface is based on a combination of the paper-based interface and the projected interface and utilizes both the hover gesture and the pinch gesture for interaction.

The EDI and EII designs have several advantages. First, they encourage the finger, one-hand, or even two-hand interactions, which are based on the natural, intuitive and touchless interaction modality. With respect to maintenance, it often requires one hand performing real tasks while another hand browses and retrieves information. Thus, the one-hand free interaction can help technicians work with real tasks and digital data simultaneously. With regard to interacting in the operating room (Wachs et al., 2008), the bare hand input modality can provide a sterile access approach for interaction. Second, since the EDI and EII share the same wearable configurations, the user is free to switch between the in-environment information and personal information. They allow interaction of personal data as well as public information. On the one hand, EDI provides the user with well-timed and well-located public information, while on the other, EII provides users with the opportunity to browse and retrieve information, search the schedule, handle personal tasks, etc., as they require. Third, they use our everyday skills of mouse pointing or even multi-touch gestures, which lower the burden of learning for users. Finally, the micro viewer display provides a private view for the user, and the pico-projector display provides a scalable interface, which has multi-scale projection sizes and satisfies the different interface size requirements.

To conclude, this thesis aims to investigate research issues involved in design, implementation and user performance of environment dependent and environment independent interface. To achieve these two interfaces in mobility, we employ the camera as input, which supports hand and finger gestures as input. And we leverage the goggle with a small screen and the projector as output, the former providing a private visual feedback and the latter providing the scalable visual output experience. With the aim of studying innovative wearable user interfaces which can help the user interact freely in an environment, we investigated the theoretical analysis, the innovative design, practical development, and both the quantitative and qualitative user studies. Our current work is a many-sided investigation into innovative wearable interfaces, as well as input and output techniques, which will pave the way for future research into wearable interfaces.

1.3 Thesis Structure

In Chapter 2, we first review the relevant research work inspiring our study on wearable EDI and EII in relation with ubiquitous computing. Then, we start the actual research with the design, prototypes and implementation of the In-environment interface, the Environment Dependent Interface and the Environment Independent Interface, described in detail in Chapter 3. All the research work in this thesis is based on the global view stated in Chapter 3. We investigate feasibility of the configuration and usability of the interaction on the basis of the camera capture, as well as marker recognition and the goggle attached with a small screen. We design and implement a series of prototypes of innovative interfaces, allowing users to finally interact with the in-environment information with at least one hand free.

To go one step further, the physical paper-based interface is studied as well as its input modalities and output. In Chapter 4, we design and develop a ubiquitous paper-based system for mobile interaction, known as MobilePaperAccess. This is a wearable camera-glasses system with a tangible user interface allowing mobile interaction. We access to the in-environment digital information from the paper interface, thus extending the input space and avoiding the problem of fat finger and occlusion of finger shadow. We propose a continuum from physical interface to digital interface in relation with EDI and EII, and we present the design, implementation and evaluation aspect of our MobilePaperAccess system. In this system, two interfaces (EDI and EII) and three input techniques (the finger input, the mask input and the page input) have been studied and evaluated.

To provide users freely with personal information, we propose the wearable one-hand gesture input and scalable projected interface and PlayAllAround system. PlayAllAround is a wearable camera-projector system with the scalable interface allowing mobile interaction, which provides both the nearer small-size interface and the farther large-size interface supporting both private and public use. The system achieve our concepts of Environment Independent Interface (EII), which focuses on enabling people to access their personal data and resources at any time and in any place. In addition, we propose the design of reference-cell and the principle from decomposition of the application tasks to formation of scalable interface. We finally evaluate the hover gesture and the scalable interface globally. This study in Chapter 5 has a major impact on the design of the projected interface which has the property of scalability uniquely.

In Chapter 6, we explore our wearable input techniques in a focalized way, as well as the situation suitable for the EDI and EII respectively. We investigate the wearable hover input gesture, the pinch gesture combined with the drag-drop input method, and the bare hand gesture such as a fist-palm gesture. We focus on easiness of learning and performing different hand input techniques.

Besides the paper-based interface, we propose integrating digital information with physical information, in a more seamless manner using the camera-projector system. In Chapter 7, we study the physical-digital mixed interface based on the EDI and EII. The interface consists of the marker-based part with ARToolKit tags and the projected part. In this way, the interface can contain more information and also has more dynamical choices for selecting information.

In Chapter 8, a detailed summarization of contributions is discussed as well as the research directions for a future study. To conclude, this thesis investigates many-sided aspects of wearable innovative interfaces in augmented reality. We provide a contribution to research into the wearable Environment Dependent Interface and the Environment Independent Interface, including a study on interaction design, technical development and implementation, and both the quantitative and qualitative evaluations of these innovative interfaces. The wearable interaction of the physical paper-based interface, physical-digital interface and the digital dynamic interface are explored and achieved. We also provide a deeper understanding into using the wearable camera-glasses system and the camera-projector system as the basis of mobile interaction. The work in this thesis is all based on the interaction of a single user. In a future work, we plan to take into consideration the multi-user interaction, social behaviors impacted on multi-user interaction and cooperation for wearable interaction.

2 Related Work

2.1 Introduction

In this chapter, we outline the relevant research work that helped inspire our study on wearable interaction in relation with ubiquitous computing. We study the input and output techniques in the field of wearable computing, as well as the related research areas: augmented reality, ubiquitous computing and tangible user interface.

The chapter is structured as follows: we first present an historical overview of wearable computing in section 2.2. We focus on interaction with mobility which is achieved via wearable configurations. We then study the applications and scenarios achieved in the field of wearable computing. In section 2.3, we describe the various hand input techniques designed and implemented by previous researchers. Visual output techniques are reviewed and discussed in section 2.4. In section 2.5, we examine and summarize other related research areas that have significance on and correlation with our study: augmented reality, ubiquitous computing and tangible user interface.

2.2 Wearable Interaction

In this section, we provide the background and historical overview of wearable computing enabling the interaction from WIMP (windows, icons, menus, and a pointing device) modality to post-WIMP modalities (Van Dam, 1997), as well as the contributions of these innovative interfaces and interactions related to wearable computing. We then discuss the various applications and scenarios experimented on the aspects concerning the industrial maintenance, military, education, medical treatment, etc.

2.2.1 History of Wearable Computing

Today, an extraordinary variety of mobile devices allow people to access to the information at any time and in any place. It is unquestionable that the emergence of mobile interaction has greatly changed people's life and work. Off-the-shelf handheld mobile phones, pads and tablets have become indispensable interaction tools penetrating into the aspects of people's social activities, communications, entertainment and other areas of life. Nevertheless, the issues of seamless integrated interaction are far from being solved with these handheld devices. Firstly, compared with wearable computing, they require users' focused attention and both their hands. Secondly, it is difficult to superimpose digital information upon the physical environment in a seamless augmented way via handheld devices. Finally, limitation of miniaturization of the traditional small screen display will decrease usability and utility if large amount of information is presented on such a small-size mobile screen. However, wearable configurations and one-hand gestures as input can free at least one hand as compared with the two hands required for handheld devices. In addition, wearable computing intersected with augmented reality (Milgram & Kishino, 1994) is capable of overlaying information upon what the user actually sees. Emergence of pico-projectors provides a richer user experience than the traditional small screen. Thus, given such advantages, there is no doubt that interaction on the basis of wearable computing can play an important role as that played by handheld mobile phones, pads and tablets today.

In this subsection, we outline a brief history of wearable computing in the early stage over the last two decades. The term "ubiquitous computing" was introduced by Mark Weiser (Weiser, 1991) in the paper published in 1991, which focuses on integration of technologies into daily life with the aim of binding the user, environment and technologies as one. Ubiquitous computing eliminates the utilization restriction obliging the users to access to the IT system only with fixed or portable computers and their classical graphical user interfaces (GUIs), with WIMP style and devices (e.g., screen, keyboard and mouse). Wearable computing is an alternative approach to ubiquitous computing, allowing the user to interact with body-worn computers, affording the user seamlessly immersed in the physical world with digital information. In 1993, Thad Starner (Starner et al., 1995) , one of the wearable computing pioneers from the MIT Media Laboratory, had attempted a heads-up display integrated with his glasses and a Twiddler (Lyons et al., 2004; Lyons, Starner, & Gane, 2006) as the input device which can be located in the pocket.

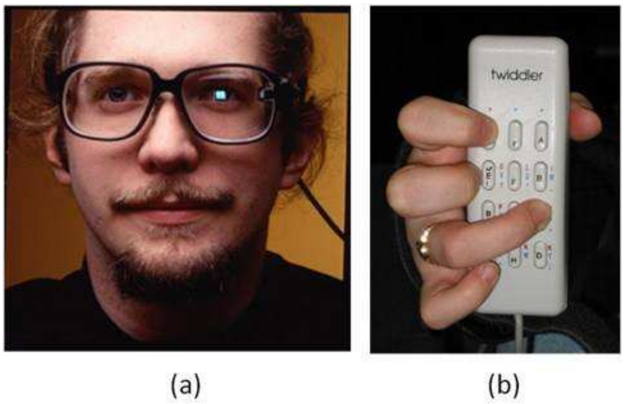


Figure 2.1 Pioneer wearable configurations.
(a). Prof. Thad Starner wearing head-worn glasses.
(b). Twiddler, a chorded keypad used for one-handed typing in wearable systems (Image from (Krumm, 2009)).

Since the mid 1990s, wearable computing has been studied by pioneer researchers, with respect to applications for maintenance engineering and military operations. The majority aimed at providing just-in-time information via wearable configurations or obviously tended to such an aim. Another important system is Remembrance Agent (Rhodes, 1997), which could act as a memory aid. This system could recommend relevant files from a database, based on whatever notes were currently written on a wearable computer. Steve Feiner and his colleagues developed KARMA (Feiner, Macintyre, & Seligmann, 1993), a knowledge-based augmented reality system for aiding maintenance. The user could wear a private eye display over one eye, giving an overlay effect when the real world was viewed with both eyes open. The system overlays the maintenance instructions on top of whatever was being repaired and used sensors attached to real objects to determine their locations. The Forget-Me-Not (Lamming & Flynn, 1994) was a wearable device which could record the interaction with people and devices and store the information of the interaction in a database for subsequent use. A key characteristic of the wearable configuration in the early stage is the heads-up display (Matsumoto et al., 1993) embedded into or combined with the user's glasses as the output, and a customized keyboard like the Twiddler as the input. Instead of this wearable configuration, a wrist computer with half-QWERTY keyboard by Edgar Matias (Matias, MacKenzie, & Buxton, 1994) was built for mobile interaction.

With the progress made in on-body sensors, novel textiles, the miniaturization of mobile devices, the powerful capacities of portable computing devices, and other mobile technologies, wearable computing has been able to provide a more natural and intuitive way to interact in recent years. Many facets of wearable interactions and innovative modalities have been studied. In terms of output, Ni and Baudisch investigated spatial interaction using the hand gesture as the input and the zero visual feedback as the output in Disappearing mobile devices (Ni & Baudisch, 2009) (see Figure 2.2). They studied the limits of miniaturization of mobile devices, what the smallest future devices might be, as well as how the user would interaction with these smallest devices.

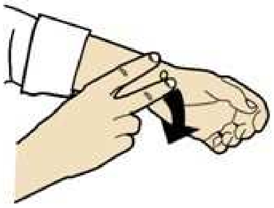


Figure 2.2 The user is entering a "2" by scanning two fingers with the disappearing mobile devices mounted on the wrist.
(Image from (Ni & Baudisch, 2009))

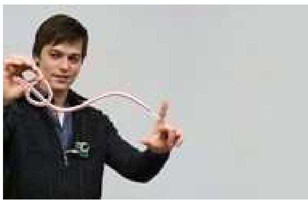


Figure 2.3 The user is sketching a stock curve using an imaginary interface.
(Image from (Gustafson, Bierwirth, & Baudisch, 2010)).

Figure 2.4 The user interacts with his imaginary phone.
(Image from (Gustafson, Holz, & Baudisch, 2011)).

In addition, Imaginary interfaces (Gustafson et al., 2010) (see Figure 2.3) and Imaginary phone (Gustafson et al., 2011) (see Figure 2.4) are also dedicated to the imaginary visual feedback, meaning that all visual feedback only takes place in the user's imagination. This research work originates from wearable computing and visual memory. Imaginary interface uses an optical tracking cameras (fixed in the environment) and gloves with markers as the input devices. The main advantage of Imaginary Interface is that it provides an alternative solution for 2D spatial interaction, allowing ultimate miniaturization. Imaginary phone prototype allows users to perform daily tasks by letting them transfer spatial memory from a real familiar device such as an iPhone. The action of pointing is tracked by the depth camera, following which the touch events are sent wirelessly to the physical iPhone.

HoverFlow (Kratz & Rohs, 2009) seeks to expand the 3D input space of mobile and wearable devices by allowing these small devices to recognize several movement-based gestures.

2.2.2 Wearable Computing Applications and Scenarios

One of the main applications of wearable computing is maintenance. Early on 1998, Siewiorek et al. (Siewiorek et al., 1998) presented a train maintenance and diagnosis system, using mobile information and communication technology to assist with maintenance tasks. In the study by Baudhuin (Baudhuin, 1996), a wearable computer combined with interactive electronic manuals is provided to support military maintenance work. Later on 2006, Nicolai et al. (Nicolai, Sindt, Witt, Reimerdes, & Kenn, 2006) described an approach to shorten the process for aircraft maintenance through a combination of wearable computer and knowledge management technology. WearIT@work (Lukowicz, Timm-Giel, Lawo, & Herzog, 2007) is a project financed by the European Union, aiming at facilitating real-life industrial deployment of wearable technology. This project focuses on wearable applications on aircraft maintenance, car production, healthcare, and emergency response.

Besides military and aircraft maintenance applications, wearable computing is also implemented in the educational sector via various wearable sensors. The research (Ngai, Chan, Cheung, & Lau, 2010) presents a platform with the aim of facilitating integration of wearable computing into computer science and engineering education. The health-based applications emerging in the area of wearable computing have become a research topic. Sung et al. (Sung, DeVaul, Jimenez, Gips, & Pentland, 2004) describe a wearable real-time shiver monitor based on a flexible distributed mobile system, which monitors the body temperature of soldiers to detect hypothermia.

2.2.3 New Advances in Wearable Computing

Four main layers of future wearable systems are envisioned in (Amft & Lukowicz, 2009) in 2009 including:

- Mobile phone-like device as a central on-body platform for general purpose computing tasks;
- Carry-on peripherals such as headsets, displays, and textile touch pads;
- Microsensors deeply embedded in accessories, such as rings, shoes, belts, etc., or encapsulated in clothing;
- Sensing, communication, and power generation infrastructure implemented in textile technology.

These layers need to interoperate seamlessly and allow automatic transitions between interfaces and sensing setups as the user changes clothing. In addition, ISWC (International Symposium on Wearable Computers) is one of the major conferences on wearable computing. From (Smailagic & Kenn, 2011), we can note that the technical program of the 15th ISWC focused on wearable context and activity recognition, research using cell phones, challenges of Human-Computer Interaction based on wearable and novel sensing modalities, electronic textiles, and wearable applications.

2.3 Vision-based Hand Input Techniques

In this section, we first present our taxonomy of hand gesture input devices and technologies based on the naturality. We then focus on research work in the aspects of vision-based hand-gesture interaction, including an introduction to hand postures and hand gestures, recognition and interaction issues within gestures, design of gestures, and pointing techniques in a spatial environment. We finally discuss the requirements of vision-based hand input techniques and their applications.

2.3.1 Hand Gesture Recognition Technology

Utilization of hands and fingers as input techniques has long been studied by researchers. Voice and gesture interactions are regarded as “natural interaction”, as described in Myron Krueger’s pioneering 1991 book “Artificial Reality” (Krueger, 1991). We classify input devices according to the naturality of hand interaction. As figure 2.5 shows, along the input axis, devices vary from the low naturality of the hand gadget; marked hand to the high naturality of the bare hand input. We define the hand gadget as handheld devices such as Wii remote (Lee, 2008) or hand-wearable devices such as gloves. In this case, the marked hand is regarded as the hand marked with colored stickers, with a less intrusive action compared with hand gadget devices. Early research into hand input mainly uses digital gloves (Grimes, 1983) or other hand gadgets such as the Wii remote to aid efficient recognition of hand gestures and postures. Recently, research into computer vision-based hand tracking has gained support from colored stickers or markers, as well as colored gloves (Wang & Popović, 2009) to detect hand gestures or have directly recognized bare hand gestures. Recognition of colored gloves or marked hands simplifies and facilitates image processing. Compared with interaction of digital gloves and other detecting sensors, interaction of vision-based hand gestures can provide an economical method, as well as ensure a natural experience for users. OmniTouch (Harrison, Benko, & Wilson, 2011) provides the touch-like pointing gesture as input via a depth-sensing camera, while SixSense (Mistry, Maes, & Chang, 2009) explores and proposes the marked fingers’ gestures as input. It focuses on gestures including those supported by multi-touch systems, freehand gestures and iconic gestures. It supports dynamical navigation such as zoom-in, zoom-out and pan, but does not really discuss how to navigate traditional selection menus. Brainy hand (Tamaki, Miyaki, & Rekimoto, 2009) includes a single color camera to detect hand gestures.

Figure 2.5 Taxonomy of hand input techniques according to naturality.

2.3.2 Vision-based Hand-gesture Interaction

The early glove-based sensing technology (Kessler, Walker, & Hodges, 1995; Kuroda, Tabata, Goto, Ikuta, & Murakami, 2004) has several drawbacks. First, it lacks the ease and naturalness required to interact; it fails to satisfy the requirement of “come as you are” in (Wachs, Kölsch, Stern, & Edan, 2011). Second, this hand gesture sensing technology requires over long calibration and setup procedures. Research into vision-based hand-gesture interaction has attracted more interest and become prevalent in recent years, as computer vision technology has the potential to provide a natural, unencumbered, and non-contact solution for Human-Computer Interaction (HCI). Two types of hand gestures have been investigated: static gestures and dynamic gestures. The former is based on posture patterns or motion patterns, while the latter is based on estimation of real 3D hand motions. Since human hands have characteristics such as a uniformly colored surface, proximity of limbs, and a concave shape, it is difficult to recognize and interpret the motion of hands with a single recognition method outside the laboratory environment. Furthermore, the main issues encountered in the design of hand pose estimation systems include the high-dimensional problem, self-occlusions, uncontrolled environments, and rapid hand motion (Erol, Bebis, Nicolescu, Boyle, & Twombly, 2007). Although the depth-sensing camera could address some issues and has been already considered in HCI, its cumbersome size still restricts its ability to be worn.

Besides borrowing the similar affordances of the gestures such as those provided by multi-touch (Elias, Westerman, & Haggerty, 2010; Roudaut, Lecolinet, & Guiard, 2009) , the design of in-the-air gestures (T. Ha & Woo, 2006; Von Hardenberg & Bérard, 2001) is usually created by experts and designers in HCI. For instance, the gestures of pen-up, pendown, frame, Namaste, etc. in WUW (Mistry et al., 2009) are all designed by researchers. The pinch gestures in (Loclair, Gustafson, & Baudisch, 2010) by Loclair and the touch pointing gesture in (Harrison et al., 2011) by Harrison are also created by researchers themselves. The manipulation technique in (Kolsch, Turk, & Hollerer, 2004) is also designed by HCI experts, including the pointer based interaction, registered technique, location-independent postures, etc. Instead of the expert design method, the end-user elicitation method could also be considered in spatial interaction as it has been considered for surface gestures (M. R. Morris, Wobbrock, & Wilson, 2010). The results of this study indicate the importance of incorporating consensus, by end-users or groups of designers, in the creation of surface gestures, and offer evidence that HCI researchers may not always create optimal gesture designs despite their expertise. These results are also valuable for gesture design of the wearable interaction.

Selection and validation of the selection is an essential action for GUIs. One of the solutions to achieve selection and validation is pointing. With respect to pointing in the spatial environment, the researcher uses hover time to validate a selection in the wearable camera-projector system (Mistry et al., 2009). The experiment in (Müller-Tomfelde, 2007) provides an empirical evidence for a possible natural dwell time (Dwell-time, namely hover time, is a certain period of time in which the user remains motionless while pointing the target object.) to select the target object. It concentrates on three-dimensional pointing interaction using hand or tool movements. This study focuses on dwell-time in a more general environment. The results recommend a feedback delay time for manual pointing actions of approximately 350 to 600 ms as a starting point for interactive application development.

The fat finger problem (Siek, Rogers, & Connelly, 2005) exists widely in the direct-touch finger input, which impacts validation and other input actions. This problem leads to two issues (Benko & Wigdor, 2010): the fat finger occludes the target (the occlusion problem) and the touch area of the finger is much larger than a pixel of the display (the precision problem). One approach is the offset cursor technique in (Potter, Weldon, & Shneiderman, 1988) proposed by Potter et al. Another solution is the Shift technique proposed by Vogel and Baudisch (Vogel & Baudisch, 2007), which offsets the area beneath the finger. Other techniques include the complex offsetting cursor technique (Albinsson & Zhai, 2003), the touch cursor technique (Wigdor, Forlines, Baudisch, Barnwell, & Shen, 2007), the Dual Finger Midpoint (Esenther & Ryall, 2006), etc.

2.3.3 Applications and Requirements of Design

Since the well-defined vision-based hand-gesture interaction could provide an intuitive, non-contact, and unencumbered user experience, applications based on computer vision technologies include medical systems and assistive technologies (Gratzel, Fong, Grange, & Baur, 2004; Wachs et al., 2008), entertainment (Höysniemi, Hämäläinen, Turkki, & Rouvi, 2005), as well as the Human-Robot Interaction (Nickel & Stiefelhagen, 2007).

Wachs et al. (Wachs et al., 2011) have discussed the basic requirements for gesture interfaces: price, responsiveness, user adaptability and feedback, learnability, accuracy, low mental load, intuitiveness, comfort, lexicon size and multi-hand systems, "come as you are", reconfigurability, interaction space, gesture spotting and the immersion syndrome, as well as ubiquity and wearability.

2.4 Visual Output Techniques

In this section, we first present the classic head-worn display together with other head attached displays. We then discuss personal projectors, including miniaturization of projection, interaction on the basis of the personal projector, taxonomy of mobile output devices along with scalability and scalable interaction, as well as social projection issues.

2.4.1 Head-worn Display (HMD)

Compared with output modalities of haptic feedback and audio feedback, the visual output provides more information to display and interact. The visual output, as the primary output mechanism, can also inherit the interactive elements of GUIs. However, compared with the fixed desktop interaction, the mobile interaction is more demanding on the complicated physical environment. The light condition is a main issue affecting not only the accuracy of hand input recognition, but also the efficiency of interaction when using the pico-projector. Also, for the projected interface, the user is more easily distracted in the state of mobility. More importantly, it requires careful design to keep a balance between protection of privacy and sufficient display experience. In light of the above, HCI researchers are working on the visual output mechanism in many-sided aspects with the aim of providing a comfortable user experience with the ubiquitous computing.

As a feedback supporter, miniaturized displays play an important role in the field of wearable computing. Researchers working on mobile interaction expect displays to be light, easy to wear, able to display multimedia information, and simultaneously support a presentation size as large as possible. As a wearable output visual display, head-worn displays (Y. Ha & Rolland, 2002; Spitzer, Rensing, McClelland, & Aquilino, 1997) have been used to present information in the early research stage of wearable computing. With these head-worn output devices, the user can obtain visual feedback and other people only see the transparent glasses without accessing the digital information. Cakmakci and Rolland (Cakmakci & Rolland, 2006) survey the current research on head-worn display technology and provide a comprehensive overview on design and development of HWDs. In general, a combination of real and virtual objects is dependent on two augmentation methods: overlaying virtual information on real objects along the space and time dimension, and offering well-timed information along the time dimension.

Figure 2.6 Head-worn displays.
(a). Lumus DK-32 wearable see-through display.
(Image from http://www.geeky-gadgets.com).
(b). Google Glass.
(Image from https://plus.google.com/111626127367496192147/posts).
(c). Micro Optical SV-6 PC viewer.
(Image from http://www.inition.co.uk/).
(d). Micro Optical EG-7 display (Salminen, 2001).
(Image from www.lumus-optical.com).

No matter which technology the HWD uses, these HWDs can be classified in two types: see-through displays and non-see-through displays according to the two augmentation approaches as we stated above. The see-through technology allows presentation of information on the transparent glasses that does not obstruct the wearer's view. In the meantime, virtual information can be superimposed on the real object via calculated augmentation. One solution of the non-see-through display is the goggle attached with a small screen, which can be clipped onto any side of the goggle or a pair of eyeglasses. Since the small screen blocks some of the user's view, this device is more suited for providing in-time information rather than for overlaying information on the real object. We list the examples of the see-through and the non-see-through displays belonging to the HWDs in the figure: the Lumus DK-32 wearable see-through display (see Figure 2.6 (a)), the Google Glass (see Figure 2.6 (b)), the Micro Optical SV-6 PC viewer (see Figure 2.6 (c)), and the Micro Optical EG-7 display (see Figure 2.6 (d). The first two are see-through displays, while the last two are non-see-through displays.

Besides the HWDs, the virtual retinal display (VRD) (Pryor, Furness, & Viirre, 1998) and the head-mounted projective display (HMPD) (Hua, Gao, Biocca, & Rolland, 2001; Rolland & Fuchs, 2000) (see Figure 2.7) have also been used as wearable configurations. The former reflects the image directly on the retina, while the latter projects the image on the target objects coated with retroreflective material.

Figure 2.7 Head-mounted projective display (Azuma et al., 2001).

These miniaturization devices normally use fixed-size screens or physical materials to present visual information. Two of the advantages of the small screen are that they can protect excellent user privacy with a small-size reading space, and that they allow high-level mobility. Also, they do not require extra physical surface to aid the display action. However, a drawback persist, namely these displays cannot avoid the limitation due to the small-size screen, in which visual output information content is restricted in a scale.

2.4.2 The Personal Projector

2.4.21 Miniaturization of Projection

The projector is the display device used to present visual images as well as project graphical user interfaces, with the property of scalability compared with the traditional screen. In recent years, miniaturization of projectors has led to the emergence of mobile devices with embedded projector or palm-size projectors. Projector components are starting to be embedded into household digital cameras (see Figure 2.8) or mobile phones (see Figure 2.9). Besides its role as an auxiliary accessory, the pico-projector as an independent device has the ability to connect with other devices and to project high quality images. Moreover, pico-projectors are small enough to be worn on the body, held in the hand or put into the pocket, which is ideal for mobility and content sharing (see Figure 2.10).

Figure 2.8 Camera with embedded projector. (2011) (Image from http://www.gadgetguy.com.au).

Figure 2.9 Mobile phone with embedded projector. (2007) (Image from http://news.tigerdirect.com).

Figure 2.10 Pico-projector size. (2009)
(Image from http://www.gadgetsarefun.com).

For instance, the pico-projector Acer C120 (see Figure 2.11) has a resolution of 1280×800 maximum. Projector size is approximately 2.54cm×11.94cm×8.13cm, with a weight of 180g. In standard mode the brightness can attain 100 lumens, while in low mode the brightness maintains 75 lumens, able to support indoor interaction.

Figure 2.11 Acer C120. (2011)
(Image from http://us.acer.com).

2.4.22 Personal Projector Interaction

To facilitate interaction with dynamic information, we also attempted to integrate utilization of a wearable projected interface based on a pico-projector. The emergence of pico-projectors and the development of ubiquitous computing have led to new-look interaction methods. Four conceptually distinct approaches for interacting with the personal pico-projector system have been identified (Rukzio, Holleis, & Gellersen, 2012) as follows: input on the pico-projector (see Figure 2.12 (a)), movement of the pico-projector (see Figure 2.12 (b)), direct interaction with the projection (see Figure 2.12 (c)) and manipulation of the projection surface (see Figure 2.12

(d)). The input approach on the projector employs mobile gadgets such as input interfaces, a two-button interface which is navigated by the user's thumb (Cao, Forlines, & Balakrishnan, 2007). The advantage of this approach is that it maintains the user's focus on the projection. Since the user cannot avoid switching the attention between the projection and the input controller, the efficiency of the interaction decreases (Greaves & Rukzio, 2008). Many researchers have investigated the movement of the projector as the input approach, which is dependent on sensing the location, orientation and other movements of the projector. The approach based on direct interaction with projected content leverages the projected interface as the interactive surface, as well as employing hand gestures, digital pens or other methods as the input technique. In this way, the projected image has a metaphor similar to the touch-screen. The approach of manipulation on the projection surface achieves input through changing the features of the projected interface, such as position, orientation, shape, etc. On-body interaction (Harrison, Ramamurthy, & Hudson, 2012) manipulates the projected interface on arms or hands via the movement of the arms or hands.

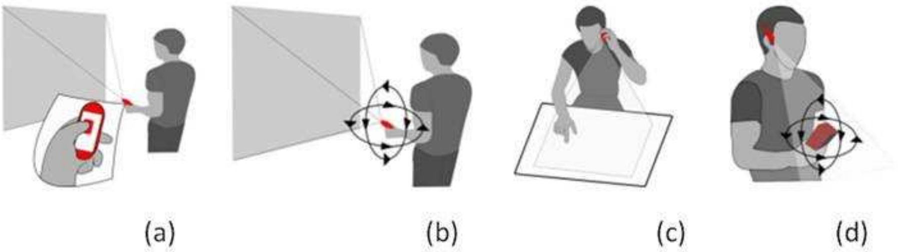


Figure 2.12 Concepts for interaction and control with a personal projector (Rukzio et al., 2012).
(a). Input on the pico-projector.
(b). Movement of the pico-projector.
(c). Direct interaction with projection.
(d). Manipulation of the projection surface.

We mainly focus our study on direct interaction with the projection in a mobile environment on daily surfaces. Kurata et al. present the BOWL ProCam (Kurata, Sakata13, Kourogi, Okuma, & Ohta, 2008) that proposes interaction techniques effectively employing both nearby projection surfaces such as the user's hands and far projection surfaces such as a tabletop and a wall. They focus on the technique for understanding where the nearby-and-far-away surfaces are located, rather than the experience of users' mobile projection. SixthSense (Mistry et al., 2009) (see Figure 2.13

(d)) is a wearable camera-projector system supporting the gestural interface with the marked finger gestures input. This project proposes superimposing the projected information onto surfaces in the real environment. Interactive Dirt (McFarlane & Wilder, 2009) is also a worn camera-projector system focusing on mobile team collaboration for military purpose. Brainy Hand (Tamaki et al., 2009) is an ear-worn hand gesture interaction device supporting the mini pico-projector display, light-weighted and practical to wear while the user is eating, talking and sleeping. The Skinput (Harrison, Tan, & Morris, 2010) technology uses the user's body as an interactive surface like the touch pad with the bio-acoustic sensors and pico-projector, providing an always-available interface (D. Morris, 2010). Skinput (see Figure 2.13 (c)) combines simple bio-acoustic sensors and some sophisticated machine learning to enable people to use their fingers or forearms as touch pads. However, the large display property of the pico-projector is limited with Skinput as it can only employ the body surface. With the goal of employing everyday surfaces as the interface, OmniTouch (Harrison et al., 2011) (see Figure 2.13 (a) (b)) allows the user to wear the depth camera and pico-projector on the shoulder to support interactive multitouch applications. It can track surfaces even if the surface has been moved. In this way, OmniTouch is able to project information on arbitrary surfaces.

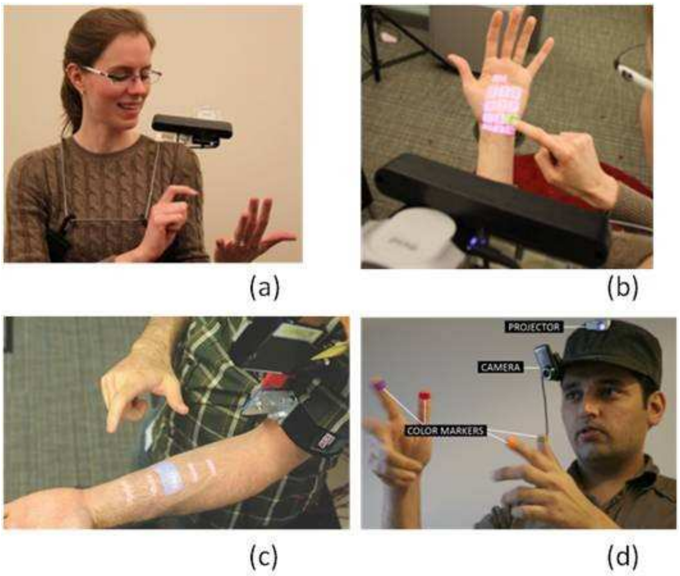


Figure 2.13 Camera-projector interaction systems.

(a). Configurations of OmniTouch (Harrison et al., 2011).

(b). Projected interface of OmniTouch (Harrison et al., 2011).

(c). Projected interface of Skinput (Harrison et al., 2010).

(d). Configurations of SixthSense (Mistry et al., 2009).

For wearable projection, pico-projector stability and projected image viewability should be considered during interaction design. The appropriate position of the wearable projector, the projected size, and the projected location are investigated and evaluated in the work of (Ota, Takegawa, Terada, & Tsukamoto, 2010), and would vary according to the different situations, contexts and projected contents. Konishi et al. (Konishi, Tajimi, Sakata, & Nishida, 2009) propose a method to stabilize projection from the shoulder or the chest to the palm in mobile settings. A hip-mounted projector for floor projection has been explored in (Tajimi, Uemura, Kajiwara, Sakata, & Nishida, 2010) by Tajimi et al.

One of the projection research topics is the study of displaying anywhere, which aims at utilizing arbitrary surfaces as projection surfaces in everyday life. Surfaces are usually slightly curved, with different textures and colors, especially in irregular shapes. These imperfect surfaces are detected (Harrison et al., 2011) to make the projected interface adapt to object surfaces. A preliminary study as a pioneer work on whether and how the projected interface is impacted by real objects and surfaces has been investigated in (Podlaseck, Pinhanez, Alvarado, Chan, & Dejesus, 2003). The iLamps (Raskar et al., 2003) technique focuses on enhanced adaptive projection on nonplanar surfaces using conformal texture mapping. Projection on arbitrary surfaces has been discussed in (Huber, Liao, Steimle, & Liu, 2011) including planar surfaces and non-planar surfaces. Another important topic is augmentation of real objects with projected information, which has some intersected part with the topic of anywhere display. Projection-based interaction can support augmentation directly located on real objects (Beardsley, Van Baar, Raskar, & Forlines, 2005; Raskar et al., 2004; Schöning, Löchtefeld, Rohs, & Krüger, 2010). Distinctive visual markers are used on the display surface in (Beardsley, Forlines, Raskar, & VanBaar, 2005) to define a coordinate frame for image stabilization while augmenting digital projected information on physical textures. In this way, electronic data can be attached to the physical object. The projection-based system AnatOnMe (Ni, Karlson, & Wigdor, 2011) projects medical imagery on patient’s injured body to facilitate medical information exchange. This augmentation is achieved by a pico-projector, webcam, near-IR camera, and modified wireless presenter control. The FACT (Liao, Tang, Liu, Chiu, & Chen, 2010) system allows the user to interact with augmented paper documents through the fine-grained physical-digital interaction mapping approach. A content based approach is used to establish homographic transformation.

Based on the survey (Rukzio et al., 2012), four areas of application of projection-based interaction have been identified: games and entertainment (Cao et al., 2007; Cao, Massimi, & Balakrishnan, 2008), augmented reality (Hosoi, Dao, Mori, & Sugimoto, 2007), data visualization and manipulation (Blask, Coriand, & Feiner, 2005), and group collaboration (Greaves & Rukzio, 2009).

2.4.23 Property of Scalability

We classify current mobile output devices according to scalability of output visual display size (see Figure 2.14). And, along the axis, devices alter from the low scalability of mobile devices with small screen, medium scalability of mobile phones combined with projector, to the high scalability of pico-projectors. For low scalability devices, the display is dependent on the physical property itself, the size of which is unchangeable. For high scalability devices, output size can be changed in a flexible way; projection on the palm leads to a small-size output while projection on the wall leads to a large-size output. The display relies on the projected surface. For medium scalability devices, the size of the visual output region can be changed to some extent, such as the output size of the mobile phone with projector which varies according to the context switch (Greaves & Rukzio, 2008; Hang, Rukzio, & Greaves, 2008).

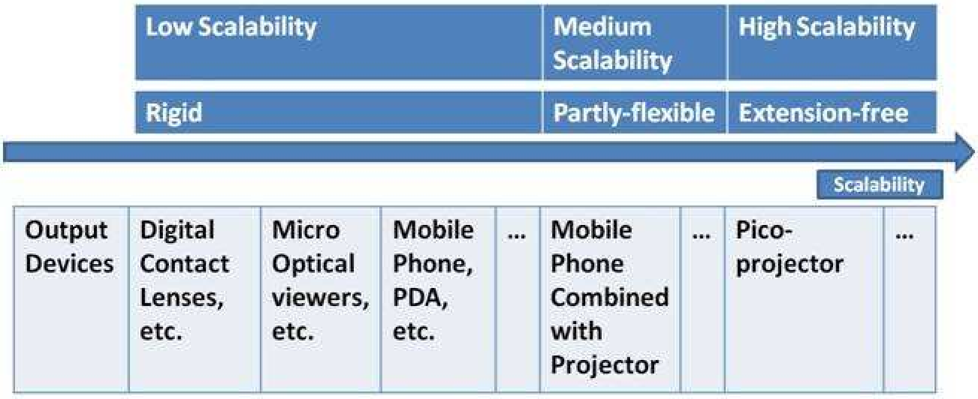


Figure 2.14 The taxonomy of mobile output techniques according to the scalability.

The mobile phone combined with the projector (see Figure 2.15) is a possible solution for achieving medium scalability of the visual output. The mobile phone and the projector are combined to display photos. Also, input capabilities are provided by the mobile phone.

Figure 2.15 The mobile phone display and projection (Greaves & Rukzio, 2008).

The pico-projector has high full scalability and supports scalable interaction. Although these systems mentioned earlier in the last subsection employ the pico-projector as the output, the scalability property of the pico-projector has not been studied, nor the layout design of the scalable interface. Proximity sensing technologies (Ballendat, Marquardt, & Greenberg, 2010) pave the way for scalable interaction. The knowledge of nearby people and other devices – their position, identity, movement, and orientation – can be investigated to design interaction techniques. We focus on integration of wearable projection and proxemic interaction to address the issues of the scalable interface.

2.4.24 Social Issues on Projection

Ju-Chun Ko et al. (Ko, Chan, & Hung, 2010) explore the rights for people to project and be projected in public spaces providing some possible solutions. The issues of applying these projected user interface (PUI) techniques in real life have been discussed. Another research path (M. L. Wilson, Robinson, Craggs, Brimble, & Jones, 2010) explores how people will want to use the projector technology, how they will feel when using it, and what social effects the researcher can expect to see. Results from this investigation showed that users are willing to project content, even when in social spaces and with other people around. One contribution indicated that projector phones should support careful control over projected content so that users can maintain privacy easily. A formative field study in (Greaves, Akerman, Rukzio, Cheverst, & Hakkila, 2009) has been explored to investigate users' reaction on public projection. The results indicated that personal projection attracts a large amount of attention, is dependent on the social context, and has been accepted socially.

2.5 Related Research Areas

In this section, we briefly summarize the research areas impacting our wearable innovative user interfaces. Firstly, we describe augmented reality, mobile AR and their applications. We then discuss ubiquitous computing and always-available mobile interaction. Finally, we survey the tangible user interface and marker-based interaction.

2.5.1 Augmented Reality (AR) and Mobile AR

Wearable computing allows users to act in mobility and in the context related to real environment (Starner et al., 1997). The real environment can be augmented consciously to support the relationship between real and virtual (digital) worlds. Wellner (Wellner, Mackay, & Gold, 1993) and Milgram (Milgram et al., 1995) contribute the pioneer work in the field of augmented reality. Milgram defined a continuum of real-to-virtual environments, in which AR is one part of the general area of mixed reality. AR systems integrate virtual and digital information into the physical environment, allowing the user to perceive information while immersed in the physical environment. Mobile augmented reality systems (MARS) achieve the goal of augmentation without the restriction of a fixed location. Also, mobile AR allows the user to interact with well-timed information without diverting attention of the user. The survey (Höllerer & Feiner, 2004) summarizes the application areas for which mobile AR prototypes have been investigated, including the areas of assembly and construction (Mizell, 2001), maintenance and inspection (Klinker et al., 2001), navigation and path finding (Furmanski, Azuma, & Daily, 2002), tourism (Cheverst, Davies, Mitchell, Friday, & Efstratiou, 2000), geographical field work (Nusser, Miller, Clarke, & Goodchild, 2003), journalism (Höllerer, Feiner, & Pavlik, 1999), architecture and archaeology (Vlahakis et al., 2002), urban modeling (Baillot, Brown, & Julier, 2001), entertainment (Starner, Leibe, Singletary, & Pair, 2000), medicine (Fuchs et al., 1998), military training and combat (Livingston et al., 2011), personal information management and marketing (Zhang, Navab, & Liou, 2000).

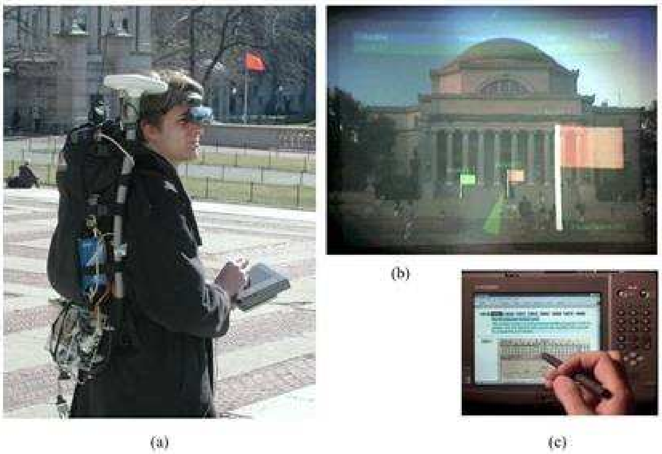

(a)
(b)
(c)

Figure 2.16 Typical wearable configurations (Höllerer & Feiner, 2004).
(a). User with Mobile AR system.
(b). View through HWD.
(c). Additional handheld interface.

To achieve mobility, mobile AR is working on wearable configurations or handheld devices such as mobile phones, PDAs or tablets (see Figure 2.16 (c)). Most of the existing wearable augmented reality systems are based on the head mounted displays (see Figure 2.16 (a)), wearable video cameras and eye/head-trackers (see Figure 2.16 (b)), input and sensing devices, and wearable PCs (Aleksy, Rissanen, Maczey, & Dix, 2011).

2.5.2 Ubiquitous Computing and Always-Available Mobile Interaction

Ubiquitous computing and pervasive systems are important evolutions of information technology (IT) allowing new utilization in everyday life. The term "ubiquitous computing" was introduced by Mark Weiser (Weiser, 1991), and focuses on the integration of technologies into daily life with the aim of binding the user, environment and technologies as one. The goal of ubiquitous computing is to eliminate the utilization restriction obliging users to access the IT system only with fixed or portable computers and their classical GUIs, with WIMP mode and devices (e.g., screen, keyboard and mouse).

Although ubiquitous computing covers a large number of aspects, we shall only address always-available mobile interaction in this subsection. Dan Morris and his colleagues (D. Morris, 2010) survey the properties of sensors and input systems that may enable a shift from traditional desktop computers to always-available computing. Several requirements for always-available mobile input have been outlined such as:

- It requires a cognitive shift to the task for which the user demands input, but the input should not disrupt cognition.
- It requires the in and out transition of always-available input to be as rapid as transitioning out visual attention from one task to another.
- It requires the always-available input to be portable to any environment.
- It requires the always-available input to be at least compatible with the use of our hands for non-computer-based tasks.

Always-available input with respect to on-body configurations has an intersection with wearable computing. Always-available input technologies include inertial motion sensing, touch sensing, computer vision, mouth-based interfaces, brain-computer interfaces, muscle-computer interfaces and other emerging sensors. Furthermore, always-available output technologies include haptic feedback, audio feedback and glasses and other mobile displays. The challenges of always-available interaction have also been discussed such as systematically handling ambiguity on input, sensor fusion, gesture design and usability, as well as cognitive interference.

2.5.3 Tangible User Interface (TUI) and Marker-based Interaction

Ishii and Ullmer (Ishii & Ullmer, 1997) have defined the tangible user interface at CHI 1997, the definition of which is to “augment the real physical world by coupling digital information to everyday physical objects and environments”. Even though the terms related to TUIs have varied, they share the same basic paradigm (Fishkin, 2004): users use their hands to manipulate some physical objects via physical gestures, and a computer system detects this, alters its state, and gives feedback accordingly.

Paper interaction is one of the tangible user interfaces (Ishii, 2008). Studies on paper interaction and paper interfaces (Akaoka, Ginn, & Vertegaal, 2010; Holman, Vertegaal, Altosaar, Troje, & Johns, 2005; Mistry & Maes, 2008) focused on augmented reality, attempt to merge the use of paper with digital information and data. Researchers mark the paper with special markers, and then use the camera to recognize and detect both the motion of paper and other input techniques. Paper Windows (Holman et al., 2005) describes a projecting window prototype able to simulate the manipulation of digital paper displays. This system takes the paper motion and finger pointing gestures as the input. The user can thus perform tasks by interacting with paper documents using the fingers, hands and stylus. The Quickies (Mistry & Maes, 2008) system users the augmenting sticky notes as an I/O interface. The DisplayObjects (Akaoka et al., 2010) proposes a workbench allowing the user to interact with projected information on the physical object. Whereas these studies are all investigated the large display interaction or the desktop interaction, we choose to focus on paper interaction in the mobile situation.

In addition to the paper-based interface, tangible objects are employed as tags and reminders, and utilized to trigger the digital information. The link between the physical world and the digital world needs to be triggered via explicit interaction such as placing a particular object in the proximity of a reader (Shaer & Hornecker, 2010) or in the target area. RFID, ARToolKit markers, and QR codes are most often used for link tagging. In the context of TUIs, computer vision is often used to sense position of markers, as well as orientation, color, shape, etc. The algorithm can interpret the marker pattern to identify markers. In recent years, there have been a large number and variety of marker-based interactions (Hornecker & Psik, 2005; Rekimoto & Ayatsuka, 2000; Rouillard, 2008) that have made it possible to use contextual markers in a mobile environment. The CyberCode system (Rekimoto & Ayatsuka, 2000) is a tagging system designed for augmented reality, based on CyberCode. The PerZoovasive (Rouillard & Laroussi, 2008) learning environment is an adaptive and context-aware system providing the support for learners in mobile pervasive situations by using QR codes (Quick Response codes). In addition, compared with other detection technologies such as RFID (Kubicki, Lepreux, & Kolski, 2012), the ARToolKit tags (see Figure 2.17) (or QR code) is based on the vision-based interaction, easy to stabilize in the environment, and less expensive. Our approach is inspired by these contextual markers, which can bridge the digital world and the real world in a light and economical way.

Figure 2.17 ARToolKit Markers and interaction. (2003)
(Image from http://www.hitl.washington.edu/artoolkit/).

2.6 Summary

In this chapter we have discussed the basic technologies and relevant research for our study. In the first section, we briefly review the history of wearable computing, the related wearable applications, and the new advances. We then describe the relevant hand input techniques in the second section. In the third section, we discuss the output techniques on the basis of the HWD and pico-projector. Finally, we summarize other related research areas. This chapter paves the way for understanding the following chapters. We will describe our research in chapters 4, 5, 6, and 7. The conclusion will be stated in the last chapter, namely chapter 8.

3 Innovative User Interfaces in Augmented Environment

3.1 Introduction

To help the user interact all around and access to information freely in the environment, we have proposed three interfaces (Zhou et al., 2011): In-environment Interface (IEI), Environment Dependent Interface (EDI) and Environment Independent Interface (EII). With the IEI, the user is in the nomadic state, i.e. without any personal IT device. The environment provides all the interaction support required for input and output devices. The EDI and EII are both based on the user's wearable computer devices, allowing the user to interact in mobility. In this chapter, we are concerned mainly with the light, wearable and cheap user interfaces. We first provide an overview of IEI, EDI and EII and then explain our concepts of EDI and EII, their background and their distinct characteristics. Next, by proposing a series of innovative interfaces based on the webcam capture, we present our approach for concretizing the EDI and EII; we first present the in-environment and fixed interface, we then discuss a series of innovative wearable interfaces based on the webcam capture. With the aim of allowing the user to have at least one hand free in the mobile augmented reality environment in the context, we employ the wearable configurations as follows: a webcam to grasp the input signal; a small screen attached to a goggle to provide visual information, or for a larger field of vision and interface, a pico-projector is used; and a laptop merely for calculating. Our goal is to create a true contextualization which is more effective and adaptive to users' information needs by taking advantage of dynamic and physical environmental characteristics, based on or independent of user's location.

3.2 Overview of Innovative User Interfaces

One of the possible solutions to enable ubiquitous interaction and end the limitation of the desktop mode is nomadism, in the state of which the user is physically mobile and not equipped with any wearable or mobile device. Another possible solution for this problem is mobility, in the state of which the user does not have any classical portable devices such as laptops, but has wearable computing devices, such as the camera-glasses unit or the camera-projector unit. In traditional mobile computing, for example, when the user is moving and wants to use his/her portable laptop, he/she needs to stop before interacting. However, wearable computing can support interaction and mobility seamlessly. The former solution can be achieved by interacting with the IEI, while the latter solution can be achieved by interacting with the EDI and EII.

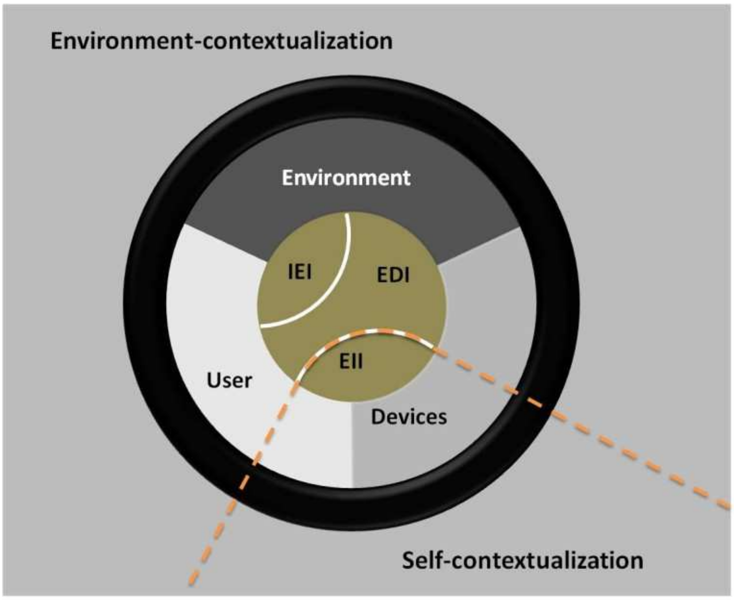


Figure 3.1 An overview of the IEI, EDI and EII, with their elements and contextualization style.

Figure 3.1 represents the relationships among the three interfaces (IEI, EDI and EII), the contextualization provided by these interfaces, as well as three main elements: User, Devices and Environment. In the situation of IEI, the webcam and the wall video projector are appropriately located to allow in-environment interaction. The user uses his/her hands to interact with the public information presented on a public wall like searching and browsing. The environment generates the contextualization, for example the physical location and the application used (i.e. public transportation information). Similarly, the EDI also focuses on the in-environment interaction which is dependent on the in-environment indication and information. As the figure illustrates, both the IEI and EDI rely on the environment, the former requiring the environment and the actual user to support the interaction (The environment provides the devices, and the actual user interacts with his/her hands or body.), and the latter requiring the environment, the wearable devices and the user. Since both the IEI and EDI are dependent on the in-environment information, contextualization can be achieved through the rearrangement of the environment. Furthermore, the EII is independent on the environment, namely, it relies neither on the in-environment information nor on the environment configurations. In this way, users can interact with any digital information by themselves, or for a more sophisticated independent interface, they can interact by showing the webcam the predefined contextualizing indications, which we called self-contextualization as shown in the figure.

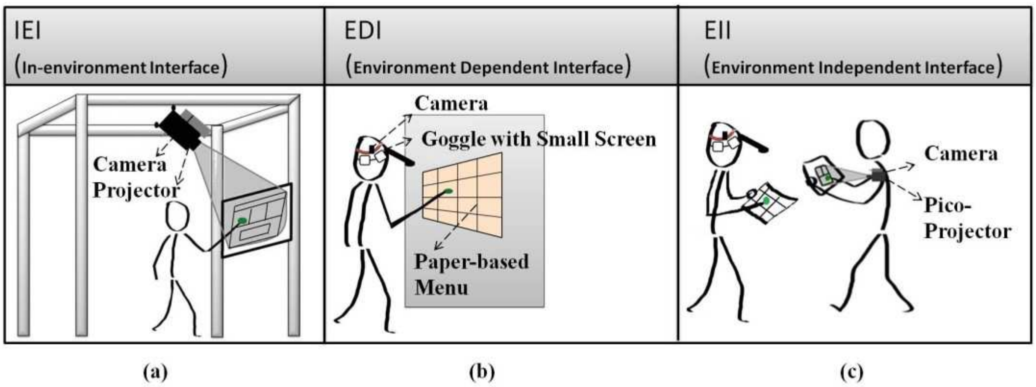


Figure 3.2 Innovative user interfaces.
(a). IEI (In-environment Interface)
(b). EDI (Environment Dependent Interface)
(c). EII (Environment Independent Interface)

In the smart city (David, Zhou, Xu, & Chalon, 2011), the IEI, EDI and EII can solve the same problems that the user encounters, as well as solve distinct problems respectively. For example, the user wants to check whether he/she can have an appointment with a professor in the lab. In the nomadic user scenario with IEI, he/she walks into the public place outside the lab, and browses the public information via hand gestures. As figure 3.2

(a) shows, the public place is configured by a camera and a large projector. In this way, he/she manipulates the professor's schedule information and decides which day is most appropriate for appointment. When the user is equipped with wearable devices, he/she can interact with EDI or EII. In the mobile scenario with EDI, the user comes to the professor's office, but unfortunately, the professor isn't there and the door is closed. So the user interacts with the paper interface posted on the door and makes an appointment with the professor through reviewing the information by a goggle attached with a small screen (see Figure 3.2 (b)). In the mobile user scenario with EII (see Figure 3.2 (c)), the user can project the dynamic information on his/her palm or on one page of the notebook.

3.3 Design of EDI and EII

In this thesis we exclude IEI, and principally focus on the last two interfaces: Environment Dependent Interface and Environment Independent Interface. As we stated above, two other user interfaces (EDI and EII) are based on users' wearable computing devices, allowing them to interact in mobility. The EDI aims to provide users with information determined by the environment, i.e. the environment provides users with information. In other words, the EDI refers to the strong relationship between in-environment information and the interface. The environment can be pre-contextualized by markers, and the markers can be pasted on appliance, wall, book, door, etc. In this way, public and professional guiding information can be used for contextualization. To go one step further, we propose the Environment Independent Interface (EII), which refers to the relationship between the interface and personal information. Contextualization is achieved by the actual user by showing the webcam appropriate contextualized markers or menus, or for a more free way, the user can merely project the personal interface to interact. Thus, the user can contextualize his/her working environment by himself/herself without any contact with the environment.

With respect to configuration, the EDI and EII employ the same configuration, allowing users to switch freely between the EDI and EII and to interact in the context, in the self-context or without any context in the ubiquitous environment.

3.3.1 Environment Dependent Interface

For several years now, we have been working in the research field of augmented reality in relation to mobility for several years. This can be characterized by two acronyms: MOCOCO (MObility, COntextualization and COoperation) and IMERA (David & Chalon, 2007) (French acronym for

Mobile Interaction with Real Augmented Environment). Real environment augmentation can be unconscious or can be conscious passively or actively. Recognition by the IT system of objects, actors or situations of interest without markers falls into the case of unconscious augmentation. The other case is augmentation with use of passive or active markers. In passive marker augmentation, the IT system discovers these passive markers and uses them in the treatment process. In active marker augmentation, active markers (e.g. the active RFID "Radio Frequency Identification" stickers) can address the IT system according to their own decisions. The IT system can, for its part, either be deployed in the environment with its sensors, or be dependent on the user interaction devices which build a unique relationship between the real environment and the IT system.

In this thesis, we are mainly concerned with the latter approach: conscious augmentation using passive markers. For the purpose of providing the use with the in-environment information and interface with the environment-contextualization, we investigate passive marker augmentation which can be achieved by computer vision-based tags and the webcam. Taking the ARToolKit tags as an example, the webcam recognizes the unique pattern of the tag and provides the linked information. In this way, our Environment Dependent Interface is concretized through the passive marker augmentation method. The markers act as the bridges linking the real environment and the digital information. The markers can be pasted on a wall, a notice-board, an information board of a bus shelter, and appliances or a doorplate (Figure 3.3 (a) (b)).

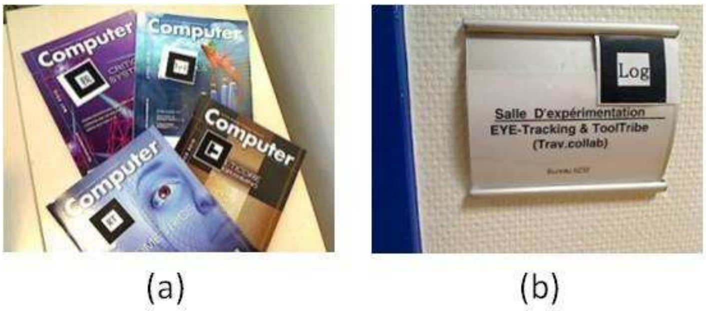


Figure 3.3 Marker-based interactions.

(a). Markers on books.

(b). Markers on the doorplate.

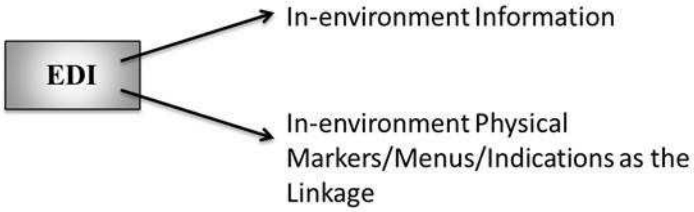


Figure 3.4 The Principal and essential characteristics of EDI.

We also define the distinct characteristics for the EDI. The Environment Dependent Interface must be closely related to in-environment information (see Figure 3.4), which is dependent on the specific location. The location can be identified through either the passive in-environment physical markers or the specific menus or indications which are dependent on the environment. It is impossible to remove the linkage (i.e. the markers) for the EDI. In other words, the linkage is essential in that it is one of the components for building the Environment Dependent Interface.

3.3.2 Environment Independent Interface

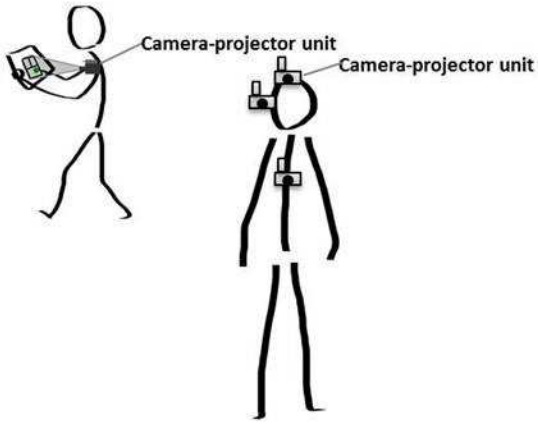


Figure 3.5 The concept and configuration of the EII.

Going one step further, we also explore both marker augmentation and non-marker augmentation to support and concretize our concept of Environment Independent Interface. As shown in figure 3.5, the user with wearable equipment interacts with EII in the mobile setting. Wearable units can be stabilized on different points of the body, and the projected information is summoned via the actual user or the indications in the user's possession.

The user of EII can interact with projected dynamic information on the situation of non-marker augmentation. As regards marker augmentation, in the scenario of reading the augmented newspaper, the user holds the newspaper and navigates the predefined markers or indications to watch the augmented video or multimedia information overlaid on the paper.

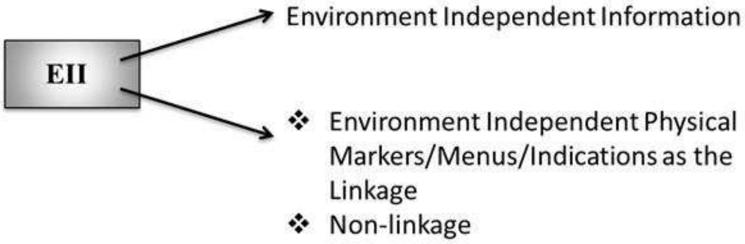


Figure 3.6 Principal and essential characteristics of the EII.

Environment independent information plays an important role in EII (see Figure 3.6). Digital information in EII is summoned with no relation to the environment, and is not dependent on the location. Environment independent markers, menus or indications can be pasted on any handheld surfaces, such as plane tickets, books, newspapers, booklets, or personal notebooks, which are completely independent on the location of the environment. Linkage for the EII is an optional part; non-linkage augmentation can be achieved by pure digital personal projection.

3.4 Light Mobile User Interfaces

We have designed and implemented a series of innovative interface prototypes, allowing the user to interact in-environmentally and beyond-environmentally with at least one hand free. Thus, wearable configurations are mainly based on: a webcam for perceiving the context and user interactions, and a goggle attached small screen for visualizing text, image and video. Otherwise, to allow a larger field of vision and interface, the goggle-attached screen is replaced by a pico-projector for dynamic information projection in the context. A computing device for calculation, as small as possible, is located in the pocket or the backpack. As shown in figure 3.5, the fixed point of the camera and the pico-projector can be fixed together and be stabilized on the head, next to the ear, as well as on the chest. With respect to the camera and glasses group, the camera can be fixed on the forehead. The software for these prototypes was developed on the Microsoft Windows platform using the C, C++, OpenCV, GTK toolkit and ARToolKit.

3.4.1 An In-environment and Fixed Interaction Support

We have distinguished between mobility and nomadism. Mobility uses wearable computing devices, while nomadism does not. In other words, in the latter case, the environment provides interaction support. We thus fixed the webcam with a plastic bracket on the desk of the workplace (Figure 3.7 (a) (b)), or in another situation we fixed the webcam on the wall of a bus shelter. With the help of this configuration based on webcam capture of interaction and large screen or wall video projection, we can establish in-environment interaction and study the feasibility and effectiveness of the camera interaction technique using a finger selection technique on a passive grid. The principle of interaction is based on x-y coordinate detection of finger position. We segmented the whole visual field of the webcam into rectangular zones, and related each zone with a unique event. The user can trigger the required action by entering the related zone. The program processes the real-time video stream data captured by the webcam, and tracks the positions of the colored sticker located on the index finger.

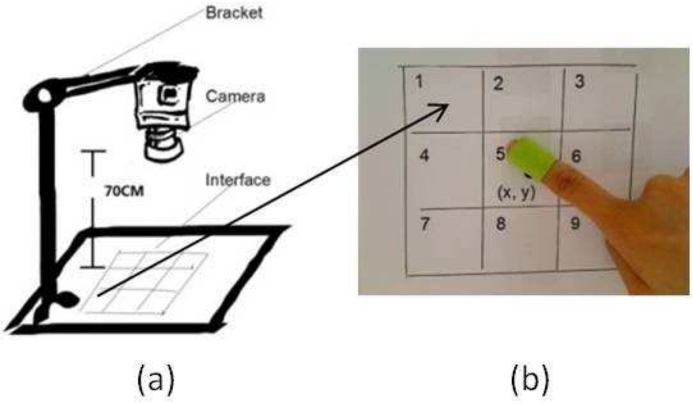
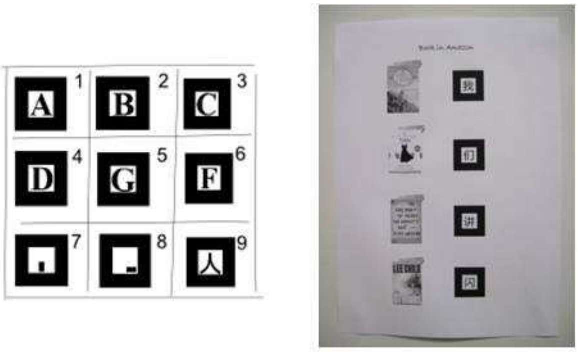


Figure 3.7 Interaction in a fixed environment
(a). Fixed webcam.
(b). Passive grid.

In our “Intuitive Dictionary” test application, the user can directly obtain the image explanation of the words by pointing the words, or can check the corresponding pictures relating to the words by pointing the pictures. The paper grid support has been divided into 12 areas. Six words and six pictures are printed on the grid. When the user places his/her finger over the corresponding areas of the words, a related picture appears on the screen. Similarly, the application also lets the user obtain the word by tapping on the picture.

Besides the grid indications, we also select the markers that can be recognized by ARToolKit to link the real world and the digital one. A fixed configuration is still used but with added markers integrated into each grid. The ARToolKit is then used to recognize these markers (see Figure 3.8 (a)). In this way, through the arrangement of the ARToolKit tags, we can design the paper physical interface with one or more markers to satisfy EDI needs. To demonstrate the feasibility we implemented, an application has been created allowing the user to choose information of interest from a list of book titles ((see Figure 3.8 (b))). Each book title has a corresponding marker printed next to it. When the user selects the marker, he obtains information about this book.



(a) (b)

Figure 3.8 Marker augmented Interaction.
(a). Passive grid with AR tags.
(b). Booklist with AR tags

3.4.2 Mobile Environment Dependent Interaction

To move the interaction from nomadic to mobile mode (interaction based on wearable devices), we abandoned the plastic bracket and equipped the user with a head-band attached webcam and a small screen attached on his/her goggle (see Figure 3.9 (a)). A laptop in the backpack allows the freedom of user in a mobile situation. The selection feedback and information related with the tasks are presented on the small screen. The user can use the menus and markers located in the working or casual environment, mainly posted on the wall in vertical position and at an appropriate distance to facilitate capture by the webcam on the head (see Figure 3.9 (a) (b)).

Figure 3.9 Interaction in a mobile environment.
(a). Camera with a goggle.
(b). Interaction with a grid.

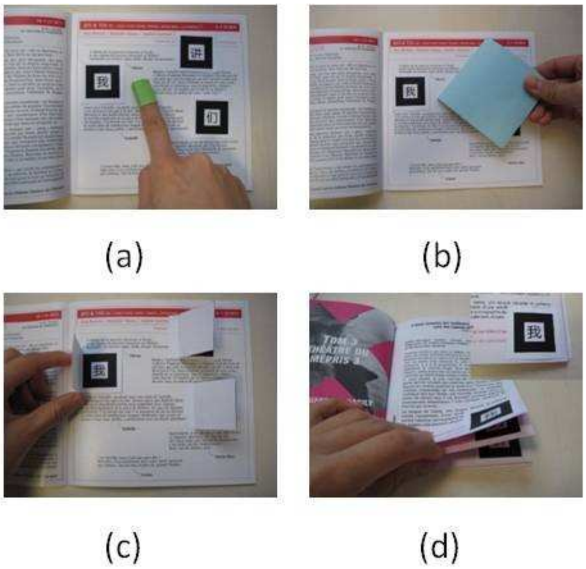


Figure 3.10 Input techniques.
(a). Finger input.
(b). Mask input.
(c). Mask input.
(d). Page input.

Besides the finger selection (see Figure 3.10 (a)), we also design the mask selection technique by presenting the webcam with only one marker at a time. Mask concretization has several solutions. One solution is to create the mask as the object which is used for shading other markers and in the form of a piece of paper, a hand-made stuff or even a notebook with pages to flip (see Figure 3.10 (b)). In interaction with the mask selection technique, the user moves the mask casually and freely. With another solution for the artificial mask, the user switches the paper shelter of the markers in the same way as he/she opens or closes a door (see Figure 3.10 (c)). A large set of markers located on a sheet can be organized and arranged for interaction. In addition, in-environment fixed menus with AR tags are also considered for user interaction. Thus, we obtain the dynamicity required to take a large set of tasks into account. The advantage of marker-based mobile environment dependent interaction is that it is in contact with the environment (i.e. contextual interaction), while the disadvantage is that these menu sheets or markers may be spoiled by vandalism.

Apart from the two selection techniques as we stated above, we also introduced the page selection technique (see Figure 3.10 (d)). We arranged only one marker on each page. The user can show the webcam one marker at a time by flipping through the pages. We also used a pack of cards in Rolodex® or in the flip card mode. Only one card can be shown at a time.

Furthermore, we improve the interface by enlarging the interactive surface, and provide a mobile environment independent interface. Our investigation applies a similar methodology as to the goggle-based prototype to study the projected mobile interface. We replace the goggle configuration with a pico-projector, which projects the interactive menus, the text, the image, the videos, and other contents on any flat surface in the environment as illustrated in figure 3.11. In this way, the user is, not only freed from the limited screen size, but also given an interaction space. We employ the webcam on the user's body as an input device, and the pico-projector as an output device. The webcam and the pico-projector are combined together as a whole unit, the position of which seriously affects the interaction efficiency. The fixing point of the configuration (Kurata et al., 2008) should be settled to capture the person's field of view and especially to distinctly and completely recognize the marker, in order to sustain the vibrations caused by the physical movement, facilitate wearing, observe the hand gestures, etc.

Figure 3.11 Projected interface.

We evaluated the position of the fixing point in an empirical way by using a “quick and dirty” observation, with the aim of finding an appropriate point on the basis of the small existing equipment; if the equipment had been wireless, as small and light as a button, we may have had more choices for locating it. Four points have been studied (see Figure 3.12): the top of the head, the side of the ear, the shoulder, and the front of the chest. Wearing the configuration fixed on these four points respectively, the user’s task was to interact with the markers pasted on the wall. We tested the configuration using five evaluation components: the webcam’s filed of view, view stability, view flexibility, distortion of the webcam and projector caused by the angle, and facility of the fixation on the body. Observation of the user’s actions showed that firm and stable fixing on the shoulder or on front of the chest was not easy without any physical support, particularly when the person is walking. Finally, we decided, in relation with the previous discussion, to fix the webcam and the projector on the top of the head using a bike (or an ice hockey) helmet.

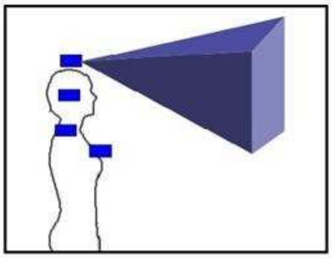


Figure 3.12 Fixed wearable configuration points.

3.5 Scenarios and Applications

We created several scenarios and implemented the relevant applications in order to present how these prototypes could promote interaction in the context and facilitate access by the mobile user to the information freely in a specific environment or a general setting. Two types of interface aforementioned are proposed on the basis of the relationship with the environment: an Environment Dependent Interface and an Environment Independent Interface. On the one hand, we define the environment dependent interface as the interface closely related to in-environment information and markers (on walls, on doors, appliances or any other surfaces). This interface has the ability to provide intuitive interaction techniques, which can recognize and understand the situation of the user and the real environment around him/her. The information supports, i.e. the tangible markers, are static and protected against vandalism. In this way, public or professional guiding information can be used for contextualization based on webcam recognition. On the other hand, we also designed an environment independent interface via which the user can acquire the right information at any time and in any place. Contextualization is achieved by the actual user by showing the webcam the appropriate contextualizing indications. These indications can be grouped on a menu grid and selected by finger, by a mask, or by flipping the pages of a notebook. On each flip card or Rolodex® card, we can find a textual and selected description of the actions to be performed, thus allowing users to contextualize their working environment. With the pico-projector based independent interface, the user can project the menus, schedules, websites, videos, and other information on a plane surface (such as a wall) in the environment or on a small personal projection surface (such as a sheet of paper, a cardboard, or part of the human body). The latter case is mainly a way to solve the problem as frequently the user cannot find an appropriate surface to project in a public place, as mentioned in (Ko et al., 2010). The user is completely free to move his/her working environment and then obtains contextual information independently from the environment. Scenarios and applications were created, to compare selection techniques and appreciate their usability.

3.5.1 Environment Dependent Interface Applications

The first scenario is an indoor way-finding instruction application, which can assist users in finding their destination when they enter a new building for the first time and want to find a laboratory office. Usually, the lab logos are fixed next to the entrance gate of the building. The markers are pasted on one side of the logo, and can be read by the webcam configuration, and then identified to extract the contextual instruction for the indoor way-finding option. When the user points to the right side of the marker with his/her finger with a colored sticker, an interface with the video information or the image instruction will pop up on the small screen attached on the goggles or be projected on the wall. Moreover, when the user reads a newspaper, he/she can only read the text and the image. However, via the pico-projector and the markers, he/she can watch a vivid video augmented on the newspaper (see Figure 3.13 (a)). In another more sophisticated interactive application called as “Research Team Interaction System”, the user can interact with a piece of paper containing markers and grids. The user arrives at the lab and wants to ask the teacher a question but the teacher is not available. The user can then consult the teacher’s schedule to visit again, looking for the appropriate time by using the interface pasted on the door. First, the user selects the action “to consult the timetable”. Next, he/she clicks the photo of the teacher and chooses the date of interest. Finally, the user views the timetable and decides the next time he can come back again (Figure 3.13 (b)). In some public places, like at the bus stop or the gas station, the user can obtain useful advice about nearby restaurants appropriate for their needs. The users can be proposed a location, opening hours, and a telephone number via the markers. In addition, when the tourist is waiting for the bus, if he suddenly wants to know the local weather forecast for that day, he only needs to select the marker and read the information. In this way, the user can obtain instant information quicker and more easily than by acquiring information with several text inputs and menu selections via his/her phone.

Figure 3.13 Interaction in mobility.

(a). Video augmentation.

(b). Dependent interface.

3.5.2 Environment Independent Interface Applications

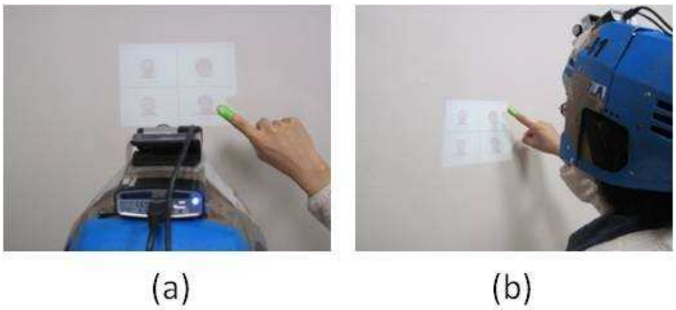


Figure 3.14 Environment independent interaction.

(a). Independent interface.

(b). Selecting the person of interest.

Apart from the general environment, applications are created according to the specific working situation, with an environment independent interface. In the case of maintenance activity in augmented reality we provide the user with the appropriate information in relation with his/her activities by using a pico-projector. For example, in an industrial scenario, a novice technician needs to replace a system board on a laptop computer. The task steps are then projected and superimposed on the real objects. The technician then contextually reads and views the sequence of actions and guidelines. This is not only the process of task completion, but also the process of mobile learning based on the context. Moreover, the pico-projector allows more dynamic behaviors by projecting information on the wall and then allowing the user to interact with it freely (Figure 3.14 (a) (b)).

3.6 Evaluation of Selection Techniques

We conducted a preliminary experiment to compare the finger selection technique with the mask selection technique on a small scale. The goal was to compare the efficiency and usability of two selection techniques. We predefine two questions to study as follows:

- Question 1: Which input technique is quicker?
- Question 2: Which input technique is preferred by the user and why?

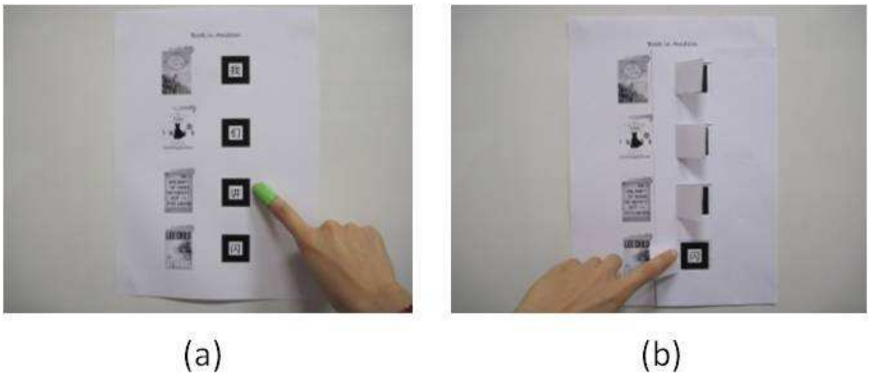


Figure 3.15 Interaction with booklist.
(a). Booklist selected by finger.
(b). Booklist selected by mask.

The experiment equipment includes a book list which has four markers to select, a mask of A4 paper size, and a green sticker which can be located on the index finger (Figure 3.15 (a) (b)). The experimental task consists in selecting items of interest from a list of book titles. Each book title has a corresponding marker printed next to it. When the user selects the marker, he will obtain the information on the book, which will be shown on the goggle screen.

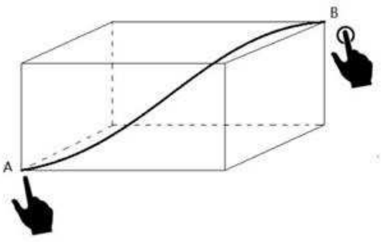


Figure 3.16 The pathway of the finger in the space.

We compared these two techniques by evaluating selection time. Six volunteer participants performed five trials with each technique. Obviously, these two techniques take almost the same distance to reach the target (Figure 3.16): in spatial situation, distance is the length of the curve between A and B. We defined that point A belongs to plane A and point B belongs to plane B. However, since the mask opening and closing actions take more time than the finger selection technique, we assumed that the selection time of the former is longer than the latter. We added a log recorder to the programs to track time. Selection time consists of three parts: reaction time, internal decision time, and execution time. Reaction time is the time between the appearance of the book list and the start of the action, i.e. when the user starts to search for an item in the book list. Internal decision time includes the process of searching for an item of interest and deciding the target item. Execution starts from the first movement of the finger or the mask, and ends when the user clicks (finger selection technique) or presents (mask selection technique) the target marker. However, it is unreasonable to test selection time only. Thus we gave subjects a task, asking them to check the information beneath the four markers in a random sequence as quickly as possible. We recorded the elapsed time of the task, instead of the selection time. The results (Figures 3.17, 3.18) suggested that selection time by finger is less than selection time by mask method. The average selection time by finger and by mask is 16.17 seconds and 19.2 seconds. Though selection time by mask is longer than with the finger, the results of the user comments indicated that users preferred the mask selection technique because they find it easier to flip the mask rather than pointing at a target carefully with the index finger in a space. It was also observed that two participants usually started from a point on plane A and then moved to the target on plane B as the first item selection process, and then kept the finger on plane B. When selecting the next item, they started from plane B each time until completion of the whole process. However, the other participants often started from a point in plane A, which entails a longer selection time.

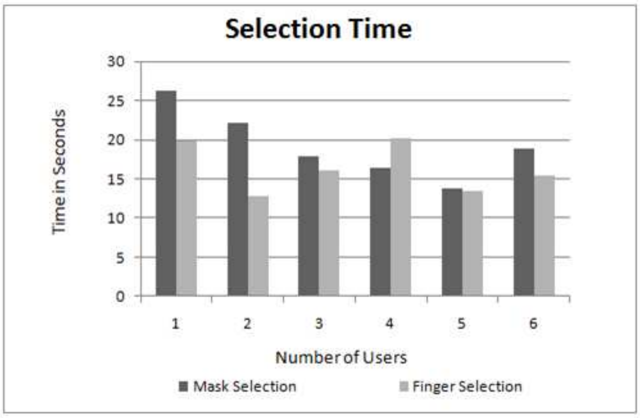


Figure 3.17 Selection time for mask and finger selection techniques.

Figure 3.18 Average selection time for mask and finger selection techniques.

3.7 Summary

This chapter describes our approach for exploring innovative user interfaces and wearable configurations, enabling the user to access to information freely and via simple quick interaction techniques. We present the design and implementation of our prototypes, which are based on camera-capture and computer vision techniques. Wearable configurations mainly consist of a webcam, a small screen attached to the goggles, and a wearable computing unit. For more advanced applications, to acquire a larger interaction output, the user can replace the screen by a pico-projector. To test and demonstrate usability of mobile interactions in the context, we created several scenarios and implemented related applications, on the basis of our different prototypes. Meanwhile, we also explored three selection techniques: finger selection, mask selection, and page selection techniques. Finally, we evaluated the efficiency of mask and finger entering selection techniques. The results of the experiment led us to improve interaction techniques.

Though the limits exist, our methods and prototypes still have advantages. Our mobile AR system allows interaction in mobility with, if possible, at least one hand free. With the marker, the user can obtain information in the context, both simply and quickly. Moreover, more dynamic interactions can be achieved using the pico-projector.

In a conclusion, we stated our design, implementation and evaluation preliminarily in this chapter. In the following chapters, we will investigate in detail a camera-projector and a camera-glasses device unit as configuration, as well as the input techniques, the two interfaces, the related concretized interfaces according to the basic continuum, and devices positions, as shown in figure 3.19.

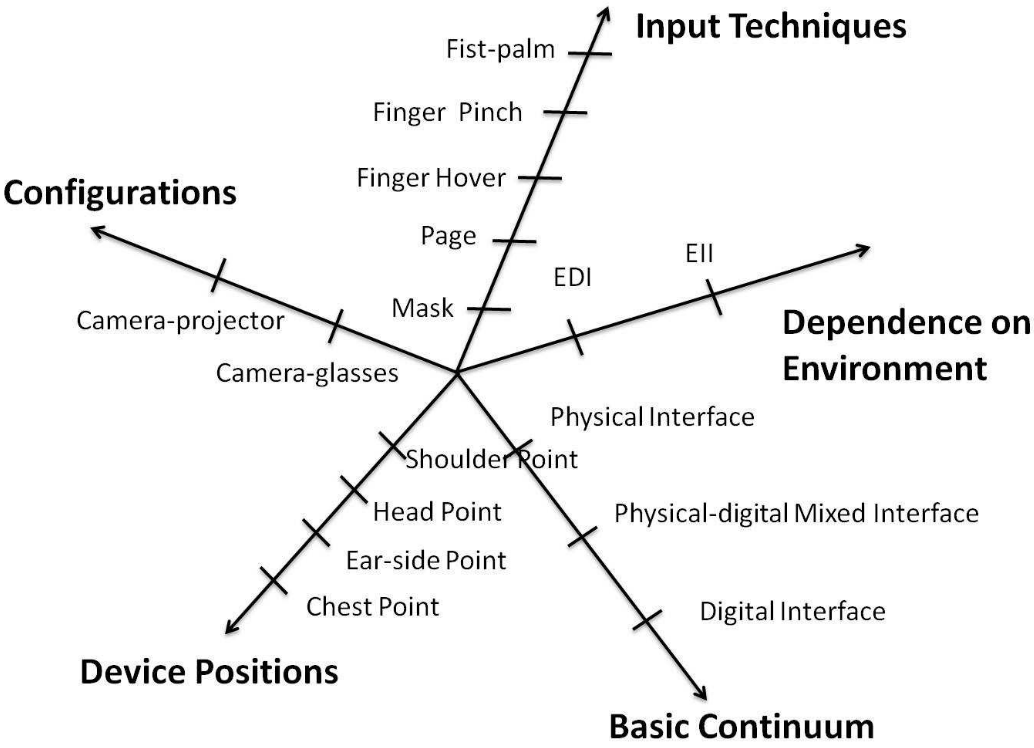


Figure 3.19 Configuration, input techniques, innovative user interfaces, basic continuum, and device positions in this thesis.

To ensure the user a more comfortable interactive experience, in the next chapter we will investigate in greater depth three input techniques with more complicated interactive items.

4 Paper-based Interfaces for Mobile Interactions

4.1 Introduction

In this chapter, we explain our approach for concretizing both the EDI and EII concepts by proposing MobilePaperAccess, which is a ubiquitous paper-based interface for mobile interactions. We employ wearable configurations: a small screen attached to a goggle to provide visual information, a webcam to capture the input signal, and a laptop as the calculating device. Our goal is to create a true contextualization which is more effective and adaptive to users' information needs by taking advantages of dynamic and physical environmental characteristics, based on or independent of the user's location.

4.2 MobilePaperAccess

4.2.1 Input Techniques

We propose three input techniques in this chapter as shown in figure 4: the hover finger input technique, the frame mask input technique and the page input technique. Different from the finger input technique described in chapter 3, the hover finger input technique in chapter lets the user's finger hover for a second and the selection signal can be generated via a span. Thus, when the user points at a button, he needs to remain within the region of the button for an interval of time. As figure 4.2 shows, unlike finger input selection which is distinguished by entering the target area or beyond the area, the hover input distinguishes the non-selection state and the selection state by slipping rapidly and hovering for a while. The items of the interface with the former input technique should be arranged in a sparse way to avoid the undesired item region such as the waylaid button in figure 4.2 (a). However, the interface with the latter input technique could contain more items because slipping is considered as the non-selection, and it is not necessary to avoid the undesired target.

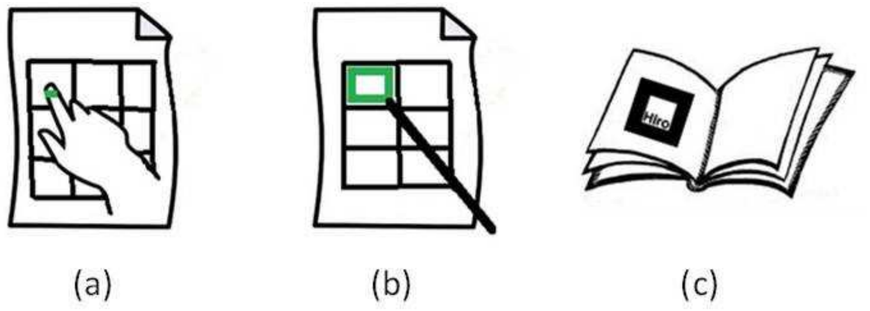


Figure 4.1 Three input techniques.
(a). Hover input technique.
(b). Frame mask input technique.
(c). Page input technique.

Besides the hover finger input, we also propose a mask consisting of a green frame and a wand, which is quite different from the mask input technique in chapter 2. The frame mask in this chapter shares the same interactive method with the hover finger input. For the user, the real physical information on the paper can be read in the center of the mask as tips or indications. On the page input, we have a booklet with several pages and each page has an ARToolKit tag (Kato, Billinghurst, & Poupyrev, 2000). To help the user select, the index in front is for him to read.

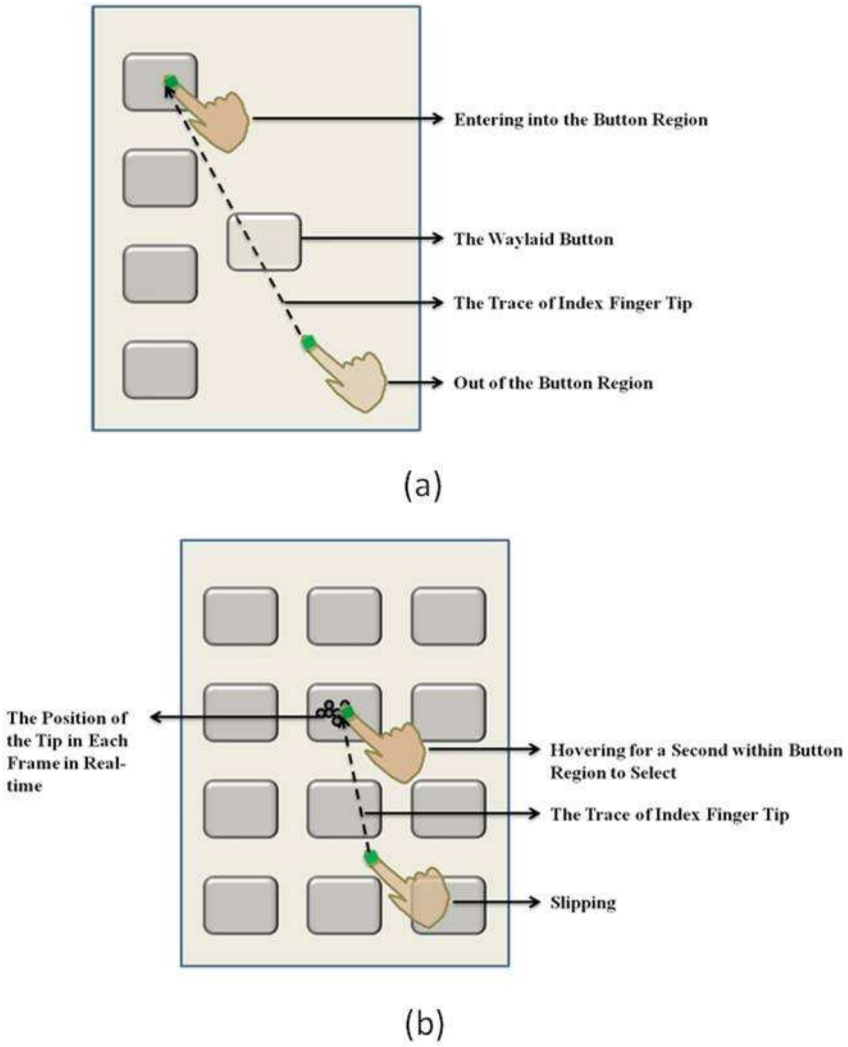


Figure 4.2 Two finger input techniques.
(a). The previous input technique in chapter 3.
(b). Hover input technique.

4.2.2 Paper Surface

According to human factors, the angle of eye rolling is 30^0 up and 35^0 down vertically (see Figure 4.3(a)) and 15^0 comfort and 35^0 maximum horizontally (see Figure 4.3(b)). The average of forward grip reach is 74.3cm (Dul & Weerdmeester, 2008). The interactive surface held in hand should be less than the size of 34.64 cm × 16.08 cm if reading distance is 30cm.

Thus, we choose an A4 (29.7 cm × 21.0 cm) paper as interface, and organize the layout within the comfortable range.

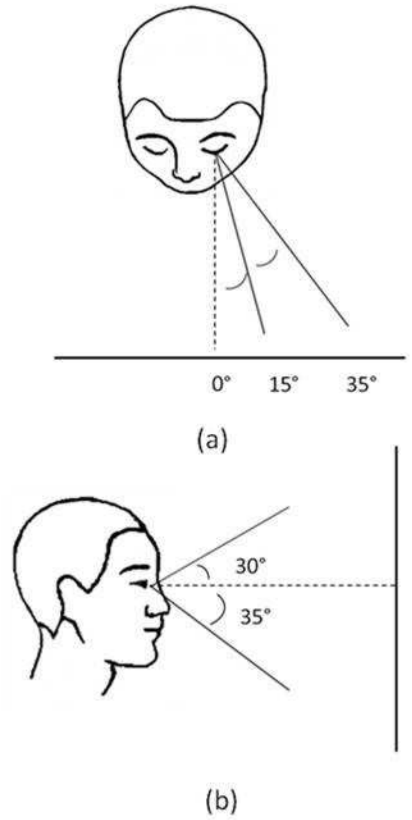


Figure 4.3 The angle of eye rolling vertically (a) and horizontally (b).

4.3 Implementation

The perspective of our MobilePaperAccess includes the paper interactive surfaces augmented with the color markers, a colored sticker located on the user's index finger, ARToolKit tags, the webcam to capture the motion of the marked index finger or to capture the ARToolKit tags (see Figure 4.4 (c)), the goggle with small screen to present the digital information (see Figure 4.4 (b)), and a laptop for calculating. The wearable configuration is illustrated in figure 4.4 (a). We will discuss the details of implementation below.

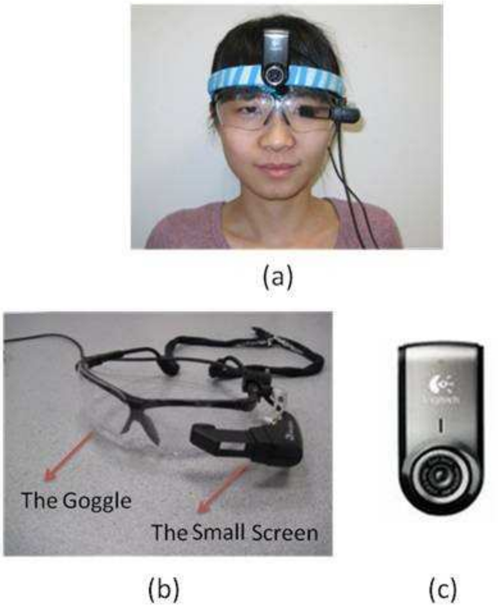


Figure 4.4 Wearable configurations.
(a). Input and output devices.
(b). Goggle with a small screen.
(c). Webcam.

4.3.1 Augmented Paper

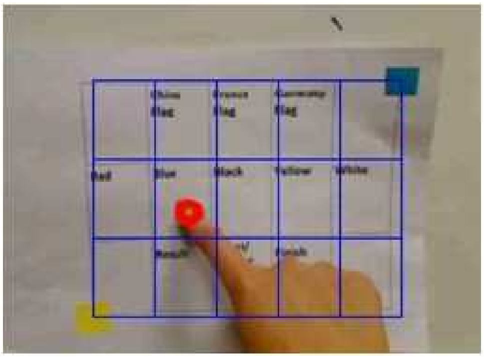


Figure 4.5 Colored augmented paper.

Unlike devices where the input takes place directly on the display surface, our display and the input are separated. Each paper interactive surface is either augmented with the color markers or with ARToolKit tags. Two color markers in a diagonal position (Figure 4.5) shape a rectangle, which can be tracked by the webcam. As long as the webcam recognizes this rectangle shape, the grid within the shape is considered as the icons and can be activated by pointing. The user is unaffected even if he/she rotates or moves the paper slightly without forethought during interaction. The booklet for interaction in EII is augmented with ARToolKit tags, and each page has a tag as the identity.

4.3.2 Goggle with Small Screen

Feedbacks are presented on the small screen fixed on the goggle, either on the right or the left side. Limited by the size of the small screen, we divide the whole display area into two parts: the main display area and the auxiliary area (figure 4.6). The main display area displays the information completely, while the auxiliary one displays the brief response of the information in the form of a keyword or tips, understood by the user in a quick and just-in-time way.

Reset

You start from beginning, please select the action.

Figure 4.6 Visual feedbacks in the small display.

4.3.3 Motion of Finger and Mask

We fix a unique color marker on the index finger of the user, which can be tracked by the Camshift algorithm (G. R. Bradski, 1998) in real-time. As shown in figure 4.7, we record the trace of the color marker and count the number of points in one area such as the grey zone. If the number meets our predefined condition, we regard this action as a pointing. For mask input, we calculate the central point of the mask as the tracking point, which is counted in the same way as finger input.

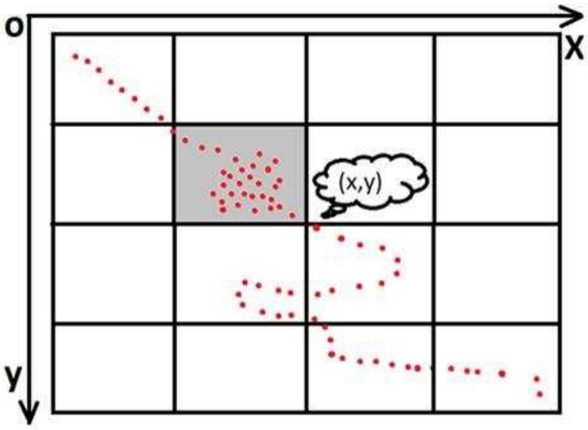


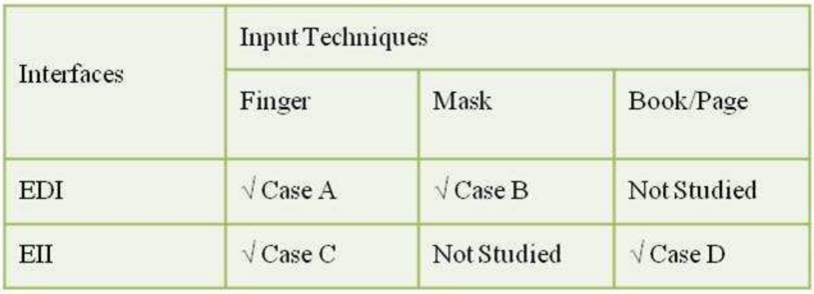
Figure 4.7 The motion of the index finger.

4.4 Research Team Management Application (RTMA)

To prove the concepts of EDI and EII, we developed an application with the aim of managing research team members' exchanges in the form of a paper-based interface. The scenario with EDI is as follows. A research team member wants to consult another member. When he/she arrives at the lab, he/she finds this person is out. He/she then walks close to the door and starts to use RTMA. He/she interacts with the paper on the door, looks for an appropriate time, and selects to check the schedule. After marking the decision, he/she asks for the appointment and obtains a feedback from the system. For the same scenario with EII, the user takes out a customized paper or a booklet held in his/her hand to interact.

4.5 User Study

To obtain a more profound understanding of our input modalities, we organized a structured evaluation to compare our three input techniques (finger input, mask input and page input) and two interfaces (the EDI and the EII). We formed four cases as shown in figure 4.8. For the cases A and B, the participants stood, and for the cases C and D, the users sat or stood in a free way to simulate mobility.



Interfaces	Input Techniques		
	Finger	Mask	Book/Page
EDI	√ Case A	√ Case B	Not Studied
EII	√ Case C	Not Studied	√ Case D

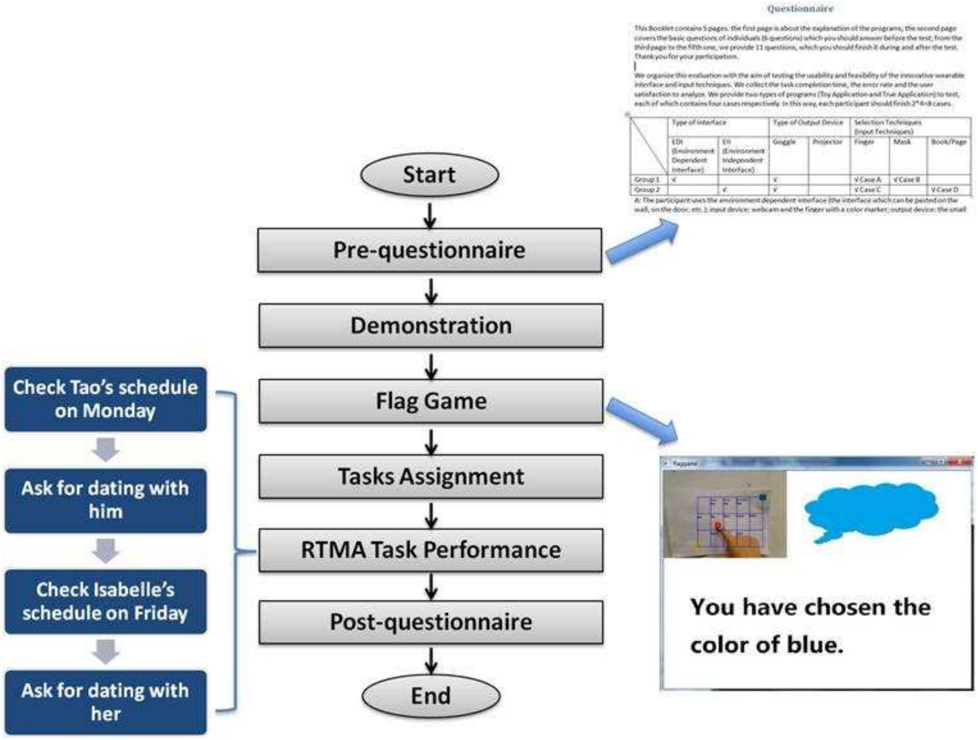
Figure 4.8 Four test cases (A, B, C and D).

4.5.1 Participants

12 student participants were recruited, including 7 males and 5 females. They were aged between 19 and 29 with an average age of 23.2, and their heights ranged from 157cm to 188cm with an average of 171.8 cm. All participants had experience in using mobile devices, but only 6 of them had knowledge of HCI. All of them except one were right-handed.

4.5.2 Procedure

We provided two types of program (Toy Application and True Application), each of which contained four cases respectively. Thus, each participant had to finish 2×4=8 cases. The toy application was a flag game for practicing, in which participants could choose the flag of interest, and then choose its color composition, and finally check the results. The goal of introducing the toy application was to help participants familiarize themselves with the interaction techniques and interfaces. They could play the flag game several times until they felt competent in the following true tasks. The evaluation began with an explanation of the protocol by the text form. The questionnaire attached in the protocol contained two parts: the first part (pre-questionnaire) covered the background information on their familiarity with mobile devices and HCI, as well as the basic data of individuals, to be answered by users before the test; the second part provided questions in Likert scale form (Likert, 1932), which the users had to finish during and after the test (post-questionnaire). Next, after practicing several times with the toy application, the users started the true RTMA test. We employed a within-subjects design, and the order of the cases was counterbalanced with a 4×4 Latin Square (Grant, 1948). For the task completion stage, we asked the participants to perform two tasks in each case once only. All the participants performed the tasks respectively. As in the procedure shown in figure 4.9, they were instructed to check two different researchers' schedules and ask for an appointment with them as accurately and quickly as possible.



Questionnaire

This Booklet contains 5 pages: the first page is about the explanation of the programs; the second page covers the basic questions of individuals (6 questions) which you should answer before the test; from the third page to the fifth one, we provide 11 questions, which you should finish it during and after the test. Thank you for your participation.

We organize this evaluation with the aim of testing the usability and feasibility of the innovative wearable interface and input techniques. We collect the task completion time, the error rate and the user satisfaction to analyze. We provide two types of programs (Toy Application and True Application) to test, each of which contains four cases respectively. In this way, each participant should finish 2*4=8 cases.

	Type of Interface		Type of Output Device		Selection Techniques (Input Techniques)		
	EDI (Environment Dependent Interface)	EII (Environment Independent Interface)	Goggle	Projector	Finger	Mask	Book/Page
Group 1	√		√		√ Case A	√ Case B	
Group 2		√	√		√ Case C		√ Case D

A: The participant uses the environment dependent interface (the interface which can be pasted on the wall, on the door, etc.); input device: webcam and the finger with a color marker; output device: the small

You have chosen the color of blue.

Figure 4.9 The flow chart of the evaluation process.

4.5.3 Variables

Independent variables are input techniques (finger input, mask input and page input) and interfaces (the EDI and the EII). Dependent variables are interaction time. Interaction time starts from the user's correct interaction to his/her correctly stopping each task. We drew the access time and regarded it as the span from starting the application to the user's first interactive action. We also recorded all the errors and found out the reason for the interaction error. The applications recorded the user's input automatically.

4.6 Main Results

4.6.1 Interaction Time

To discover whether there are any significant difference among three input techniques and between interfaces, we used the Mann Whitney U test (Lehmann, 2006) of the nonparametric tests. There are no statistically significant differences (p>0.05) between finger input and mask input for the same EDI case and between the EII and EDI with the same finger input. However, there are the significant differences (p<0.05) between finger input and book input for the same EII case. Figure 4.10 shows the mean interaction time. The interaction time of page input with EII obviously took longer than the others.

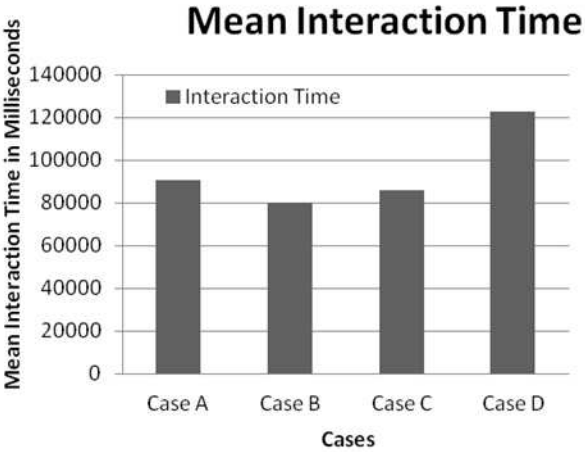


Figure 4.10 The mean interaction time for each case.

4.6.2 Access Time

We measured access time in each trial. Figure 4.11 shows the mean access time in each case. There is no significant difference between cases A and B, between A and C, and between B and C; which only took less than 8 seconds to access. However, the access time of case D is nearly two times longer than the other cases.

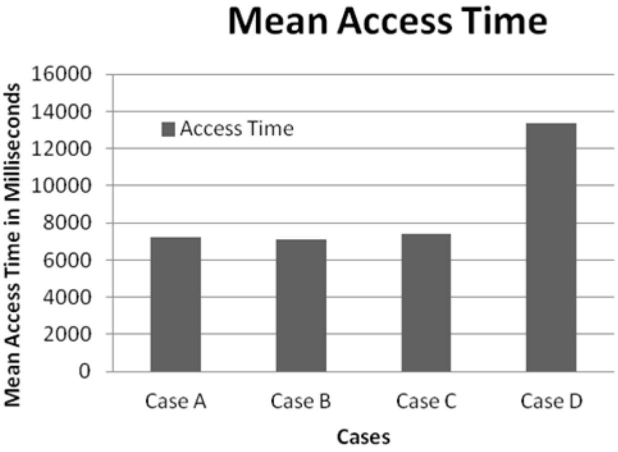


Figure 4.11 The mean access time for each case.

4.6.3 Interaction Errors

In this evaluation, all the participants have finished their tasks even though some of them have made errors. Through observation, we found that the reasons for the errors are mainly due to the locomotion of users, misunderstandings of tasks, and the attempt to do more than the tasks. These three errors were counted respectively. Among these errors, locomotion is the main cause of the interaction error. Thus, we counted the number of the locomotion errors of all participants in four cases as shown in 4.12. The number of interaction errors with EDI is less than with EII, and with finger input less than with page input.

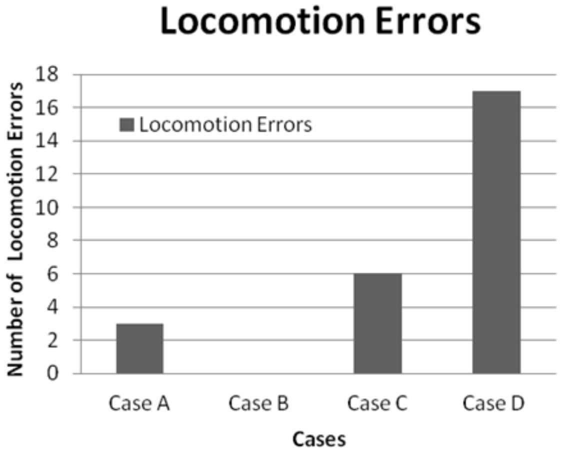
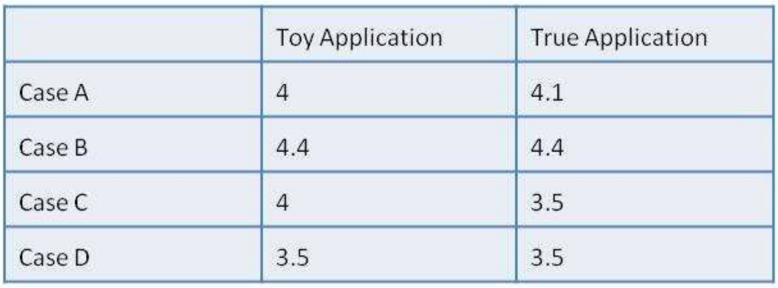


Figure 4.12 The locomotion errors for each case.

4.6.4 User Satisfaction

We had five levels (1-Strongly disagree, 2-Disagree, 3-Neither agree nor disagree, 4-Agree, 5-Strongly agree) for the Likert items to describe easiness of learning and convenience of interaction. Figure 4.13 gives the average scores of each case. It showed that all the participants thought it was not hard to learn and perform (Mean scores are all more than 3). Also, interaction in case B is easiest to learn and most convenient to perform.



	Toy Application	True Application
Case A	4	4.1
Case B	4.4	4.4
Case C	4	3.5
Case D	3.5	3.5

Figure 4.13 The mean score of users' satisfaction with the toy application and true application in four cases.

4.6.5 User Comments

For case A, i.e. the finger input with the EDI, four participants felt their lifting arm was tired after operating for a while, which led them to interact unsteadily with their finger. However, two participants expressed that the fixed position was efficient and convenient for interaction. Moreover, two participants commented that when they moved their arms and fingers the physical chain reaction effect resulted in a tiny movement of the camera fixed on the head. For case B, one person said that the frame of the mask made it easy to choose and select items, while another person was unable to work well with the frame angle of the marker. Two people felt their arms tired. For case C, more than half of the participants reported the long time involved in lifting their arms and the unsteadiness of their fingers. They thought that it was not easy to hold the interface in the hand steadily. Also, two participants reported the chain reaction effect. For case D, four participants explained that when there were more pages in the booklet, they found it less easy to search for the right page to interact; it is not convenient to return to the index each time. Only one mentioned the chain reaction. One person preferred marker interaction for the faster and more sensitive interactive experience. For the devices, six participants felt the screen was small to read, which made them feel a little faint or tired.

4.7 Discussions

Case B has the best performance of all with the shortest interaction time, access time, no locomotion error and best satisfaction. Compared with in case A, in case B fewer participants reported a tired arm because the band with mask is more comfortable than the lifted hand. Case A has a better score than case C; they have almost the same interaction time and access time except that case A has a better satisfaction score. Also, in case A, fewer participants reported a tired arm, the reason for this being the fixed and stable interface. Case C has a better score than case D due to its shorter interaction time, shorter access time, fewer interaction errors, and better satisfaction score. Case D is influenced by locomotion errors more than the others. We found that the more pages there are, the harder the selection action will be. Due to the action of searching pages via returning to the index, the input technique in case D leads users to an unsteady interaction state.

To reduce locomotion errors in the system with EDI and EII, we propose the two following solutions. One of the solutions is to reduce paper size and increase paper hardness. Expanded paper size can avoid the problem of the fat finger, but it is easy to carelessly leave part of the paper out of the webcam range. In addition, some users hold the paper with different degrees of strength that can result in bending of the paper, thus reducing webcam recognition and leading to the same interaction problem as the locomotion errors. However, paper hardness can compensate for this effect. We can choose cardboard as the paper interactive surface of the EII. The physical paper interface has a low multiplexed ability; the selected items are physical and cannot be changed dynamically. If we reduce the space and size of the paper-based interface, the content also decreases. To balance size and content, we propose introducing the ARToolKit tags into the paper interface to provide the EDI with the aforementioned physical and dynamic interface. We also found that it was tiring for participants to raise hands in the same position as the eyes after a certain period of time. Also, the chain reaction reduced interaction efficiency. Thus, we propose changing the position of the webcam from the forehead to the chest to lower the elevation of users' hands and ensure stability. We will discuss the physical and dynamic mixed interface and the chest position in chapter 5.

4.8 Summary

In this chapter, we propose, design and implement a MobilePaperAccess system based on a webcam, a small screen attached to a goggle and a laptop as a calculating device. MobilePaperAccess is a wearable camera-glasses system with a tangible user interface allowing mobile interaction. The users access to the digital information from a paper interface extending the input space, and can interact with the paper using fingers, masks and pages. This interaction modality supports privacy and avoids the problem of fat fingers and focus point blocking by the shadow and the actual finger. The system is devised to validate our concepts of Environment Dependent Interface (EDI) and Environment Independent Interface (EII), which focus on enabling people to access their personal data as well as public resources at any time and in any place. We compare two interfaces (EDI and EII) and three input techniques (finger input, mask input and page input). The quantitative and qualitative results show that the main interaction error is the locomotion error and that the mask input with the EDI has the best performance. To conclude, as shown in figure 4.14, we investigated camera-glasses device unit, three input techniques, both EDI and EII, physical interface, and the head point in this chapter.

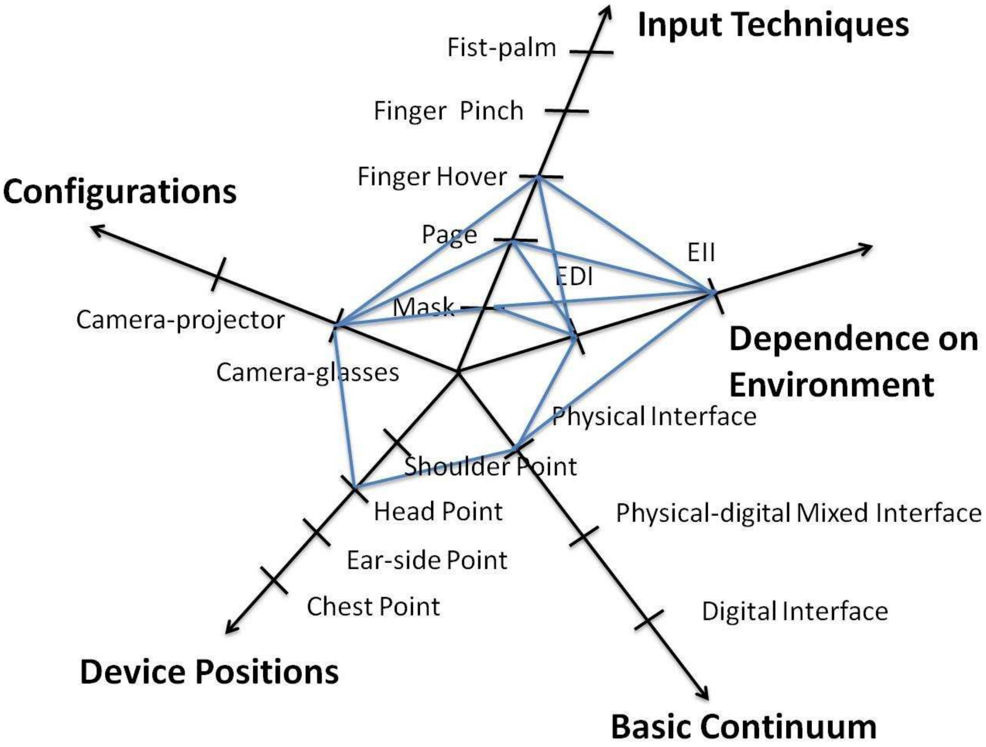


Figure 4.14 The spider figure of paper-based interfaces for mobile interactions.

In the next chapter, we will investigate the hover gesture in greater depth, and explore the scalability of the projected interface in a preliminary evaluation.

5 Wearable One-hand Gesture Input and Scalable Projected Interface

5.1 Introduction

With rapid development of portable projection technology and computer vision techniques, the interaction modality of the camera-projector system provides an alternative method to ubiquitous access to mobile interaction and communication. While studying this interaction modality, we found that the projector has the property of scalability, enabling it to display different sizes of interfaces according to surface size and the distance between the projection surface and the projector. Unlike the screen with nonscalability, if we provide the same content and layout for the different size interfaces, usability will decrease. This problem occurs commonly with adaptive interfaces: usability will be less if we transfer directly the same elements and layout from the web browser of a traditional desktop screen to the small screen of mobile devices. The difference between scalability and adaptability is that the scalable interface is presented by one device, while the adaptive interface exists in several devices. With the aim of solving the aforementioned problem of scalability, we propose an approach for providing the appropriate interfaces by detecting the distance between the surface and the pico-projector.

In this chapter, we propose the PlayAllAround system, a wearable system with one-hand gesture input and scalable projected interfaces to concretize our concept of the Environment Independent Interface (EII) defined in our previous work. The EII refers to the relationship between the interface and personal information which could be used for personal self-contextualization, with which the user can acquire the right information freely. Our approach is implemented by employing the following wearable configurations: a pico-projector, a webcam and a wearable computer-like tablet used only for calculating and computing. PlayAllAround can be adapted to different sizes of the interactive surface: the larger farther projective surface on the wall and the smaller nearer one on the notebook or other surfaces. Via image processing methods based on computer vision, we also extract the size of the projective area for automatic calibration, as well as provide offset and non-offset of cursors. We then infer the appropriate adaptive content and layout to produce an appropriate interface via the ARToolKit tags on the surface. Furthermore, we employ the hover hand gesture as the input makes interaction more natural and intuitive. We study hover time to find the bearable region and the satisfied hover time. Our goal is not only to ensure a transition from existing WIMP to post WIMP, but also to study the appropriate interaction techniques, the easy-to-learn mechanism, and the usability of the innovative modality.

5.2 Overview of Camera-projector Interaction Techniques

In this section, we first describe the design of the scalable projected interface before discussing the design of the hover gesture and its interaction technique. Although the projection image has its size on a continuous range, only the nearer small-size interface and the farther large-size interface are defined and selected to investigate.

5.2.1 Scalable Projected Interfaces

As the projector approaches the wall or other surfaces, the image becomes smaller and brighter, and has a higher resolution, but the size of interactive items decreases. Inversely, as the projector recedes from the surfaces, the image becomes larger and less bright, and has a lower resolution, but size of the items increases. Obviously, item sizes increase linearly as the distance between the surface and the projector increases as illustrated in figure 5.1 (a) and (b). Thus we reduce the size of the items in the larger projected interface and increase the number of items as shown in figure 5.1 (c), which can give full advantage to the larger display capacity of projector.

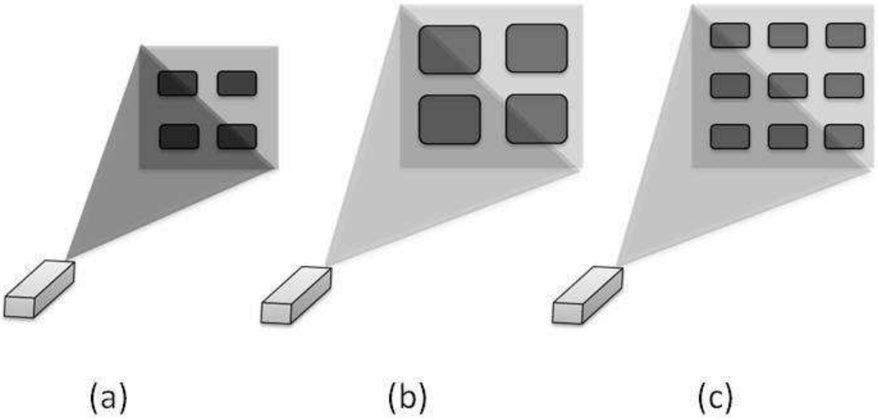


Figure 5.1 Multi-scale interfaces.
(a). Nearer projected interface.
(b). Farther projected interface.
(c). Farther interface containing more items.

With the property of scalability, it is inappropriate to introduce the same elements and layout into the different size interfaces. To provide the different elements and layout without altering the function of the interface, we define the reference-cell to specify the size of interactive elements and items, and formulate the rules to specify the process from the application tasks to the organization of the scalable interface.

5.2.11 Reference-cell and Scalability Threshold

Hands are the articulated objects, the skeleton system of which is formed by bones and joints. For at least the lowest convenient selection by the tip of the index finger, it is preferable to consider the width of the distal interphalangeal (DIP) part of the index finger as the length of the reference-cell for pointing. DIP width is the distance between the ulnar side and the radial side of the distal interphalangeal joint crease of the index finger (Rogers, Barr, Kasemsontitum, & Rempel, 2008). According to hand anthropometric data, the mean (SD) of DIP width is approximately 17.0mm (±1.9mm). Thus, we define 20.0mm as the length of the reference-cell (see Figure 5.2); the smallest possible size of the pointing icon is limited as the square of 20.0mm×20.0mm. Our interface is organized arbitrarily via these reference-cells as shown in figure 5.2. Also, one pointing icon or selection icon can run across several reference-cells. Compared with shift technology (Vogel & Baudisch, 2007), we avoid the occlusion of the actual finger and the shadow by formulating reference-cell size and cursor offset.

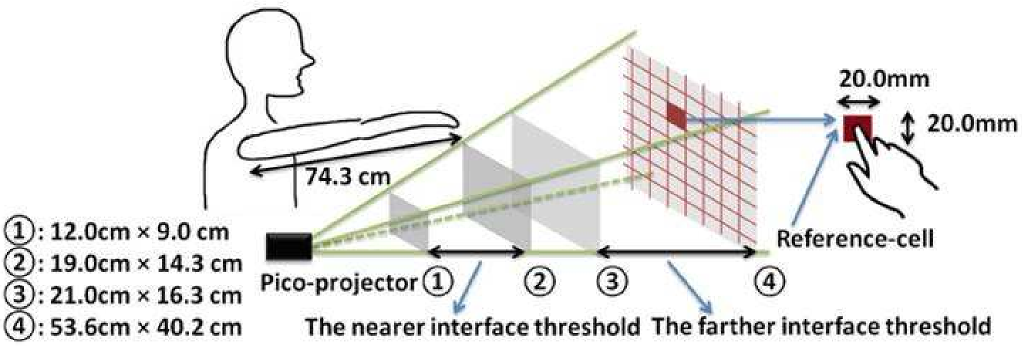


Figure 5.2 The reference-cell and the threshold of scalability.

With the data from (M. L. Wilson et al., 2010), we collect examples of a number of common projection surfaces in everyday life, including door, wall, whiteboard, fridge, notebook, etc., all of which have the appropriate texture and simple unified color to make projection effective. As shown in figure 5.2, the size of 12.0cm×9.0cm is the minimum projection size of our pico-projector, which is the lower threshold of the nearer interface. We measured projection size as approximately 19.0cm×14.3cm (the higher threshold of the nearer interface), when the average arm forward grip reach is 74.3cm. We then add 2cm to width and height respectively, which is the minimum projection size for the threshold of the farther interface. For maximum projection size, we specify 100cm as the perpendicular distance between the surface and the user's feet. When reading distance is 100cm, the comfortable reading region is approximately 53.6cm×115.5cm, and we cut part of 53.6cm×40.2cm as the maximum projection surface, which is less than the projection size limitation of the pico-projector (101.6cm×76.2cm).

5.2.12 Process from Application Tasks to Scalable Interfaces

We break down the application tasks and assign the labels of weightiness for ranking property to each sub application task as illustrated in figure 5.3. The tree created via CTTE (Mori, Paternò, & Santoro, 2002) demonstrates the application tasks. First, we defined a specific ranking property for each application task. Weightiness of this property varies: weightiness is ranked as 1, 2, 3… where the first rank is marked as #1, the second rank as #2, etc. Then we assign the weightiness for each application task. Next, we use the Min rule to process the weightiness of the property of the interaction task relating to the leaf node of the application task. Taking the interaction task "selection" relating to the application task "SelectPerson", for example, the related application task is in the position of the leaf node. The property of its parent node is #1, and its actual property is also ranked as #1. Thus, according to the Min rule, we obtain the property of the interaction task "selection" relating to "SelectPerson" as #1. Similarly, we obtain the property of the interaction task "visual output" relating to "AskforMeeting" as #3.

After attaching the weightiness of property labels to the interaction task, we use the layout filters to decide the interface layouts. Two filters are considered: the hierarchical filter and the disposal filter. With the former, the projected interfaces are made up of the unfolded and folded interactive items. With the latter, the projected interfaces are made up of the retained and the discarded items. When the scalable interface reaches the small size, the layout is organized with the unfolded elements and folded items in a hierarchical structure with the hierarchical filter. However, only the interaction tasks with the primary weightiness of property can be left in the layout with the disposal filter. In this way, we can obtain the appropriate layout and relative size of the elements for two scale interfaces.

In other words, we define the hierarchical layout filter as the view-altering filter method, which changes only the layout and the view of the interface. Furthermore, to keep the wearable small-size interface concise, we offer the disposal layout filter as the content-altering filter method, which not only changes the layout and the view of the interface but also removes the secondary parts and leaves the primary parts concisely in the small-size interface. In this way, we obtain the layout and the relative size of the elements for each interface: both the farther large-size one and the nearer small-size one.

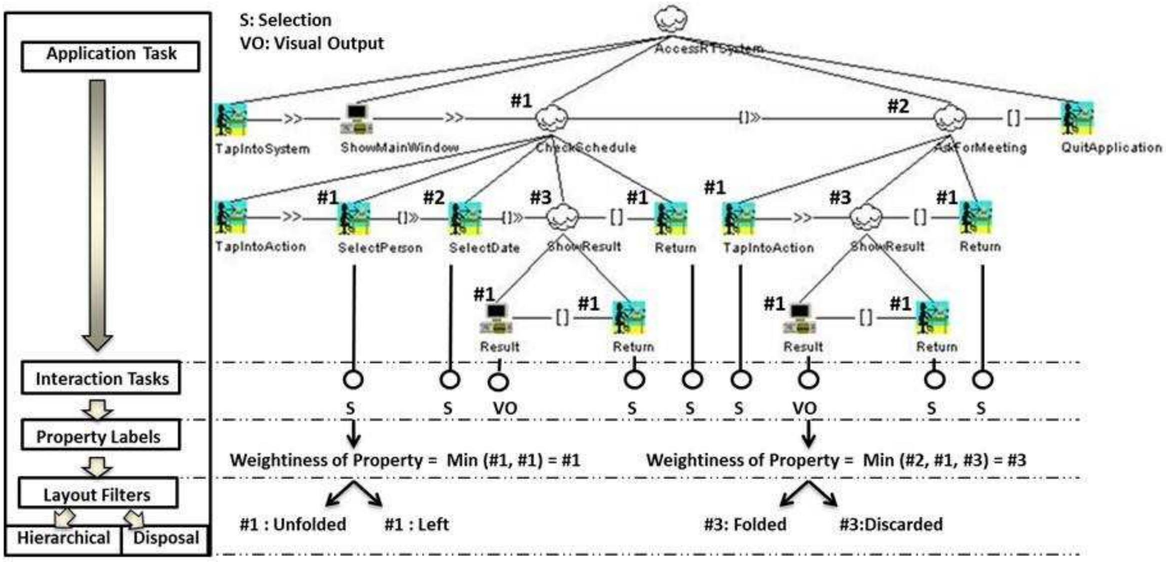


Figure 5.3 The model from decomposition of application tasks to composition of interface layouts.

5.2.2 Hover Input Technique

One solution of hand gestures for user's input is to let the user hover for a second with his/her finger, and the selection signal can be generated via a span. When the user points at a button, he needs to remain in the position of the button for a time period (namely a time interval). In this way, the button is considered as selected. Buxton specifies a three-state input model (Buxton, 1990), which provides a conceptualization of some of the basic properties of input devices and interactive techniques. We propose a hover gesture with a three-state input model, illustrated in figure 5.4. The first state, (state 0), is what we will call "out of range". In this state, the finger is beyond the reach of the webcam's vision, so any movement of the finger has no effect on the system. The system starts to track and the tracking symbol is the tip of the user's index finger as the finger is entering the region of the webcam (state 1). The two actions "Hovering for a Second" and "Stop Hovering", are closely linked, similar to the relationship between opening the door and closing it. In this way, the "Stop Hovering" action is non-substitutable and strongly linked to the preceding action. Thus, the return path from state 2 to state 1 is drawn in gray as illustrated in figure 5.4.

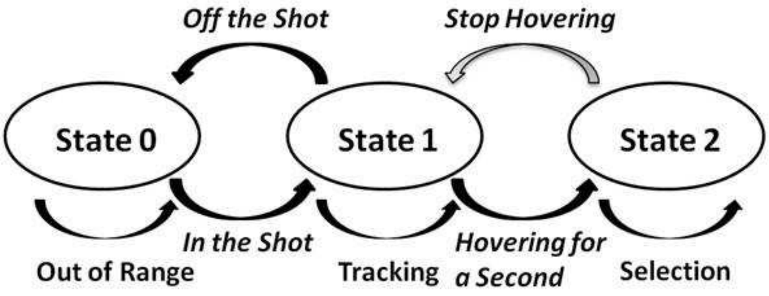


Figure 5.4 The three-state model of the hover gesture input.

5.3 Implementation

As we stated above, with the purpose of providing the user with an Environment Independent Interface for more freedom, the user can project the personal interface to interact. In this chapter, we discuss the projected dynamic digital interface which can be overlaid on daily arbitrary surfaces. Although many researchers endeavor in the research of camera-projector systems, the property of scalability of the pico-projector has not been studied as thoroughly as the design of the layout of the scalable interface. And as shown in figure 5.5, the camera-projector unit is stabilized in front of the user' chest, and the same configuration is employed to study scalability of the projected interface and hovering time of the hover gesture.

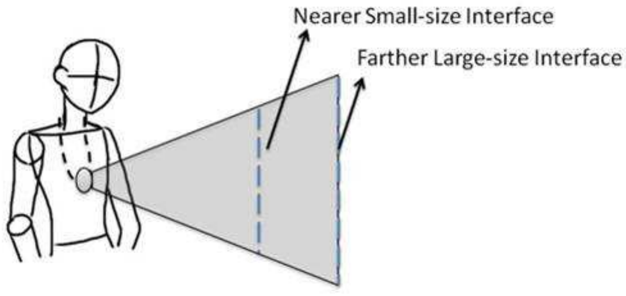


Figure 5.5 System concept and Implementation.

5.3.1 Wearable Configuration

Our wearable configuration contains a webcam, a pico-projector and a tablet for calculating. The Logitech webcam can obtain a 640×480 RGB frame. The pico-projector has a resolution of 640×480 pixels and a projection size (diagonal) of 127cm maximum and of 15cm minimum. The tablet is equipped with a multi-touch screen, which can be carried on the back or in a messenger bag along the body. We fix the webcam and pico-projector together on the light foam support (see Figure 5.6 (a)) and choose the chest as the worn point (see Figure 5.6 (b)). The strap hanging on the neck is adjustable according to user stature. When the user is walking, we also use a chest strap to prevent the configurations from swaying. The projection angle of the foam support can be readjusted up to 15° upwards to accommodate the human vision line. This chest location shifts the burden compared with the head-worn configurations, and is stable enough to allow interaction by hands.

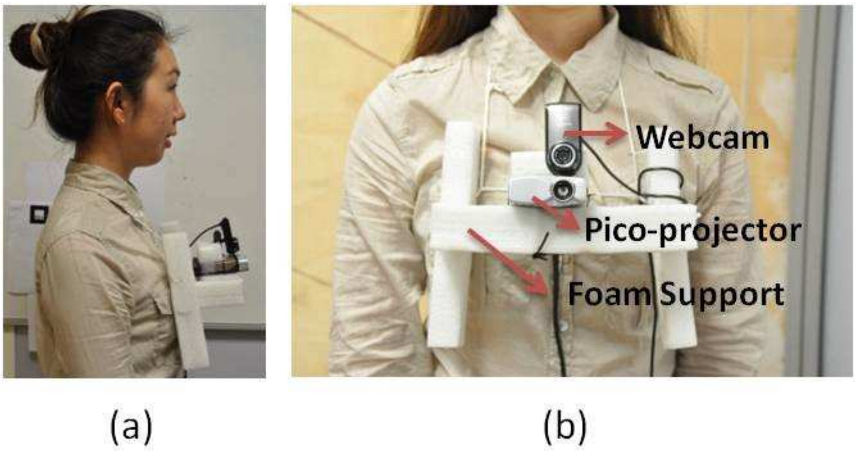


Figure 5.6 The chest mounted configurations.
(a). Chest position.
(b). Foam support and camera-projector unit.

5.3.2 Recognition of Hover Gesture

Our research is based on the computer vision-based hand tracking method. We fix a unique color marker on the user's index finger, which can be tracked by the Camshift algorithm (G. R. Bradski, 1998) in real-time. As shown in figure 5.7, we record the trace of the color marker by noting the x and y coordinates of the color marker in each frame. We define a moving region in charge of distinguishing the non-pointing action and the pointing action. This moving region follows the motion of the finger. In a continuous time interval, if the number of points of the color marker reaches the predefined value in this moving region, we regard this action as a pointing. Otherwise, we consider the action as a slip action. The click event is then transferred to the related located items such as the buttons or other interactive widgets, which are superimposed on the moving region in position.

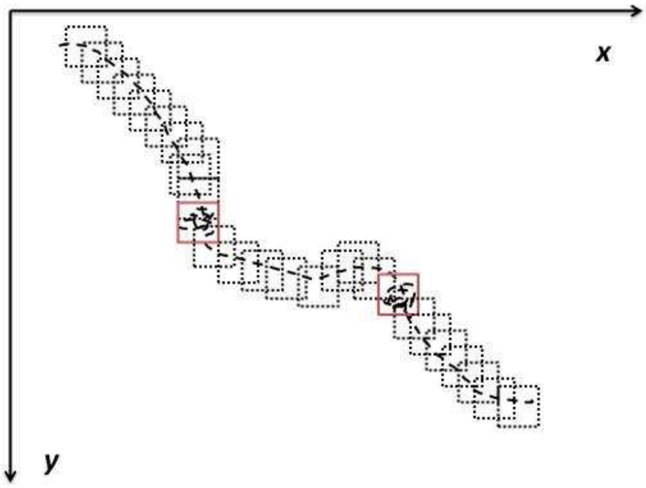


Figure 5.7 The finger trace and the pointing action.

For the arm-reached interface, namely the nearer small-size interface, we take the direct input technique, with which the user's finger is superimposed by the cursor (see Figure 5.8 (a)). In this way, the input takes place directly on the projected surface and the user does not need to pay attention to the cursor, but only to his finger. For the beyond arm-reached surface, that is to say the farther large-size interface, we take the indirect input technique, with which the position of the finger separates with the cursor. To avoid the occlusion of the actual finger and the shadow beyond the finger, we set an offset of the cursor, which is higher than the finger as shown in the figure 5.8 (b).

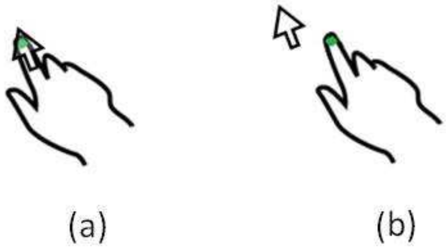


Figure 5.8 The position of cursors.
(a). Non-offset of the cursor.
(b). Offset of the cursor.

5.3.3 Auto-calibration of Projector and Camera Coordinates

Although the camera and the pico-projector are linked, the disunity between the coordinate system of the camera and that of the projected surface continues to result in a mismatch between the finger position and the cursor. Instead of the traditional manual-calibration method, we calibrate automatically by detecting the lightest area. It is obvious that the projected area has the greatest brightness as shown in figure 5.9 (a). We extract the contour of the projected lightest area in real-time (see Figure 5.9 (b)) and only process the image of this lightest area for tracking. In this way, we unify the coordinates of the camera and the projected window, and provide the non-offset or offset of the cursor for the nearer small-size interface or the farther large-size interface.

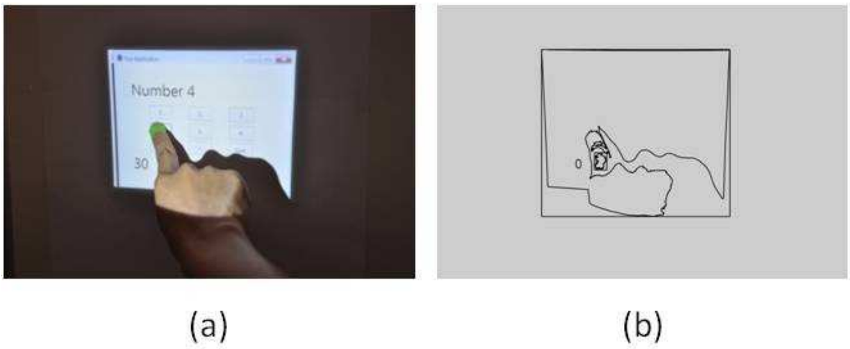


Figure 5.9 Auto-calibration.
(a). Original Image.
(b). Contour of the lightest area.

5.3.4 Depth Sensing via ARToolKit Tags

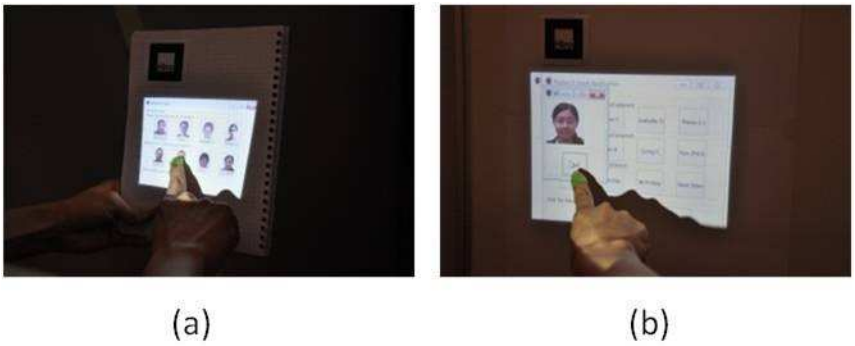


Figure 5.10 Depth sensing via ARToolKit tags.
(a). AR tags on the pad.
(b). AR tags on the wall.

Contrary to the high price of the depth sensors or depth-camera, we present an economical approach for detecting distance between the projected surface and the user. We place the paper-based ARToolKit tags on the surface of the wall, notebook, pad, or door (see figure 5.10 (a) (b)). We can then use them to automatically obtain the distance from the camera to the surface. The ARToolKit has the ability to detect the distance between the marker and the camera. We define a threshold to distinguish the nearer interface and the farther interface even though distance is a continuous value. Thus, the user can switch freely between the two different threshold interfaces via the distance between him/her and the surface. Since recognition of ARToolKit tags is less efficient in a darker environment but the pico-projector works better with insufficient lumens in the current technology, we continue to select a darker environment as the interaction environment. In the interaction process, we asked users to leave the ARToolKit tags within the projection area at first and then instructed them to access to the system and to move the projected direction slightly to avoid the tags.

5.4 Research Team Interaction System (RTIS)

5.4.1 RTIS Scenario

As we stated above, in the smart city (David et al., 2011), the IEI, EDI and EII can solve the same problems faced by the user, as well as different problem respectively. To concretize these three interfaces, we implemented an application known as the Research Team Interaction System (RTIS), which varies slightly according to the interface, and has virtually the same function as the RTMA that we mentioned in chapter 4. In this chapter, we focus on our concept of EII and PlayAllAround System, and continue to employ the RTIS.

The scenario with large size scalable interface is as follows. One day, a research team member wants to consult another member. When he/she arrives at the lab, he/she finds the person in question is out. Thus, he/she approaches the door with a marker pasted in advance in the lab and starts to use the RTIS. He/she interacts with the projected interface on the door, selecting the relevant person and looking for his/her schedule. When he/she decides on the appointment date and time, he/she asks for this appointment and obtains feedback from the system. For the same scenario with a small size scalable interface, the user takes out a paper or a booklet held in his/her hand as the projected surface, projects the interface, and completes the action of checking the schedule as well as asking for an appointment with the person in question.

5.4.2 Scalable Interface Creation Process

While continuing to explore our method for creating the appropriate scalable interface, we provide an illustrative example based on the research team interaction system scenario as stated earlier. Since the RTIS function is essential, we select the hierarchical filter to reorganize the small size interface, which is a view-altering method without any change in content.

The formation process of the nearer interface is as follows. The researche team interaction system is in charge of checking the schedule and making appointments with the target person on the target day. The tree created via CTTE shows the application tasks. We select the application tasks appropriate for the nearer small-size interface and mark them with the different weightiness for property as (#1, #2 and #3) as illustrated in figure 5.11. The interaction tasks are then defined and labeled with the final calculated weightiness of the property, which are marked with the bottom line. According to these final calculated weightiness of property, we use the letters within the black filled circle, which are alphabetically related to the weightiness values in a hierachical sequence. These letters represent the layout and the placment order of interactive items within an interface.

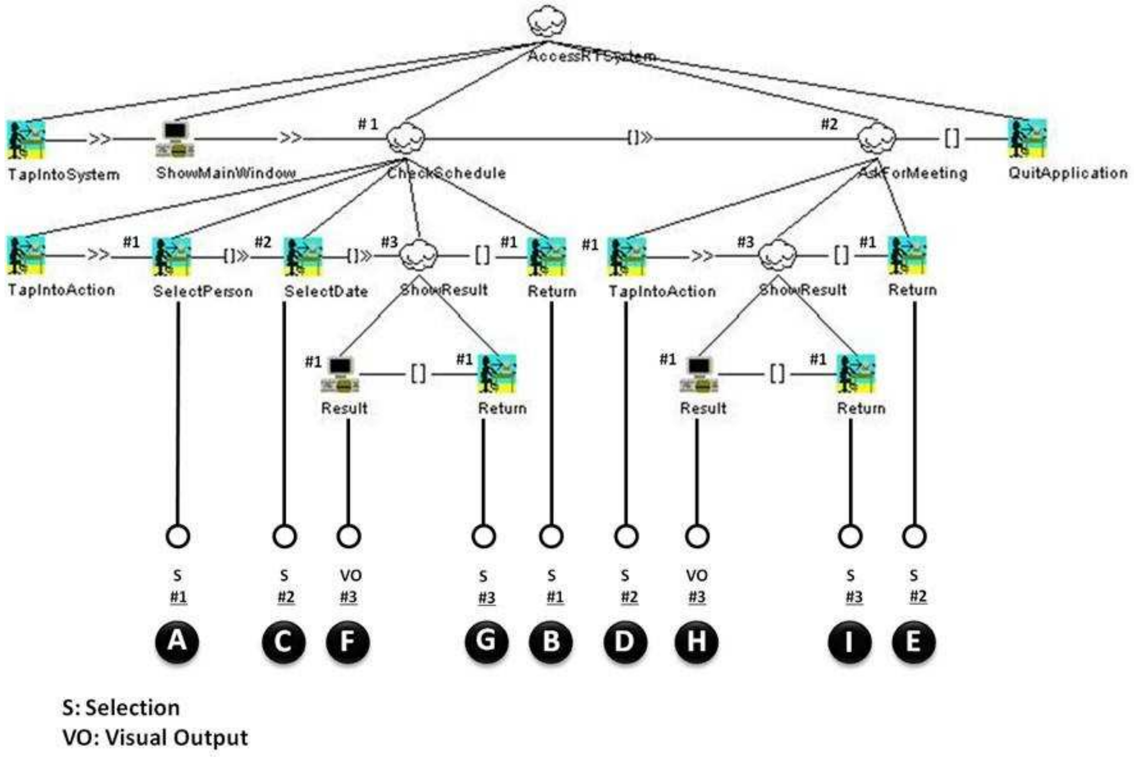


Figure 5.11 The formation of the nearer small-size interface.

We then place the letters in the interface as shown in figure 5.11. In the condition for view-altering without content altering, the interface is divided into parts and reorganized as the multi-window form. Besides title part, the letters A and B are ranged in the primary window (see Figure 5.12 (a)). Furthermore, letters C, D and E are ranged in the secondary window (see Figure 5.12 (b)), and letters F, G, H and I are ranged in the pop-window with the property of (3) (see Figure 5.12 (c) (d)).

The interface formation process for the farther large-size interface is similar to the nearer small-size interface, illustrated in figure 5.13 and figure 5.14. Since the large-size interface can contain more content, items with lower weightiness of the property in the small-size interface are observed as having high weightiness in the large-size interface. In the condition for the view-altering mode with the large-size interface, the interface is divided into parts and reorganized in the multi-window form, but each window including both primary window and sub-windows contains more interactive items. A, B, C and D are all ranged in the primary window (see Figure 5.14 (a)), whereas E, F, G and H are assigned to the secondary pop-windows (see Figure 5.14 (b)).

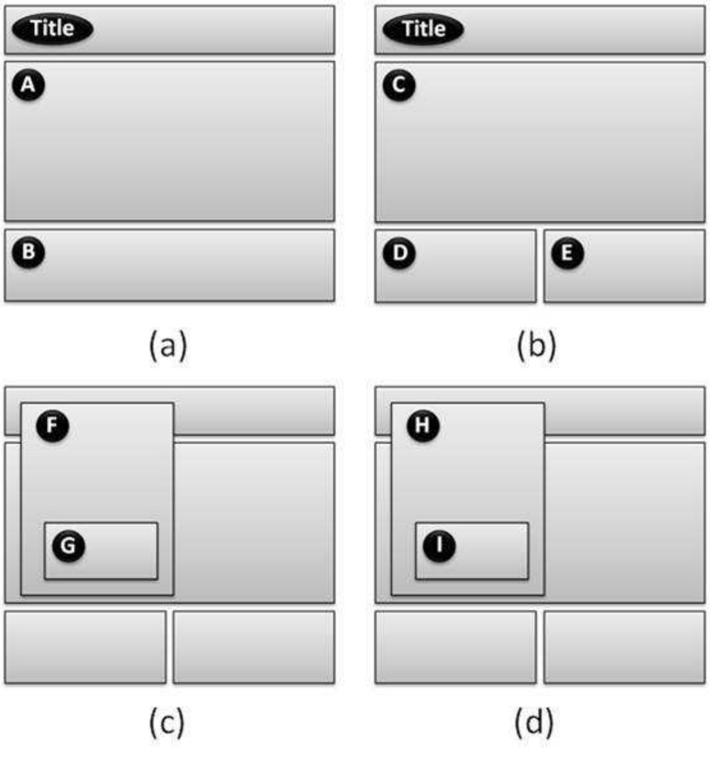
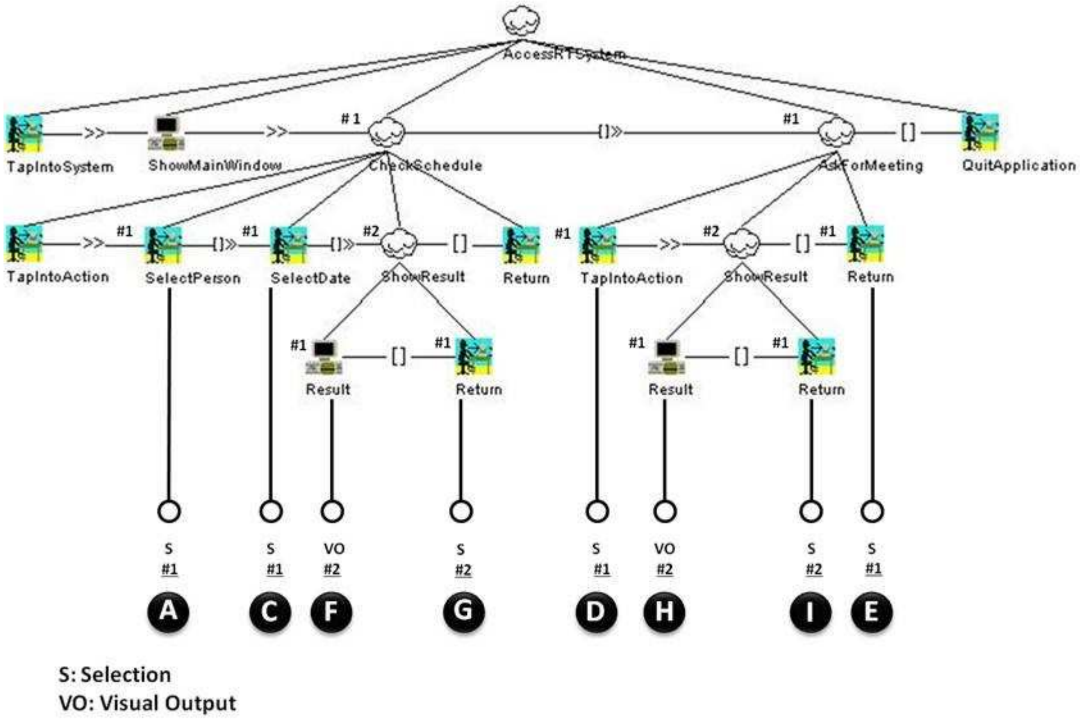


Figure 5.12 The layout of nearer interfaces.
(a). The primary window with weightiness #1.
(b). The secondary window with weightiness #2.
(c). (d). The pop-window with weightiness #3.



S: Selection
VO: Visual Output

Figure 5.13 The formation of the farther large-size interface.

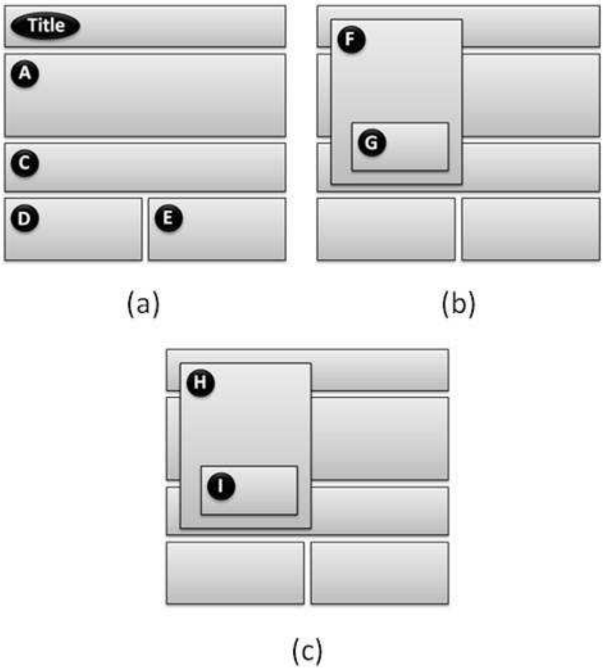


Figure 5.14 The layout of the farther interface.

(a). The primary window with weightiness #1.

(b). (c). The secondary pop-window with weightiness #2.

In this way, the specific views for the nearer interface and the farther interface with the RTIS are created respectively and are appropriate for the change of distance.

5.5 User Study

5.5.1 Questions for Hover Gesture

We use the hover gesture as the input. When people interact with a wearable interface based on the camera-projection techniques, a number of unexpected behaviors and noise gestures exist in the background. Thus, it is important to define exactly and implicitly the gestures to avoid ambiguity as well as to ensure the return and restart mechanism. The selection signal of hover gesture can be generated via a time period. When the user points at a button, he needs to remain in the position of the button for a certain time. On the one hand, interaction sensitivity can be decreased by this longer period of time, and the time spent waiting is likely to result in user inconveniency and impatience. But on the other hand, the interaction efficiency is susceptible to too short interval. Thus, to support this hand gesture, it is essential to find out an appropriate hover time interval. In order to find a bearable area and a satisfactory value for hover time as shown in figure 5. 15, we studied the subjective feeling of participants on the trade-off between speed and accurate rates. We explored the four research questions as follows:

- Questions 1: which time is the user’s bearable maximum for the shortest hover time?
- Questions 2: which time is the user’s bearable minimum for the longest hover time?
- Questions 3: which time is the user’s most satisfactory hover time?
- In addition to these three questions, we also asked the fourth question: is the hover gesture easy to learn or not?

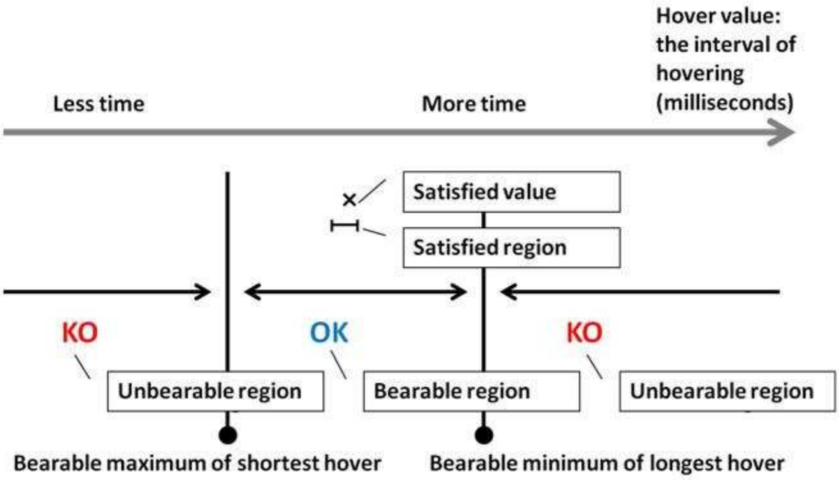


Figure 5.15 The bearable and unbearable region for hovering.

5.5.2 Questions for Scalable Interface

We use the hierarchical layout filter to process application tasks with our RTIS. We explored the following three research questions:

- Question 1: is there any difference between the farther large-size interface and the nearer small-size interface with scalability on the task completion time? With the pointing interactive elements of the same reference-cell size, the small interface folds some of the selection items. We assumed that it would take more time for the user to interact with multiple windows. Although the large interface has just one single window or fewer windows than the small interface, it would still take more time to focus on the target element due to the longer pointing distance.
- Question 2: which scale of interface do the users prefer?
- And to go a step further, we now ask the third question: in which situation do users prefer to use the small interface, and in which situations do they prefer to use the large interface?

5.5.3 Participants and Procedure

Our evaluation consists of two steps and we provided two programs (the hover gesture application and the RTIS) to be tested: the former is for the test of hover gesture and learning, while the latter is for the test of the scalable interface. In the first step, we organized the hover gesture test with the nearer projected interface. The hover gesture application consists of two parts: an upper part called the number panel for interacting, and a lower part called the control panel for controlling hover time. Via the number panel, the user can select any number to interact, and can obtain the feedback from the text view area. Via the control panel, the user can increase or decrease numbers in the text view area to speed up or slow down pointing speed. The default number of frames is 30, corresponding to 1363 milliseconds (The frame rate is 22 frames per second.). During this test, the user can try several times to adjust and find his/her shortest hover time, longest hover time and the most satisfactory time. This test includes the learning process; the user can use the number panel for practicing. We also provided help for users by discussing with them and demonstrating. To answer the hover gesture questions, we recruited 12 participants, including 10 males and 2 females, in the experiment. Participants were aged between 22 and 48 with an average age of 28.8. Their heights ranged from 165 cm to 185 cm with an average of 173.1 cm. All participants had experience in using mobile devices, but only 7 of them had knowledge of HCI. All of them except one were right-handed. This evaluation step began with an explanation of the protocol by text form. We explained orally when users had questions or problems. The questionnaire attached in the protocol contained two parts: the first part covered the background information on their familiarity of mobile devices and HCI, as well as basic data of individuals, to be answered by users before the test; the second part provided some questions in Likert scale form (Likert, 1932), to be answered by users during and after the test.

Next, in the second step, we started the scalable interface test with the RTIS. To answer the scalable interface questions as stated above, we selected 10 participants, 6 males and 4 females, who had taken part in the hover gesture experiment. Participants were aged between 23 and 28 with an average age of 26.5. Their heights ranged from 164 cm to 176 cm with an average of 170.9 cm. All participants had experience in using mobile devices, but only 5 of them had knowledge of HCI. All of them were right-handed. We asked each participant to start the interaction at a nearer or farther distance: this would be detected automatically. After completing the task with one interface, the user switched automatically to another one and performed the same task. The order of interacting with two scale interfaces was counterbalanced with a 2×2 balanced Latin Square. All participants performed the tasks respectively. Finally, we also asked participants to answer the questions in Likert scale form and share their subjective opinions during this evaluation.

5.6 Results and Findings

5.6.1 Results on Hover Gesture

We recorded the user's bearable maximum for the shortest hover time, the bearable minimum for the longest hover time and the most satisfactory hover time. As shown in the figure 5, we found that the interval from 7 to 10 (from 318 milliseconds to 455 milliseconds) can satisfy all participants' needs. Also the mean value of the most satisfactory hover time is 7.25 (330 milliseconds). Thus we select one value within this interval as the hover time of the scalable interface.

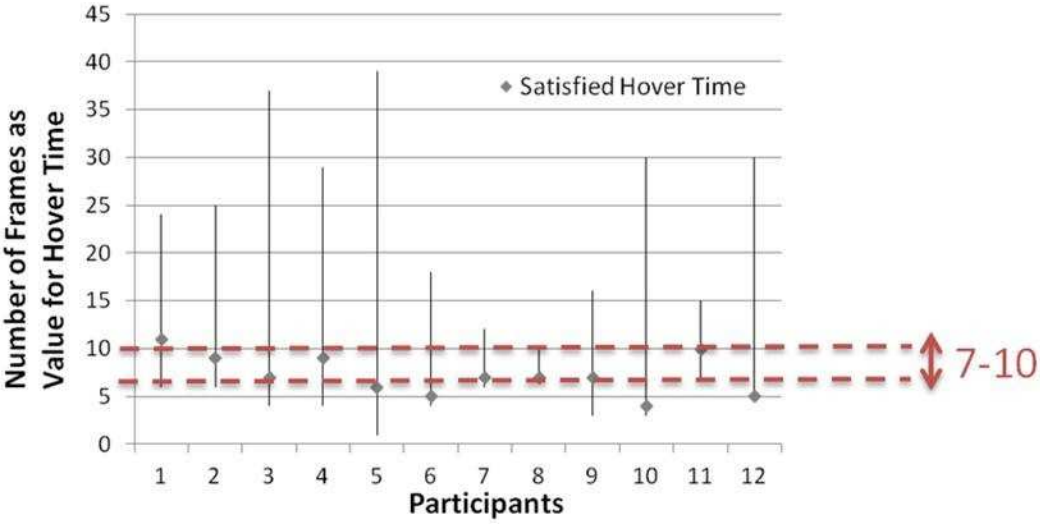


Figure 5.16 The hover time of each participant.

To obtain subjective opinions technically, we asked participants to respond to the Likert questionnaire items with respect to easiness of learning of the hover gesture. We had five levels (1-Strongly disagree, 2-Disagree, 3-Neither agree nor disagree, 4-Agree, 5-Strongly agree) to describe the easiness of learning. It showed that all the participants thought it was not hard to learn (The mean scores is 4.18). Two participants reported long time involved in lifting their arms and the unsteadiness of their fingers. Also, two participants commented that the shadow would at times occlude some buttons and feedback. Only one person reported that the finger occludes the buttons.

While obtaining the bearable maximum value of the shortest hover time, 7 participants stated that stopping bearing is for the reason of making the mistaken pointing actions under the too fast hover time. The rest of them stated that even if they were able to interact with the fast speed of hover time, they were worried about accidentally carrying out false pointing. While obtaining the bearable minimum value of the longest hover time, three subjects could not wait longer as they were both losing patience and their fingers and arms were feeling tired. Also, three subjects expressed that they stopped bearing it was only because they were losing patience, while the rest of them (six subjects) said it was due to their tried arm.

5.6.2 Results on Scalable Interface

5.6.21 Interaction Time

To know whether there were any significant interaction differences between the nearer interface and the farther interface, we used the Mann Whitney U test (Lehmann, 2006) of the nonparametric tests. There was no statistically significant differences ($p = 0.761 > 0.05$) between the nearer interface and the farther interface under the condition of the same task performed. The mean interaction time of the nearer interface is 29.89 seconds, while that of the farther interface is 29.45 seconds. Figure 5.17 shows interaction time with the nearer interface and the farther interface of 10 participants.

5.6.22 User Preference

To obtain subjective opinions technically, we asked participants to respond to the Likert questionnaire items concerning easiness of interacting with the nearer interface and the farther interface, and the switch between the two interfaces. It showed that all participants thought it was not hard to interact and switch (The mean score of easiness of interacting with nearer interface is 4.3, and the mean score of easiness of interacting with farther interface is 3.6, and the mean score of switching between the two interfaces is 3.9). Concerning their preference for the nearer interface or the farther interface, 7 participants reported that they preferred the nearer interface, while 3 participants expressed that they preferred the farther interface. Participants had a preference for the nearer interface because they can acquire a touch-like experience of the touch screen, and they do not need to carefully aim at the targets compared with the act of aiming with the farther interface. Participants preferred the farther interface due to the larger display and larger size of interaction items such as the buttons. Also, two participants reported that the shadow generated by the hand occluded some of the interaction items with the farther interaction.

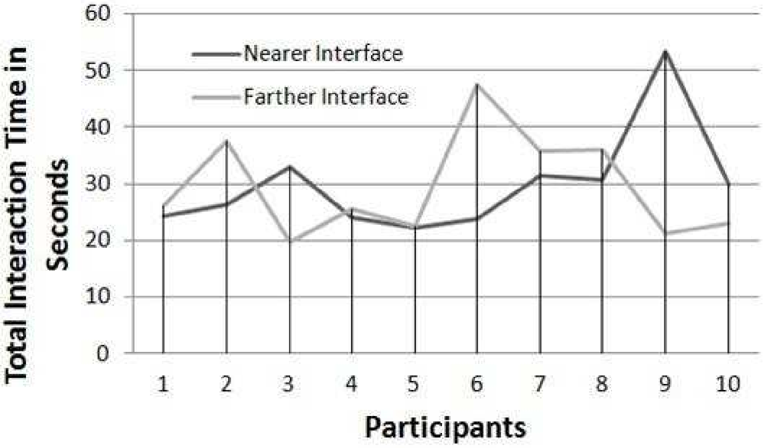
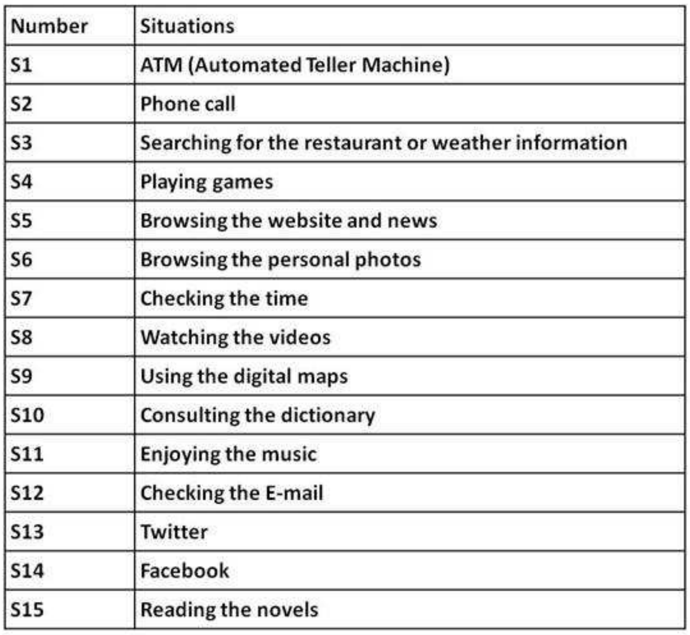


Figure 5.17 The interaction time of two interfaces by participants.

5.6.23 *User Comments on the Situations*

To go one step further, we provide a list with fifteen daily interaction situations in the questionnaire (see Figure 5.18) to study in which situation participants will prefer to use the small nearer interface, the large farther one or both of them (see Figure 5.19). We also asked them to give the reasons for their choice.



Number	Situations
S1	ATM (Automated Teller Machine)
S2	Phone call
S3	Searching for the restaurant or weather information
S4	Playing games
S5	Browsing the website and news
S6	Browsing the personal photos
S7	Checking the time
S8	Watching the videos
S9	Using the digital maps
S10	Consulting the dictionary
S11	Enjoying the music
S12	Checking the E-mail
S13	Twitter
S14	Facebook
S15	Reading the novels

Figure 5.18 The 15 situations for interaction.

For S1, all the participants stated that protecting their privacy was the reason for their choice. For S2, all the participants chose the nearer interface because of privacy or their customary use of traditional phone keyboards. For S3, participants who chose the farther interface stated that the larger projected interface could support more information while at the same time searching and sharing with others, while those who chose the nearer interface explained that the direct input method is more accurate for searching compared with the indirect input method, and it is easier to glance through the small-size interface to obtain all the results. For S4, some participants prefer the nearer interface because of the accurate direct input approach, while some prefer the farther one because of the larger display. For S5, participants chose the nearer interface because they preferred reading the text at a nearer distance, while the farther interface attracted some people because of the larger display and the possibility of sharing views with others. For S6, those who chose the nearer interface stated that they preferred reading personal information at a nearer distance due to privacy and accuracy, while the farther interface attracted some people for the same reason as in S5. For S7, those who chose the nearer interface wanted to watch the time clearly, while those who chose the farther interface preferred glancing quickly with a larger display. Those who chose “both” explained that they felt no difference between the nearer interface and the farther interface with S7. For S4, S6, and S7, the number of people choosing the different interfaces is nearly the same. For S8, nearly all chose the farther interface simply because of the larger display. For S9, the farther interface attracted participants for the same reason as in S8, while two participants preferred the nearer one because they could grasp information clearly and carefully. For S10, nearly all selected the nearer interface because they preferred reading the consulting results carefully. For S11, those who chose the nearer interface needed to select items carefully, while those who chose the farther interface stated that they preferred using the larger display without attracting attention. For S12, participants wanted to protect their privacy and read the details of e-mails via the nearer interface. For S13 and S14, participants are accustomed to the small surface of the mobile phone using Twitter and Facebook. For S15, those who chose the nearer interface like acquiring the details in novels. Those who chose the farther interface in S12, S13, S14 and S15 wanted to obtain a larger display experience. Two participants stated that it is easier to find a public wall than a private space.

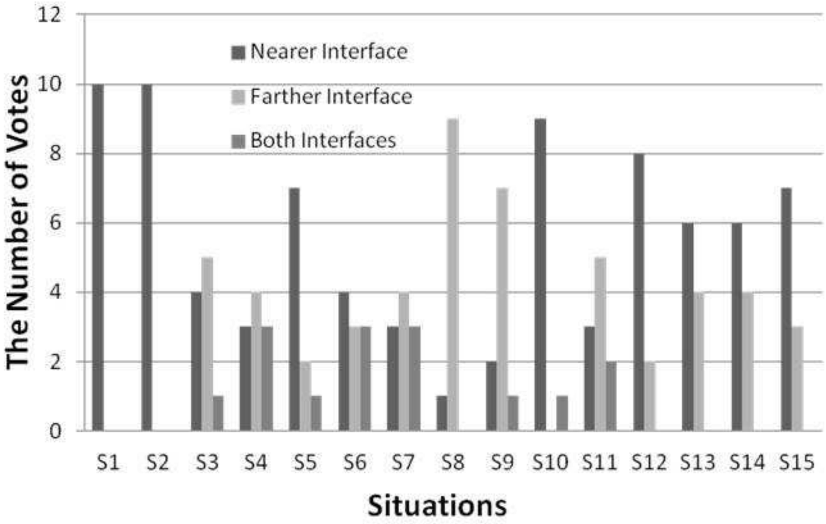


Figure 5.19 The situation interface votes.

5.7 Discussions

The findings in the hover gesture study showed that there is a common hover time in wearable pointing action. However, we cannot exclude that special hover time escapes from the common range. We also found that the interval from 318 milliseconds to 455 milliseconds can satisfy all participants' needs. The research (Müller-Tomfelde, 2007) recommended a feedback delay time for manual pointing actions of approximately 350 to 600 milliseconds as a starting point for the development of interactive applications. Overly fast and slow hover times could result in mistakes and excessive attention. A preliminary evaluation of the scalable interface did not show any significant difference between the nearer interface and the farther interface. We also observed that participants performed well with the offset cursor and the non-offset cursor. In the aspects of 15 situations, S3, S4, S6, S7, S11, S13, and S14 have the same applicability for both the nearer interface and the farther interface. S1, S2, S5, S10, S12 and S15 are more applicable for the nearer interface, while S8 and S9 are more applicable for the farther interface. Thus, we will focus on the applications possessing dual applicability, and design the scalable interfaces on the basis of these applications with more than two threshold values.

In addition, based on the user commands, we found that the nearer interface can provide a phone-like experience, increased efficiency for selection, a comfortable reading visual field and fewer disturbances for privacy, while the farther interface can provide a larger display experience and the possibility of sharing.

5.8 Summary

In this chapter, we presented our approach of the wearable scalable interface and the hover input gesture. We also described the PlayAllAround system: a wearable camera-projector system allowing mobile interaction to concretize the Environment Independent Interface (EII) concept. The user can access to the digital information from the projected scalable interface, which provides both the nearer small-size interface and the farther large-size interface supporting private and public use. We proposed the design of a reference-cell in the field of ergonomics as well as the principle based on decomposition of the application tasks to allow a scalable interface design, with the aim of providing enhanced user experience. Finally, we described the evaluation methods and results of the hover gesture and the scalable interface from both the qualitative and quantitative points of view. In a conclusion, as shown in figure 5.20, we investigated the camera-projector device unit, the hover input technique, EII, the digital interface, and the chest point in this chapter.

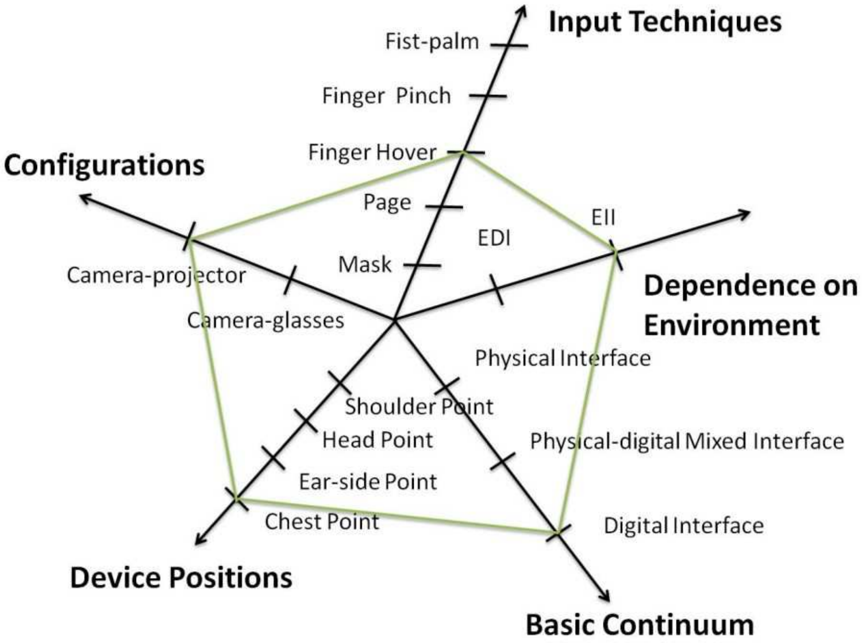


Figure 5.20 The spider figure of the wearable scalable projected interface.

In the next chapter, we will discuss our interaction techniques in both the stationary and mobile settings. Existing wearable camera-projector hand input research work focus more on the investigation in the stationary settings, which can not satisfy the requirements of interaction in the sophisticated everyday life, especially when people are walking or moving. We will present the design of two finger gestures in chapter 6: the hover gesture for pointing interaction and the pinch gesture for pointing and drag-drop interaction. We will also describe the design and the implementation aspects of bare hand gestures and our wearable system. The benefits and limitation of the head-worn configuration of the camera-projector device unit will be explored. Furthermore, to investigate how the user might interact with this system in both the stationary and mobile settings, we will compare the interactions of two finger gestures and projection output under three situations, namely standing, sitting and walking. Finally, we will describe the evaluation methods and results from both the qualitative and quantitative points of view.

6 Wearable Interaction Using Hand Gestures

6.1 Introduction

Although the projected interface has put an end to the limitation of the small screen display and enabled the scalable presentation of information, the floating property and focus variation of the projector make interaction different from the traditional display in both the stationary and mobile settings. Existing wearable camera-projector hand input research work focus more on investigation in the stationary settings, which is unable to satisfy the requirements of interaction in the sophisticated daily life, especially when people are walking or moving. To investigate the effective hand gesture input and projection output in a true ubiquitous environment, we propose the wearable camera-projector system. This system aims at putting into practice our concept of the environment independent interface (EII), which focuses on supporting people to access to their personal data as well as public resources at any time and in any place. In this chapter, we present the design of pinch gesture and bare hand gesture interaction. We employ the pinch gesture for pointing, drag-drop action and painting, and the design of bare hand interaction. We also explore the benefits and limitation of the head worn configuration of the camera-projector device unit. Furthermore, to investigate how the user might interact with this system in both the stationary and mobile setting, we compare the hover gesture interaction and the pinch gesture interaction, and also evaluate the projection output in three situations, including standing, sitting and walking. Finally, we describe our evaluation methods and results from both the qualitative and quantitative points of view.

In this chapter, we explore in greater depth the camera-projector system with EII in the true mobile settings. Our approach is implemented by employing the following wearable configurations: a pico-projector, a webcam and a wearable computer-like tablet used only for calculating and computing. We stabilize the camera-projector device unit next to the ear, the projection image of which can move as the head moves, and closely follows eye movement. In the current stage, we primarily concentrate on two input techniques: the hover gesture and the pinch gesture. Besides the solution of hovering for a second to select items, we propose also pinch gesture input to navigate the interface, which provides the selection and drag-drop action. We also extract the size of the projected area for automatically calibrating and providing the non-offset experience of the cursor for hover and for pinch gesture. To analyze the feasibility and appropriateness of mobile interaction, we conducted an evaluation by comparison with hover gesture and pinch gesture in three stationary and mobile settings: sitting, standing and walking. We discuss interaction time, the average selection time, interaction errors, as well as users' preferences. These findings imply the importance of interaction based on the hand gestures input and projected output in a mobile situation rather than only in a stable state.

6.2 Interaction Techniques

6.2.1 Pinch-gesture-based Interaction

When the user works with a WIMP system using a mouse, he/she has the feeling of the physical "pointing". However, when he/she is interacts with a wearable interface based on computer vision technique, he/she has neither the feeling of touch or contacting, nor the feeling of haptic feedback. One solution is to let the user's finger hover for a second. The selection signal can be generated via a span. When the user points at a button, he needs to remain in the position of the button for a period of time. In this way, the button is considered as selected. Interaction sensitivity may thus be limited by this hover time. However, the time spent waiting is likely to prove inconvenient for the user.

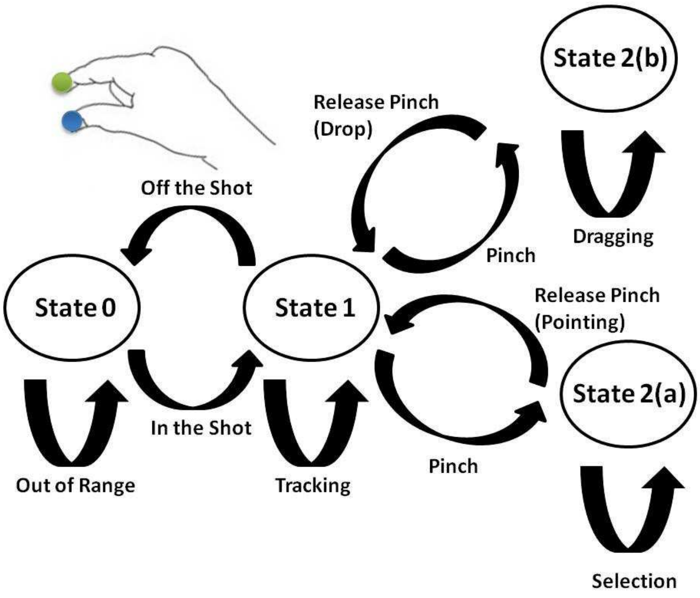


Figure 6.1 The four-state input model of the pinch gesture.

In this chapter, we propose a pinch gesture with a four-state input model, which are illustrated in figure 6.1. The first state, (state 0), is what we will call out of rang. In this state, the finger is beyond the reach of webcam's vision, so any movement of the finger has no effect on the system. As the finger enters into the region of the webcam, the system starts to track and the tracking symbol is the cursor, which corresponds with the center point of the two fingers in figure 6.2. State 2 has two states, which are dependent on the types of the widget. For example, selecting an object whose property of dragging is true (IsDrag = 1), will cause the selected object to be dragged, whereas selecting an object which can not be dragged, means that this object is clicked.

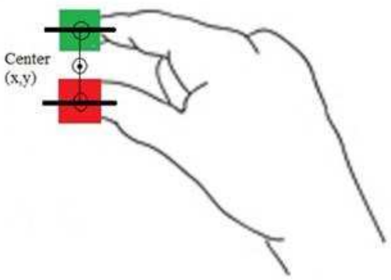


Figure 6.2 The illustration of the pinch gesture.

The standard pinch gesture is when the hand is parallel with the projected interface. The line between the tip of the index finger and thumb finger could be freely vertical or horizontal. Besides the selection and pointing action, navigations such as dragging, inking, pull-down menus require distinct motions between states. Therefore, compared with the hover gesture, the pinch gesture is able to support more tasks and support more interaction.

6.2.2 Hand-gesture-based Interaction

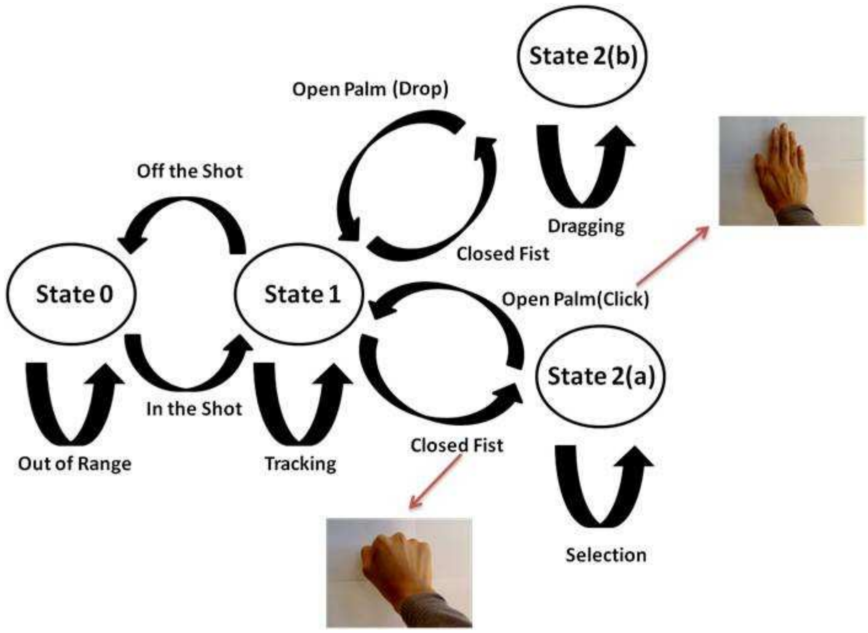


Figure 6.3 The four-state input model of the fist-palm gesture.

Rather than the solution of hover gesture or the pinch-release-pinch gesture to select the items, we propose a fist-palm gesture as input to navigate the interface. Figure 6.3 shows the set of gestures used, which include the full hand with the fingers together and the fist closed. As the four-state input model illustrates, the input mechanism of the fist-palm is similar to that of the pinch-release-pinch gesture. The palm gesture controls the movement of the cursor, while switching between the fist state and the palm state that determines the selection, the drag and the drop actions.

6.2.3 Items in Projected Interface

As chapter 5 illustrated, we define 20.0mm as the length of the reference-cell for the hover gesture; the smallest possible size of the pointing icon is limited as a square of 20.0mm×20.0mm. Furthermore, the pinch gesture employs two fingers to complete the pinch and release-pinch action. Thus, we define the double DIP width as the length of the reference-cell for the pinch gesture, namely 40.0mm×40.0mm (see Figure 6.4).

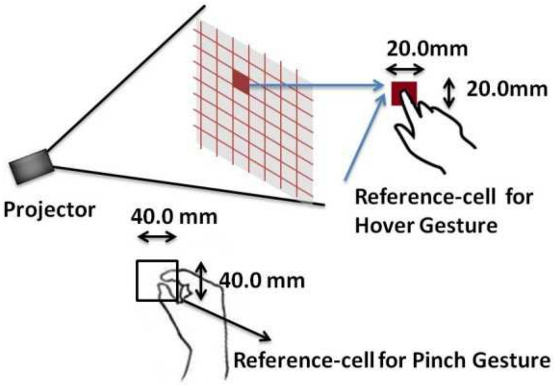


Figure 6.4 The reference-cell for the pinch gesture compared with that for the hover gesture.

6.3 Wearable Configuration and Implementation

In this chapter, our wearable configuration still contains a webcam, a pico-projector and a tablet for calculating. However, the device position is placed on the ear instead of the chest position in chapter 5.

6.3.1 Head-worn Configuration

The camera-projector unit in this chapter is still composed of a RGB 640×480 webcam and a pico-projector (see Figure 6.5(a)). We combine the camera and projector as a complete assembly and fix the unit next to the user's right ear using a plastic head band (see Figure 6.5 (b)). In this way, the camera sees what the user sees as the user turns the head, and the projector displays digital information precisely in the user's field of vision. As the user turns his/her head, the projected interface follows the required direction. The pico-projector has a resolution of 640×480 pixels and a projection size (diagonal) of 127cm maximum and of 15cm minimum. Size is 90mm×63.5mm×24.5mm with a weight of 117g, and in the manual focus mode. Compared with the overhead position, the close ear position supports the same height of the user's field of vision, thus reducing the possibility of image distortion and the inconvenience of moving eyes upward.

(a)

(b)

Figure 6.5 The wearable configurations.
(a). The camera-projector unit.
(b). The head worn configuration next to the ear.

6.3.2 Auto-calibration of Projector and Camera Coordinates

We continue to employ the same method in chapter 6 to correct disunity between the coordinate system of the camera and the projected surface (see Figure 6.6(a)). The contour of the lightest area is as shown in figure 6.6 (b). In this way, we unify the coordinates of the camera and the projected window, and can provide the non-offset cursor for the hover gesture, the pinch gesture and the fist-palm gesture.

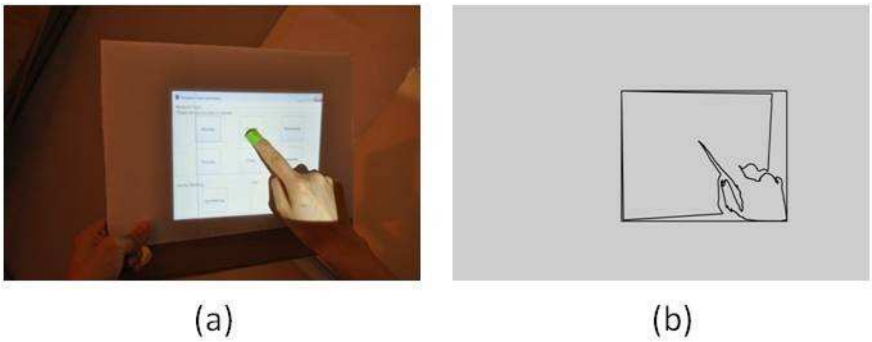


Figure 6.6 The lightest area process.
(a). The original projected image.
(b). The contour of the lightest area.

6.3.3 Recognition of Hover Gesture and Pinch Gesture

Our research is based on the computer vision-based hand tracking method. We fix the unique color markers on the tips of the index finger and the thumb, which can be tracked by the Camshift algorithm (G. R. Bradski, 1998) in real-time. We record the trace of the color markers and if the tracking points meet our predefined condition, we regard this action as a pointing for the hover gesture and the pinch gesture or as a drag-drop action for the pinch gesture.

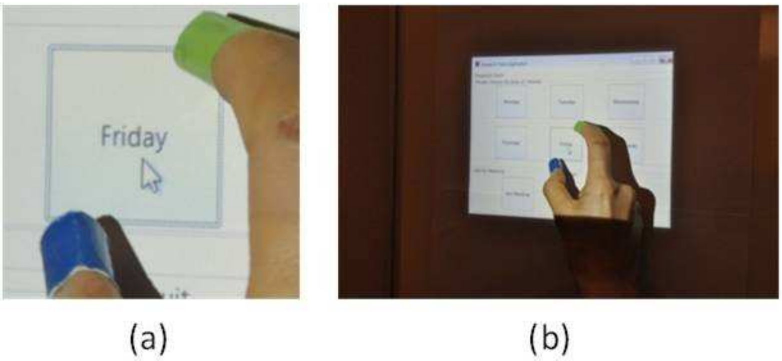


Figure 6.7 Pinch gesture interaction.
(a). The pinch gesure.
(b). The position of the cursor.

For the pinch gesture, we set the position of the cursor in the middle of the two fingers (the index finger and the thumb, see Figure 6.7 (a)).

Since the reference-cell for the pinch gesture is large enough to enclose the two tips of the index finger and the thumb, the user only needs to carry out the pinch action within the button as shown in figure 6.7 (b).

For the drag-drop action, the user only needs to carry out the drag action within the button as shown in figure 6.8 (a). Also, after dragging a distance, the user carries out the dropping action in the target area as shown in the figure 6.8 (b).

Figure 6.8 Drag-drop action interaction.
(a). The drop action.
(b). The drop area.

For the hover gesture, we set the position of the cursor on the tips of the index finger (see Figure 6.9 (a)). Since the reference-cell for the hover gesture is large enough to enclose the tip of the index finger, the user only needs to hover on the button for a while as shown in figure 6.9 (b). Frame rate is 30 frames per second, so we set 20 frames as a selection; hover time is 667 milliseconds per pointing.

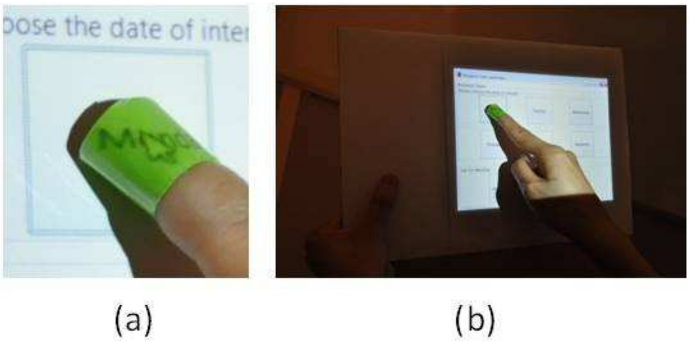


Figure 6.9 Hover gesture interaction.
(a). The non-offset cursor for hover gesture.
(b). The hover gesture.

6.3.4 Recognition of Fist-palm Gesture

Our research is based on the computer vision-based hand tracking method. We propose an alternative approach for solving the problem of hand segmentation and gesture recognition. We extract hands and their geometric features via image preprocessing, and track the input hand.

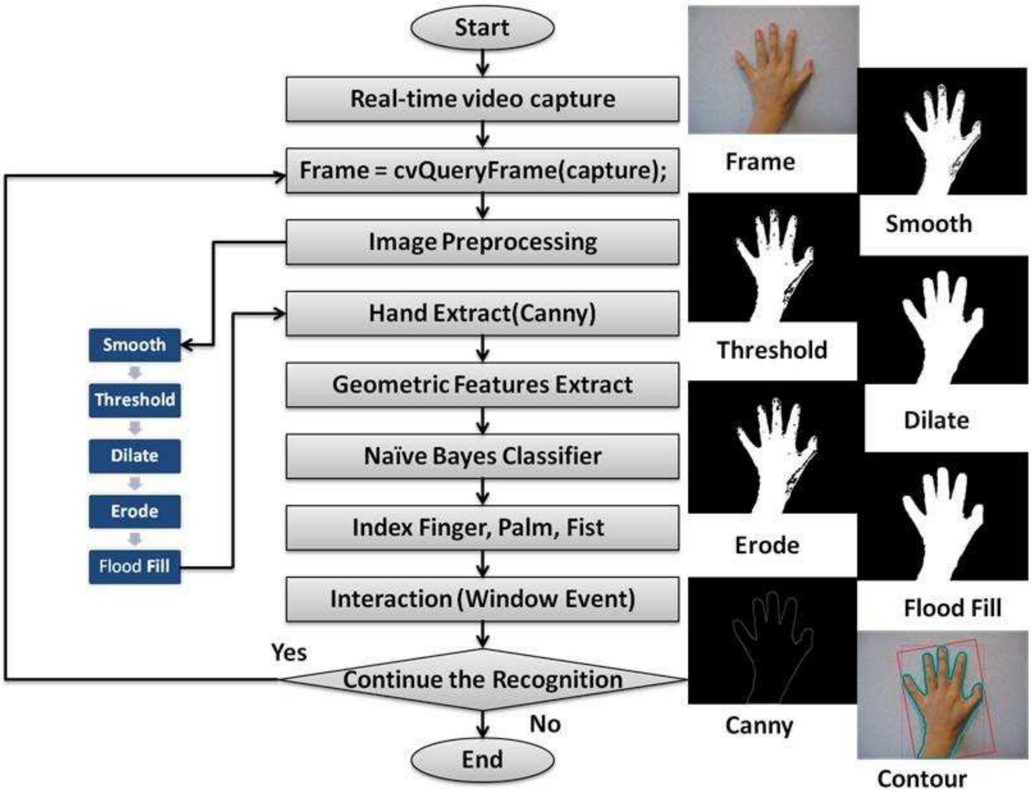


Figure 6.10 The recognition process of the fist-palm gesture.

As illustrated in figure 6.10, we propose a flow that with seven main steps as follows:

- The real-time video is captured from a webcam. The video information (frame speed, image format and image size, etc.) is also obtained in this step.
- The second step is to preprocess the image. To remove image noise, we use a series of preprocess steps including making the image smooth, applying the fixed-level threshold to the grayscale image, using dilate and erode operations, and employing flood fill for further analysis, etc.
- We then extract the hand by canny algorithm and obtain the contour of the hand.

- Based on the contour of the hand, we use geometric features to recognize hand gestures. We select two features to distinguish three hand gestures: the index finger, the open palm with all fingers together, and the closed fist. The smallest rectangle surrounding all points of the contour plays an important role in these two features (see Figure 6.11). We employ the ratio of the hand contour area to the smallest rectangle area as the first feature. The second feature is regarded as the ratio of the smallest rectangle height over its width. We thus obtain the two geometric features.
- Next, we collect the samples of the three hand gestures and train our naïve Bayes classifier based on the two geometric features. In this way, when real-time data are obtained, the classifier can recognize the type of hand gesture.
- Afterwards, interactive items can be navigated on the basis of the hand gestures.
- In addition to the training part of the classifier, the other parts are all in the image and recognition analysis loop.

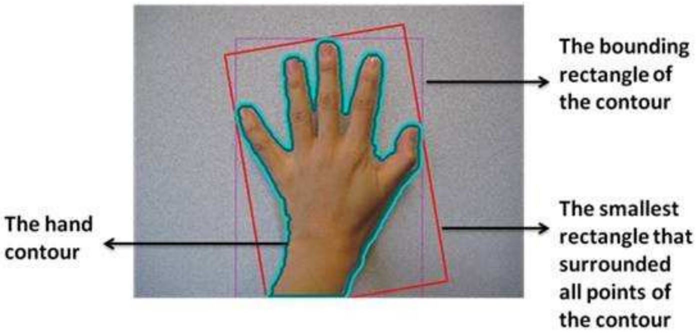


Figure 6.11 The hand contour and the smallest rectangle.

In this way, the user can interact with one of his/her bare hands to navigate the interactive items, which are projected on the arbitrary surface. We define that the cursor is located at the center of gravity of the hand. The user moves his/her palm to select and switch to the fist gesture to validate his/her selection. In brief, we use the predefined gestures, as we stated above, to perform the operation of selection, dragging and dropping, painting and scrolling.

6.4 Prototype of Research Team Interaction System (RTIS)

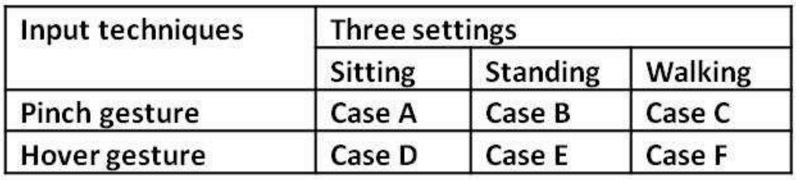
We use the application of Research Team Interaction System (RTIS) to prove our concept of EII with the pinch gesture and the hover gesture in three settings. The scenario has already been described in the previous chapters. The user can select interactive items in the menu using either hover gesture or pinch gesture. To carry out the drag-drop task, we only propose pinch gesture input. The user can drag the items and drop them into the right region by pinch gesture.

6.5 User Study

To obtain a more thorough understanding of our input and output modality in the stationary and mobile settings, we organized a structured evaluation to compare our two input techniques (hover input and pinch gesture) in three situations: sitting, standing, and walking. We explored the following three main research questions:

- Whether the pinch gesture and hover gesture are easy to learn and use for users in the stationary and mobile settings?
- Whether or not the mobility will influence the interaction?
- What are the advantages and disadvantages of using the pinch gesture input technique compared with the hover gesture input?

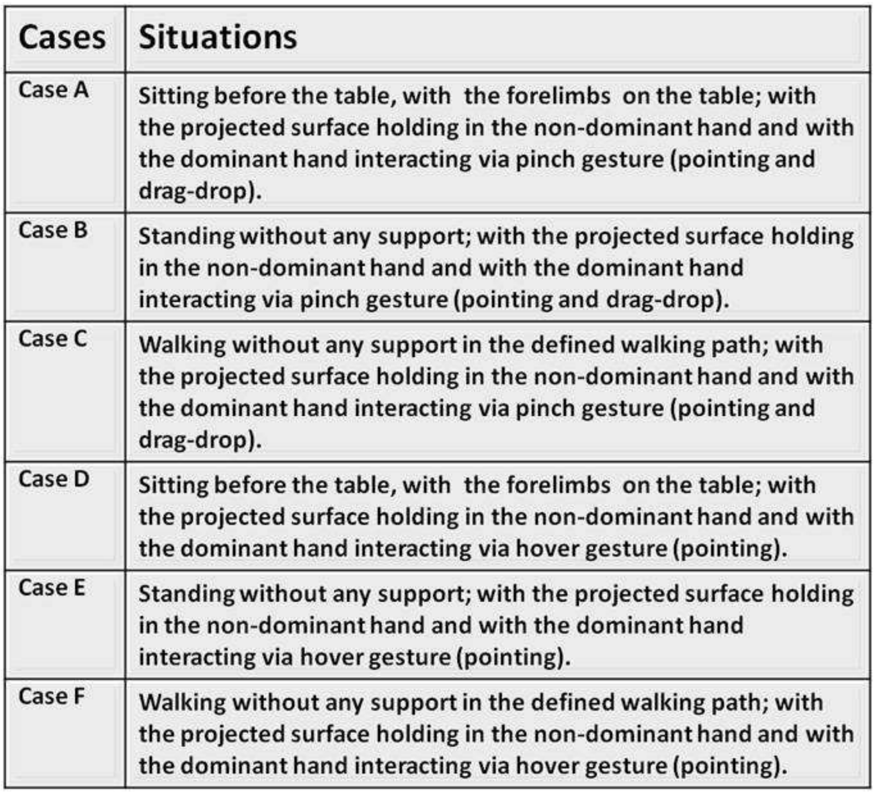
To answer the above questions, we evaluated the two input techniques in three situations to form six cases as shown in figure 6.12. We employed a within-subjects design, in which all participants used both the two techniques in all three settings; each participant would perform all six cases (see Figure 6.13). The order of the cases was counterbalanced with a 6×6 balanced Latin Square.



Input techniques	Three settings		
	Sitting	Standing	Walking
Pinch gesture	Case A	Case B	Case C
Hover gesture	Case D	Case E	Case F

Figure 6.12 The six cases with two gestures in three settings.

In cases A and D, we asked participants to sit freely in front of a table, which could provide support for their hands and arms, thus alleviating the tiredness. In cases B and E, we instructed participants to stand freely to simulate the stationary state without any support for their forearms. In cases C and F, we instructed participants to walk while learning and interacting with the aim of simulating true mobility without any support in the same room. In all cases, participants held white cardboard in their non-dominant hands to project the interface and employed their dominant hands to interact.



Cases	Situations
Case A	Sitting before the table, with the forelimbs on the table; with the projected surface holding in the non-dominant hand and with the dominant hand interacting via pinch gesture (pointing and drag-drop).
Case B	Standing without any support; with the projected surface holding in the non-dominant hand and with the dominant hand interacting via pinch gesture (pointing and drag-drop).
Case C	Walking without any support in the defined walking path; with the projected surface holding in the non-dominant hand and with the dominant hand interacting via pinch gesture (pointing and drag-drop).
Case D	Sitting before the table, with the forelimbs on the table; with the projected surface holding in the non-dominant hand and with the dominant hand interacting via hover gesture (pointing).
Case E	Standing without any support; with the projected surface holding in the non-dominant hand and with the dominant hand interacting via hover gesture (pointing).
Case F	Walking without any support in the defined walking path; with the projected surface holding in the non-dominant hand and with the dominant hand interacting via hover gesture (pointing).

Figure 6.13 The situations for six cases.

To simulate the normal walking state, we defined a walking path as shown in figure 6.14. The path was specified in a room. We asked participants to keep walking during the interaction but did not specify any specific speed. Participants walked as usual and in a casual way but without stopping. If a participant stopped without any reason, we asked him/her to continue walking.

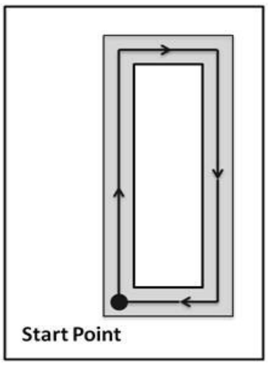


Figure 6.14 The walking path.

6.5.1 Participants

To answer the questions stated above and evaluate the utility of our prototypes, we recruited 12 participants, all students, including 9 males and 3 females. Participants were aged between 24 and 31 with an average age of 27.6. Their heights ranged from 167 cm to 178 cm with an average of 173.3 cm. All of them were right-handed.

All participants had experience in using mobile devices, but only one of them had no experience in multi-touch smart mobile device system such as the iOS system for iPhone and Android system. On the topic of Human-Computer Interaction (HCI), 5 of them said that they had not taken HCI courses or read any books on the subject. The other 7 had read the related books or taken classes in the courses of introduction for HCI.

6.5.2 Procedure

We provided two types of programs (toy application and true application) to be tested, each of which contained the six cases respectively. The toy applications were designed only for participants to learn and practice, but had the same interaction methods as the true applications. The goal of introducing the toy application was to help participants to familiarize themselves with interaction techniques and interfaces. Participants could play several times until they felt competent in carrying out the following true tasks. In the learning process, we demonstrated how to interact with the hover gesture, pinch gesture or the drag-drop action of the pinch gesture. Besides demonstrating, we also guided the users by discussing with them.

The evaluation began with an explanation of the protocol by text form. We also explained orally when users had questions or problems. The questionnaire attached in the protocol contained two parts: the first part covered background information on their familiarity with mobile devices and HCI, as well as basic data of individuals, to be answered by users before the test; the second part provided questions mainly in Likert scale form (Likert, 1932), to be completed by users after the test. We asked participants to complete the Likert scale after finishing all the trials and to select the reasons listed or to write down any unlisted ones. Once participants had completed the first part of the questionnaire and were ready to be tested, we let them enter the learning sector: to play the toy application. Next, after practicing several times, they were given the tasks by the instructor and started the actual test. In each case, participants played the toy application first and then accessed into the task part via the true application. All participants performed the tasks respectively. They were instructed to check one researcher's schedule and to ask for an appointment with this researcher as accurately and quickly as possible. Finally, we also asked participants to share their subjective opinions during this evaluation.

Each participant performed one task with 7 clicks of each input technique. The result was 42 trials per participant (2 input techniques × 1 task × 3 settings × 7 click trials = 42 trials). For pinch drag-drop action, each participant performed 12 trials (1 input techniques × 1 task × 3 settings × 4 drag-drop trials = 12 trials). Thus, each participant performed 54 trials in total. To summarize the design: 12 subjects × 54 trials = 648 trials.

6.5.3 Variables

Independent variables are two hand gestures (pinch gesture pointing and hover gesture pointing), two actions (pinch pointing and pinch drag-drop) and three settings (sitting, standing and walking). Dependent variables are interaction time, average interaction time, and task completion time. Interaction time contains the time of the correct operations. Moreover, to compare the action of pointing and drag-drop with the pinch gesture, we recorded the average interaction time. Task completion time is the time from the user's starting the task to his/her stopping, which also contains user's incorrect operations' time. Each input step and the time cells were automatically recorded by the system. We analyzed our logs to investigate the difference in time.

6.5.4 Errors

Through observation, we found that the reasons for errors while carrying out the tasks are mainly due to user locomotion and misunderstandings of tasks as we stated in chapter 4. These two errors were counted by observation and system logs in each case per participant. The error rate was calculated as the number of occurrences of the errors divided by all the operations for all 12 participants in each case.

Furthermore, we also conducted the qualitative test to obtain users' preferences for the two input gestures, and for the drag-drop action and the pointing action, easiness of learning the gestures and the drag-drop action, evaluation of the projected interface, as well as satisfaction concerning the operation of the head worn configuration.

6.6 Study Results

In this section, we present the quantitative and qualitative results of our study, including interaction time, average interaction time, task completion time, error rates as well as users' subjective preference, satisfaction and comments. To discover whether there are any significant differences between two input techniques in three situations, we used the Mann Whitney U test (Lehmann, 2006) of the nonparametric tests.

6.6.1 Interaction Time

Figure 6.15 shows mean interaction time with pinch pointing and hover pointing in three settings. In both the stationary and mobile settings, mean interaction time with hover gesture input is longer than that with the pinch gesture. For the pinch input technique, mean interaction time in the situation of sitting and standing are both (sitting: 11.833s, standing: 12.976s) shorter than that in the situation of walking (walking: 16.417s). For the hover input technique, mean interaction time under the situation of sitting and standing are almost the same (sitting: 15.308s, standing: 13.801s), and shorter than that in the situation of walking (walking: 21.336s).

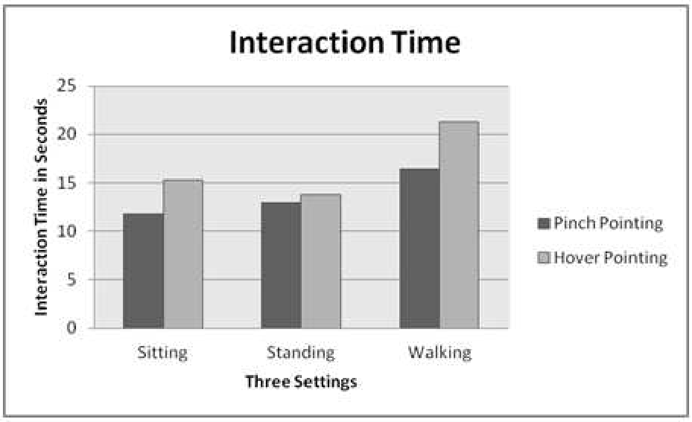


Figure 6.15 Interaction time of pointing by two gestures in three settings.

For interaction with pinch gesture and drag-drop action (see Figure 6.16), mean interaction time is almost the same (sitting: 15.308s, standing: 13.801s, walking: 21.336s) in the stationary and mobile settings. Analysis of the Mann Whitey U test on interaction time showed that there are no statistically significant differences (p>0.05) between sitting situation and standing situation, between sitting and walking, and between standing and walking with drag-drop interaction action.

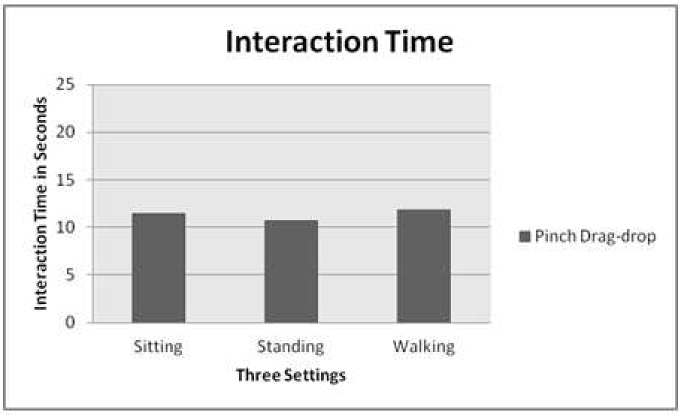


Figure 6.16 Interaction time of drag-drop action by pinch gesture in three settings.

6.6.2 Average Interaction Time

By comparing the average interaction time per participant, the Mann Whitey U test showed that there are statistically significant differences ($p<0.05$) between pinch pointing and pinch drag-drop in all three settings. The mean interaction time of the drag-drop action is longer than that of pointing with the pinch gesture in the three settings as shown in figure 6.17.

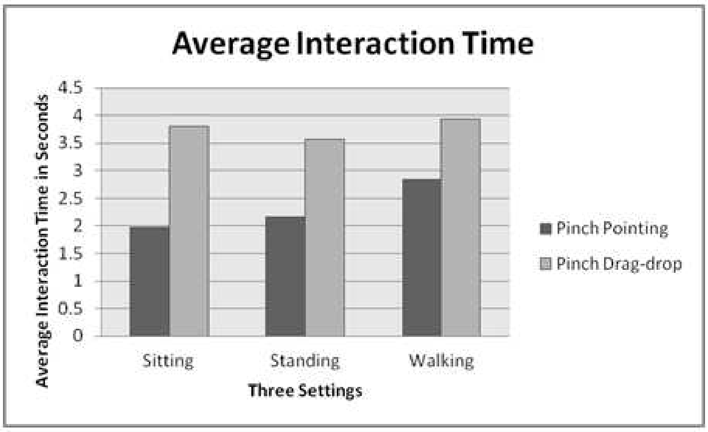


Figure 6.17 Average interaction time of pointing and drag-drop actions in three settings.

6.6.3 Task Completion Time

The result of the Mann Whitey U test indicated there are no statistically significant differences ($p>0.05$) between pinch input and hover input in all three settings. However, with the hover input gesture, there are significant differences ($p<0.05$) between sitting and walking, and between standing and walking. There is no significant difference ($p>0.05$) between the sitting situation and the standing situation with hover gesture. With pinch gesture, there are no differences ($p>0.05$) between sitting and standing, between sitting and walking, and between standing and walking.

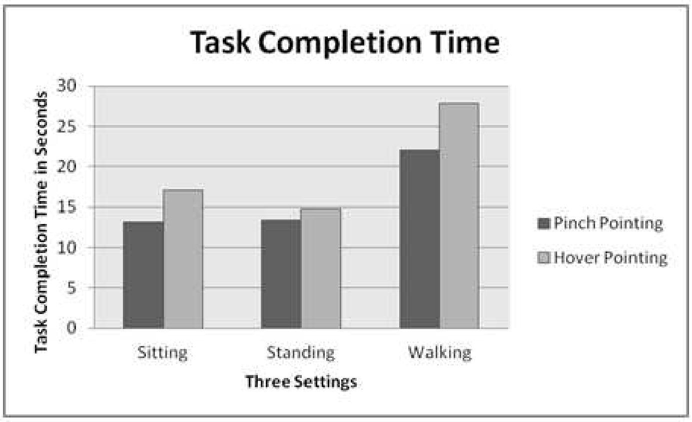


Figure 6.18 Task completion time for two gestures in three settings.

For the same task and same layout of the interface, the mean time for the task completion with pinch pointing is shorter than with hover pointing in the three settings (see Figure 6.18). With both input techniques, task completion time for walking (pinch gesture: 22.100s, hover gesture: 27.773s) is almost twice than for sitting and standing (pinch gesture: 13.160s for sitting and 13.333s for standing, hover gesture: 17.145s for sitting and 14.756s for standing).

6.6.4 Errors

As shown in figure 6.19, in all three settings, errors occur with the pinch gesture to the same extent or less than with the hover gesture. In the sitting situation, the error rate with pinch input is the same as the hover gesture. However, the error rate with hover pointing is more than twice as high as with pinch input in the standing situation. For the pinch gesture, the lowest error rate is in the standing situation (Error rate is 0%.). For the hover gesture, the lowest error rate is in both the sitting and standing situations. A relatively high error exists in the walking state for both the pinch gesture and the hover gesture. In the walking scenario, the error of the latter is nearly twice as high as the former (pinch input: 11.58%, hover input: 19.23%).

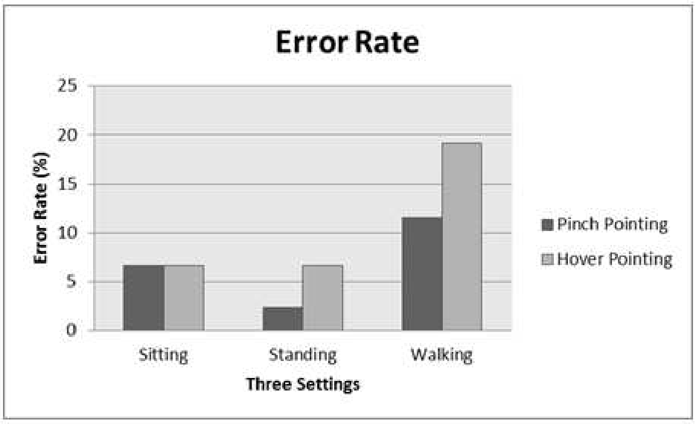


Figure 6.19 Error rates for two gestures in three settings.

We also recorded the reason for errors via observation. In the sitting situation with pinch gesture, only one participant made errors. All these errors were made due to an incorrect drop-release action. In case B with pinch pointing, only one participant made errors, ascribable to the incorrect posture of the pinch gesture. In case C with pinch pointing, three participants made errors due to locomotion reasons. In case D with the hover gesture, all errors were made by one participant as he forgot the task. In case E with hover gesture, one third of errors were made because the user forgot part of the tasks. The remaining errors were made because absence of instant feedback caused the user to hover longer to obtain visual change. In this way, he could easily falsely point interactive items in the following window when the latter is switching. In case F, all pointing mistakes were made because the projected window floated and waved excessively during walking.

In addition, we noted errors with drag-drop action as illustrated in figure 6.20. By comparing the error rates in the figure, we found that participants made less mistakes with drag-drop action than with pointing action in three settings respectively, using the pinch gesture. The error present in the walking scenario (4.00%) is slightly less than that in the sitting scenario (5.88%) with drag-drop action. In the sitting situation, errors were made due to forgetting the task, incorrect dragging, and incorrect dropping. No errors existed in the standing situation with drag-drop action. Errors occurred due to incorrect dragging in the walking scenario.

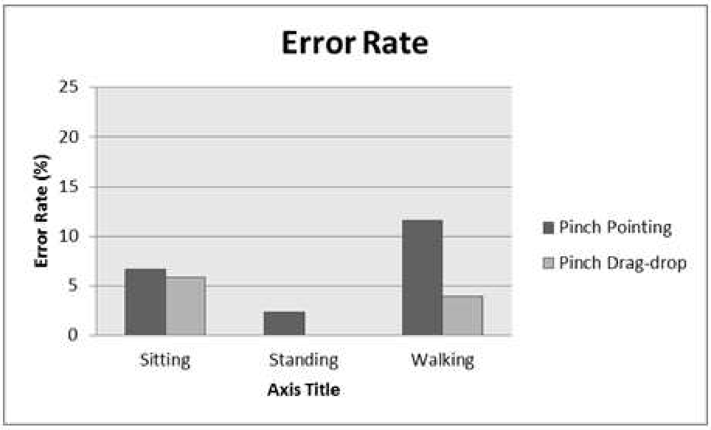


Figure 6.20 Error rates for pointing and drag-drop actions of the pinch gesture in three settings.

6.6.5 User Satisfaction and Preference

After the experiments, we asked participants to give their subjective preference and satisfaction.

To obtain subjective opinions technically, we asked participants to respond to the Likert questionnaire items concerning the easiness of learning for the pinch gesture and the hover gesture in three settings. We had five levels (1-Strongly disagree, 2-Disagree, 3-Neither agree nor disagree, 4-Agree, 5-Strongly agree) to describe easiness of learning and utilization. Figure 6.21 showed that all participants thought it was not hard to learn and utilize with two input techniques in three settings (The mean scores are all more than 3). After learning, all the true task operation scores are higher than the toy learning scores. Also, for both pinch gesture and hover gesture, the sitting scores are the highest and the walking scores are the lowest.

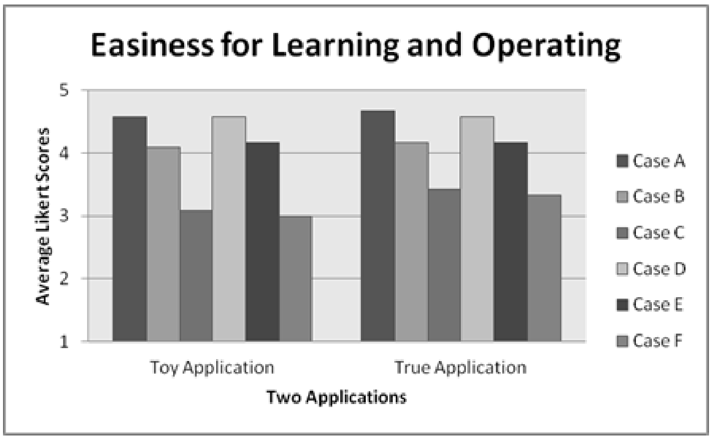
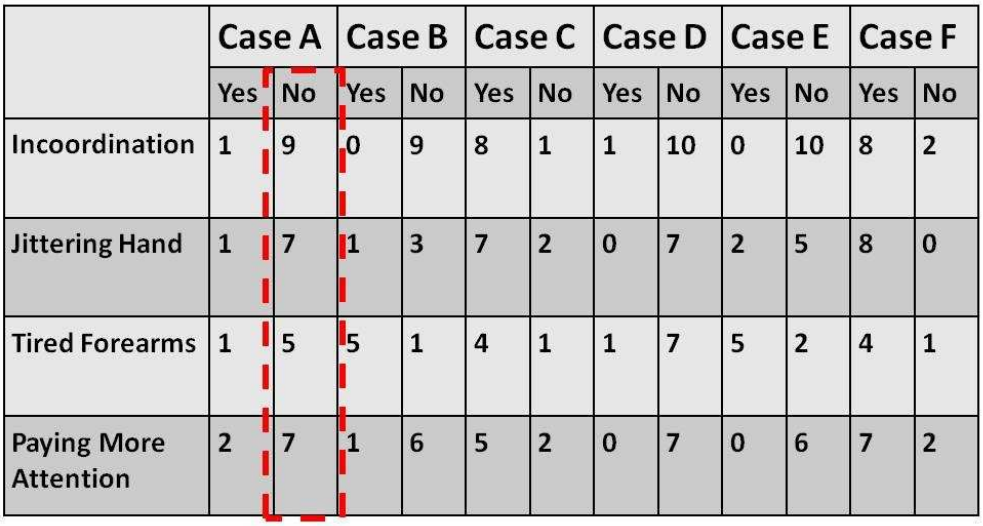


Figure 6.21 Easiness for learning and utilization.

In addition to asking participants to give the scores for the six cases, we also asked them to choose the reason for the scores from the list. From the questionnaire, we counted the number of participants for each reason as shown in figure 6.22. Our observations revealed four main reasons: incoordination, jittering hand effect, tired forearms, and extra attention paid. As the red dot frame shows, taking the pinch pointing action for sitting (case A) as an example, 9 participants reported that it was easy to coordinate the position of the projected interface and hands. 7 participants reported no jittering hand effect due to the table support table, and 5 participants reported no tired arms effect. In addition, 7 participants said that they did not need to pay extra attention to interaction. We found that the tendency of the number of participants in answering "Yes" and "No" in cases C and F is the opposite to that in cases A, B, D and E.



	Case A		Case B		Case C		Case D		Case E		Case F	
	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Incoordination	1	9	0	9	8	1	1	10	0	10	8	2
Jittering Hand	1	7	1	3	7	2	0	7	2	5	8	0
Tired Forearms	1	5	5	1	4	1	1	7	5	2	4	1
Paying More Attention	2	7	1	6	5	2	0	7	0	6	7	2

Figure 6.22 The number of participants for each reason.

We also asked participants to evaluate easiness of learning with the two input techniques, and to choose their reasons as well as to give comments. As shown in figure 6.23, the average score of easiness for learning the pinch gesture and the hover gesture is 4.333 and 4.250, respectively. From the questionnaire, 7 participants reported an easy pinch action and 6 participants reported an easy release-pinch action. Also, 5 participants stated that it was not tiring to pinch and only one gave the opposite opinion. Among them, the person who gave 2 as the score reported a tired finger, and the difficulty of dragging and dropping. For the hover gesture, 7 participants said it was easy to control hover time, while 4 participants said the opposite. Also, 6 participants reported that their arms were not tired.

Moreover, 4 participants commented that they should pay more attention to slipping the finger while 4 people said just the opposite. In addition, 5 participants said they had a touch experience when using hover gesture.

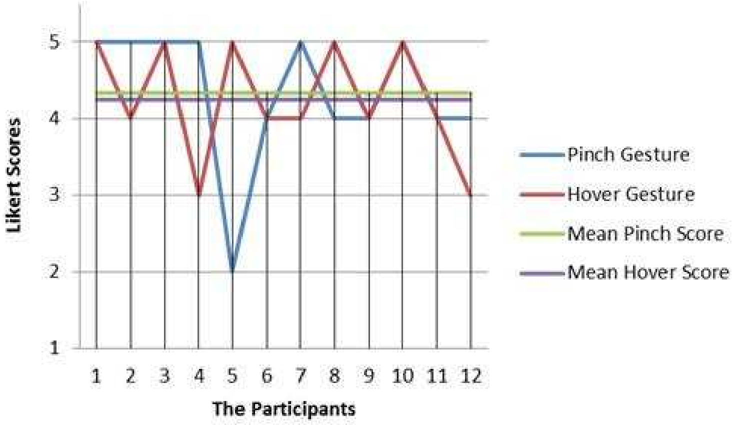


Figure 6.23 The Likert score of two gestures.

Figure 6.24 showed that the average score of easiness for learning the drag-drop action is 4.25, and that for utilization with drag-drop is 4.417. Based on the questionnaire, 10 participants reported it was easy to drag, and 9 participants thought that it was easy to drop. The low scores are mainly due to the difficulty of dropping, reported by 3 participants, and due to the difficulty on dragging, reported by one participant.

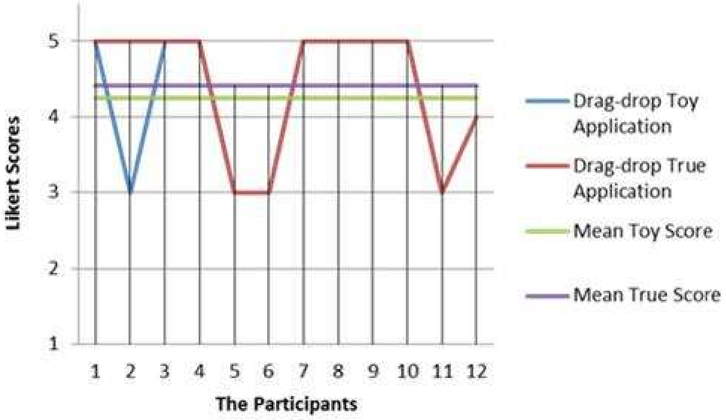


Figure 6.24 The Likert score of drag-drop action in the toy and true applications.

Most participants expressed a preference for the pinch gesture; they gave the pinch input twice as many votes as the hover one (see Figure 7.25). Most participants preferred the pinch gesture because it contains clear feedback, has a low possibility of false pointing, with no need to wait, whereas the hover gesture provides no feedback. The remainder preferred the hover gesture because it is simpler to interact with one finger by hovering.

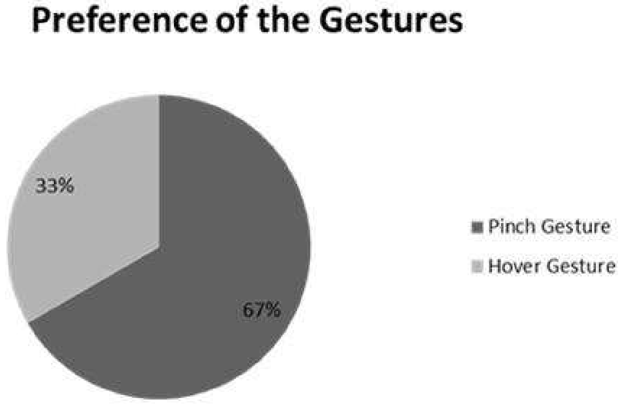


Figure 6.25 User preference of the two gestures.

For the drag-drop action and pointing via pinch gesture, 11 participants preferred the pinch pointing interaction (see Figure 6.26). Most participants expressed that it was uncomfortable to drag an item over a long distance, especially in the walking state, even though they had the ability to interact correctly.

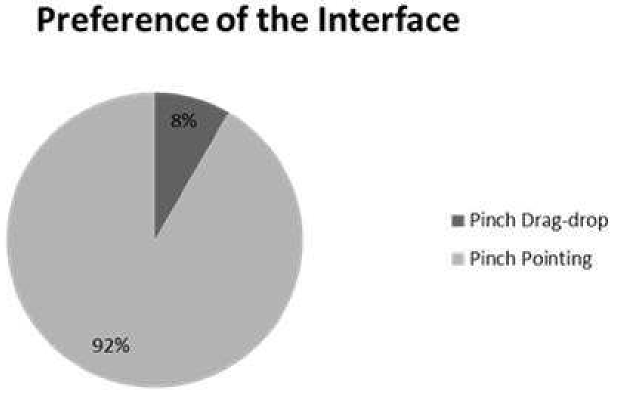


Figure 6.26 User preference of the two actions: drag-drop action and the pointing action.

For the purpose of studying the projected interface, we asked participants to give a score for the projected interface according to their satisfaction, and to select the reasons listed as well as write down any unlisted ones. The average score of the projected interface is 4.333 as shown in figure 6.27. From the questionnaire, we learned that half participants thought the manual focus could impact interaction, while the other half thought there was no impact.

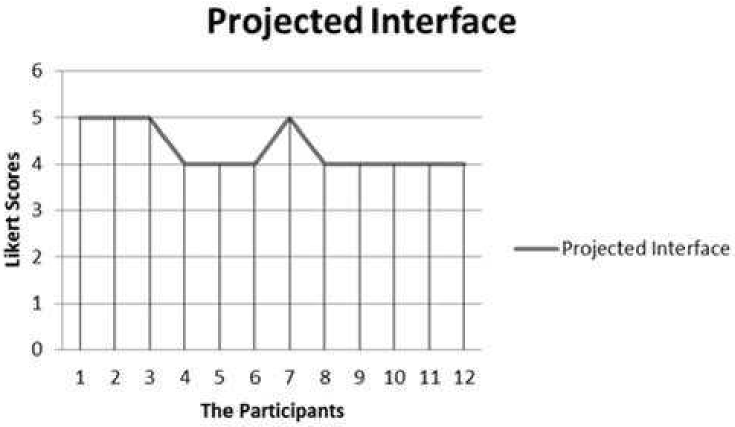


Figure 6.27 The Likert score of the projected interface.

For the configuration, the mean score of satisfaction is 3.833 as shown in figure 6.28. On the positive side, more than half participants expressed the convenience of the view, while more than half participants stated that the heavy weight would lead to discomfort after a long time use of configuration, and the locomotion of the head would impact interaction negatively.

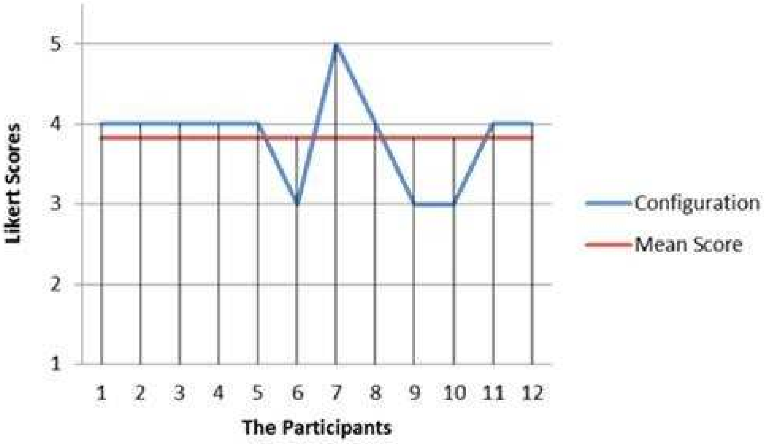


Figure 6.28 The Likert score of the wearable configuration.

6.7 Discussions

After exploring whether the pinch gesture input and hover gesture input are easy to learn and use by users in the stationary and mobile settings, we found that the pinch pointing gesture, the hover pointing gesture, and the pinch drag-drop gesture are all easy to learn and use. The average scores for easiness of learning and utilization with two input techniques in all three settings are all more than 3. Learning has raised the satisfaction of utilization; after learning, all the true task operation scores are higher than the toy learning scores. However, easiness of learning and utilization varies considerably in six cases. For the pinch gesture, the best learning and utilization scores are for sitting situation, followed by standing situation, and the third ranking scores are for walking. The hover gesture has the same tendency as the pinch gesture. Also, in the same setting (sitting, standing, and walking) the pinch gesture scores for learning and utilization are slightly higher than or the same as the hover gesture scores. For the drag-drop action, the average scores of easiness for learning and utilization are all more than 4.

Furthermore, the answer to the second question aforementioned is that both the pinch gesture and the hover gesture will be impacted by the walking setting, but the pinch gesture is affected less than the hover gesture. The Mann Whitey U test results on task completion time with the pinch gesture in three settings showed no difference. However, the same test on task completion time with the hover gesture showed that there are differences between sitting and walking, and between standing and walking. In addition, error rates in three settings with the pinch gesture are lower than with the hover gesture. However, both gestures are still impacted by locomotion. Interaction time with the pinch gesture and hover gesture are ranked in an ascendant way as sitting, standing and walking. Also, the same tendency was observed on error rates. Furthermore, concerning users' comments, incoordination, jittering hand effect, tired forelimbs and extra attention were reported more in the walking situation than in the sitting and standing situations.

This study also showed us the advantages and limitations of the pinch gesture and the hover gesture in three settings. We first discuss the pinch gesture. First, the pinch gesture was less impacted by mobility as we stated above. In particular, the drag-drop action was not noticeably affected by mobility. The average interaction time of the drag-drop action is almost the same. Also, analysis of the Mann Whitey U test on interaction time showed that there are no statistically significant differences between the sitting situation and the standing situation, between sitting and walking, and between standing and walking with drag-drop interaction action. Thus, the drag-drop action is more stable for interacting than the pointing action when the user is walking. Error rates were all low in all three settings. Second, the pinch gesture provides a clear feedback, a low possibility of false pointing, and no load for waiting; when the user releases the pinch, he knows he makes a selection or an action. Third, the pinch gesture caused fewer errors, and more people preferred the pinch gesture than the hover one. Nevertheless, the limitation of the pinch gesture is still presented. When the user makes a wrong pinch gesture, he will fail to pinch, release pinch, drag or drop. Also, the size of the reference-cell for the pinch gesture is larger than that of the hover gesture, which limits the number of selected items on one page. We next discuss the hover gesture. First, the hover gesture does not need any articulated movements; it is simpler. Second, the small-size of the reference-cell allows more items in the interface. Third, the hover gesture provokes the feeling of the touch screen for some participants. However, the limitation is obvious, namely, the user still needs to wait a short time to hover, and this gesture is impacted more by mobility.

In addition to these findings, we also found the ear side position is a good position to display but we need to improve its stability and reduce the weight. Also, manual focus would impact the scalable interface.

6.8 Summary

In this chapter, we presented our approach for interacting with the wearable camera-projector system with the pinch input gesture, the hover input gesture and the fist-palm gesture in both the stationary and mobile settings. Furthermore, we proposed the design of reference-cell for the projected interface according to the pinch gesture in this chapter. We finally evaluated the two gestures in all three settings. Results of our experiments showed that the pinch gesture is less affected than the hover gesture in the mobile setting. Mobility impacted both gestures. Also, the drag-drop action is more stable for interacting than the pointing action when the user is walking. The ear side position is a good position to display but we need to improve its stability and reduce the weight. Manual focus would impact interaction with the scalable interface.

To conclude, as shown in figure 6.29, we have investigated a camera-projector device unit, three input techniques (hover, pinch, and fist-palm gestures), EII, the digital interface, and the ear-side point in this chapter.

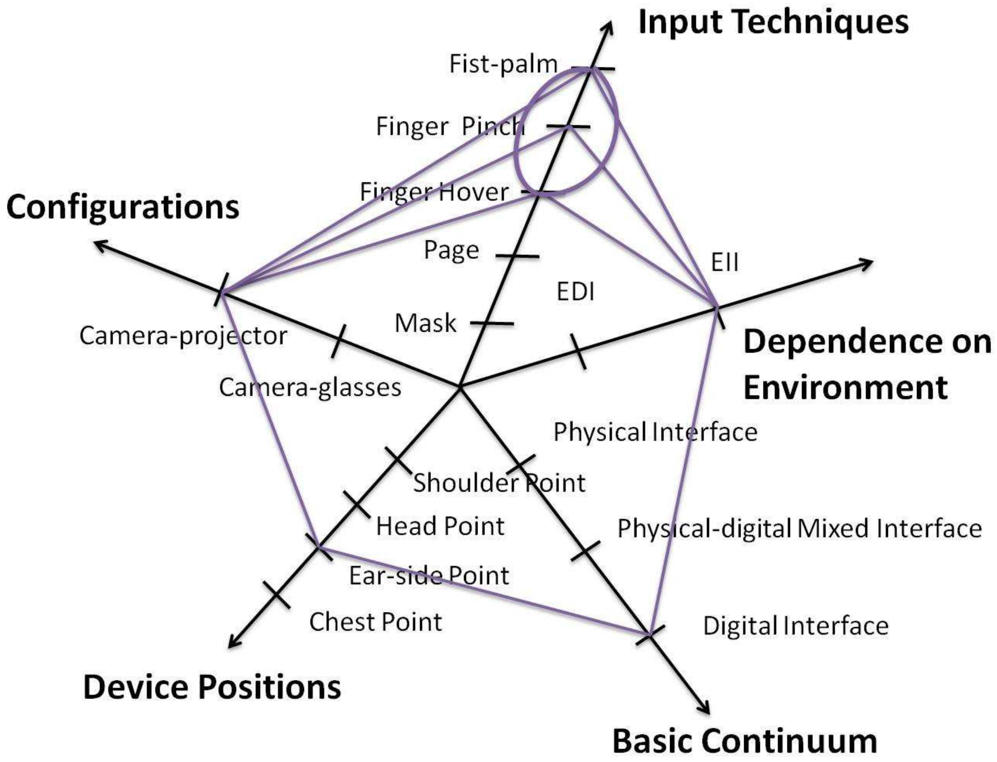


Figure 6.29 The spider figure of hand gestures used in wearable interaction.

From chapter 4 to chapter 6, we have discussed the physical interface (paper-based interface in the MobilePaperAccess system) and the digital interface (projected interface in the PlayAllAround system) according to the basic continuum which will be described in the next chapter. In chapter 7, we will discuss the mixed interface which contains the paper-based part and the projected part. The related interaction techniques will also be described.

7 Physical-digital Mixed Interface

7.1 Introduction

For many years, the combination of paper and computers has been focused and studied. The very nature of paper offers certain advantages over other materials. Paper, as a medium, is comfortable to read and quite light to carry. Its light weight and collapsible nature ensures a high degree of mobility. The texture of paper has positive qualities for both projection and vision-based detection. Marking and tagging through paper-printing is an economical method for designers, users and producers alike. Last but not least, paper documents are widely accepted in public, and have a close relationship with people's lives. Thus, in addition to the aforementioned physical interface (in chapter 4) and the digital interface (in chapter 5 and 6), we now discuss the physical-digital mixed interface. We first describe the basic continuum from physical interface to digital interface for EII and DI, and provide the definition of a mixed interface. We then explain our design and implementation of the mixed interface, which consists of the marker-based paper part and the projected digital part. Most existing paper-based interaction systems such as DigitalDesk (Wellner, 1993), PlayAnywhere (A. D. Wilson, 2005), and Docklamp (Do-Lenh, Kaplan, Sharma, & Dillenbourg, 2009), enable detection by means of markers or barcodes. FACT (Liao et al., 2010) allows a content-based approach for mapping instead of the marker-based mapping method. Since our system needs to support selection on a paper interface, we utilize ARToolKit tags to locate the position of interactive zones for selection. Finally, to demonstrate our interaction techniques, we continue to present our RTIS application according to a mixed interface.

Furthermore, with the aim of providing more information for the mixed interface and to add interactive items, we remove the configuration of the small-size goggle display and adopt the pico-projector as the output device. The projection display can be an alternative method to provide a large image presentation without any external device support. In this way, the interface has the capacity to contain more information and offer more dynamic choices.

7.2 Basic Continuum from Physical Interface to Digital Interface for EDI and EII

In the augmented reality environment, we propose a continuum that spans the range from physical interface to digital interface, based on which the interaction techniques of our EDI and EII design (see Figure 7.1). The physical interface surface is static and inflexible, usually in the form of unitary planar, regarded as the uniplanar. Since the elements in the interface are fixed and physical, these elements should all evolve in the uniplanar, rather than in the multilayer windows. In our study, we use the paper-based interface to create the concept of the physical interface, namely, all the interactive elements are predefined and printed in a piece of paper for interaction. The physical-digital interface incorporates the physical interface with the digital interface, in which the paper interface has been augmented with the projected interface in possession of the half-dynamics. Also, the digital interface is totally projected with personal information and has full dynamics. The latter two interfaces are based on the multiplanar, by means of which the interactive elements are organized logically in the dynamic multilayer windows.

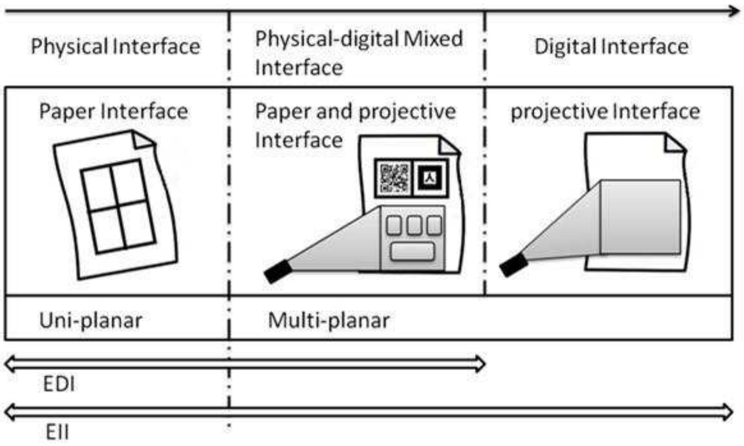


Figure 7.1 The continuum from physical interface to digital interface for EII and EDI.

In the EDI system, the interface relies closely on the environment and the context information, such as location. Namely, in this case, the presentation of the interface is not dependent on the individual's decision. Based on this dependence, the EDI builds on the physical interface and the physical-digital interface. In the EII system, the interface is determined by the actual individual and can either be augmented by the required projected information or augmented with markers or predefined menus. Thus, the EII entirely spans from the physical interface to the digital interface.

Take the "Shopping and Tourist Information" scenario for example: with the uniplanar paper interface, all shopping and tourist information choices are fragmented and reformed logically in a piece of paper as illustrated in figure 7.2 (a). As regards the multiplanar interface, we propose two forms of organization: one utilizes the paper interface and the digital projected interface, while the other leverages merely the digital projected interface. As shown in figure 7.2 (b), the interaction access for the former is the paper with markers, following which digital dynamic information is unfolded according to the signal sent from the paper-based interface. For the latter, the user enters the interaction directly by projecting the interface either on walls or on his/her palm (see Figure 7.2 (c)).

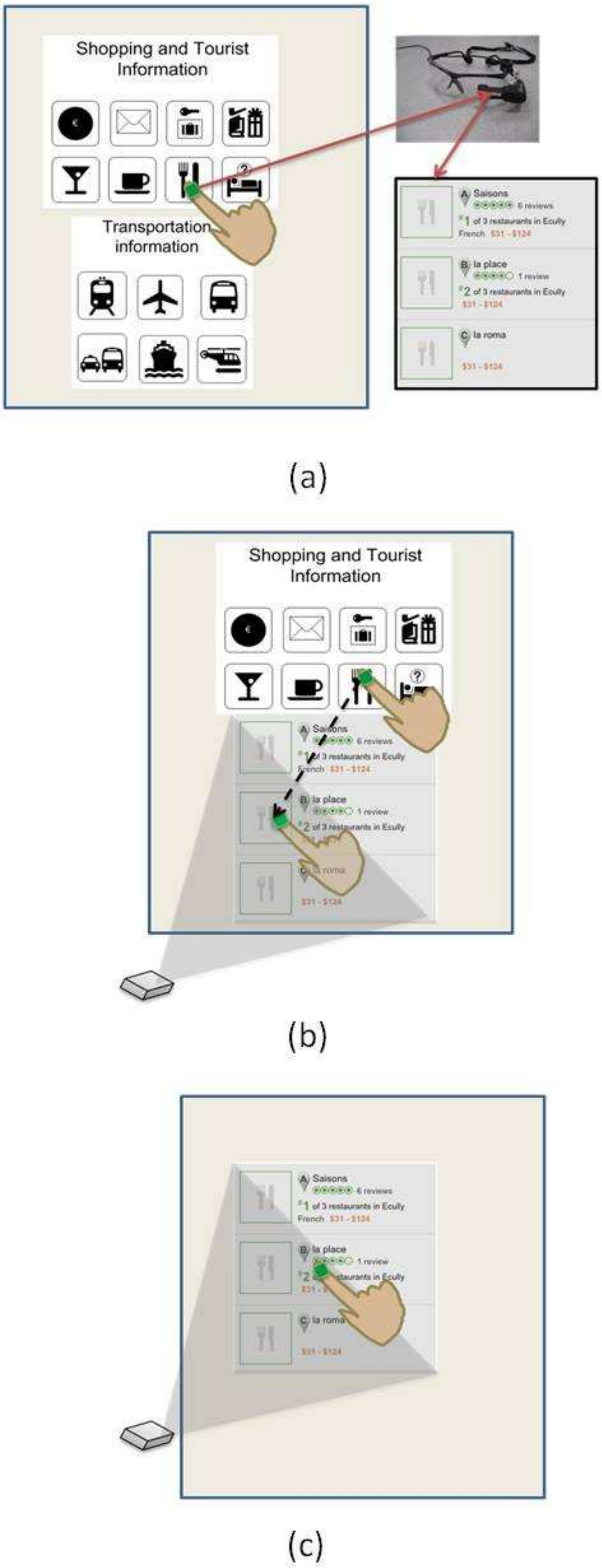


Figure 7.2 The physical-digital mixed interface of the shopping and tourist information scenario.

(a). The physical paper-based interface.

(b). The physical-digital mixed interface.

(c). The digital projected interface.

7.3 Design of Mixed Interface

In this section, we describe the design of the physical-digital mixed interface and its input technique as well as the output technique. As we stated above, our mixed interface offers a physical paper interface and a digital projected interface for interacting, which has following advantages. First, it encourages hand gesture interaction and does not require any other input sensors or devices. This input method reduces the configuration cost as well as the number of wearable devices. Also, hand input is a natural, intuitive and non-contact method compared with mask selection described in chapter 4. Next, the mixed interface can overlap with or augment digital information directly on the physical paper surface. Finally, the projected interface puts an end to limitation of the small-size presentation and offers a large-size display experience, thus increasing the amount of interaction information.

7.3.1 Place-around Wearable Camera-projector Unit

To support camera reorganization and the interface projection, we select four fixed points to stabilize the camera-projector unit as illustrated in figure 7.3: the chest point (a), the ear-side point (b), the over-head point (c), and the shoulder point (d). On the one hand, these points make it easier to place the unit compared with other positions on the body with the help of foaming or plastic materials. On the other, the positions of these points are closer to the eyes, thus reducing the possibility of distortion of the projection image as well as providing the appropriate field of view. In chapter 3, we equipped the user with the over-head unit (see Figure 7.3 (c)). The chest-point was discussed and evaluated in chapter 5, while the ear-side position was discussed and evaluated in chapter 7. In this chapter, we employ the chest-point as the wearable configuration position (see Figure 7.3 (a)).

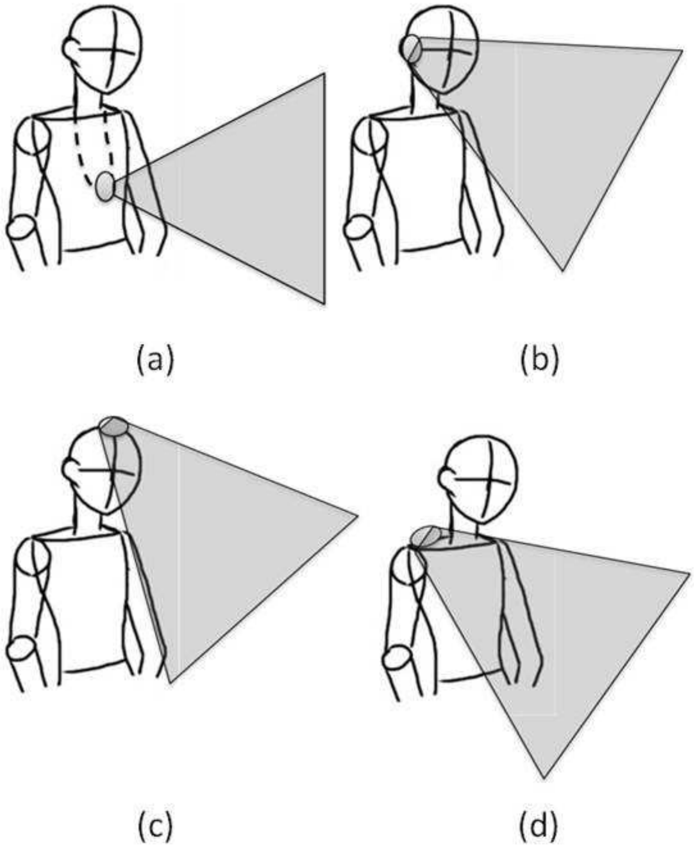


Figure 7.3 The stabilized points of the camera-projector unit on the body.
(a). Chest point.
(b). Ear-side point.
(c). Over-head point.
(d). Shoulder point.

7.3.2 Interaction Process of Mixed Interface

The mixed interface consists of two parts: the paper-based interface and the projected interface. As shown in figure 7.4, the paper-based area is made up of the physical interactive items and provides the physical tactile experience. The projected interface is formed by the digital interactive items. Thus, the mixed interface has two interaction loops: the 1st loop is the physical interaction loop and the 2nd loop is the digital interaction loop.

The 1st loop provides the user with the concise sound feedback for selection, whether or not user enters the digital interaction process. Also, if the user selects items by touching the paper with the finger input technique, he/she also obtains an instantaneous tactile feedback from the paper. The interaction of this loop is dependent on the environment indications, predefined menus and markers. We propose the ARToolKit tags as a part of the physical interface. These ARToolKit tags play a two-fold role; first, they identify the unique interface which is distinguished from the other interfaces, and second they mark the start point of the interactive zone and arrange the interface structure.

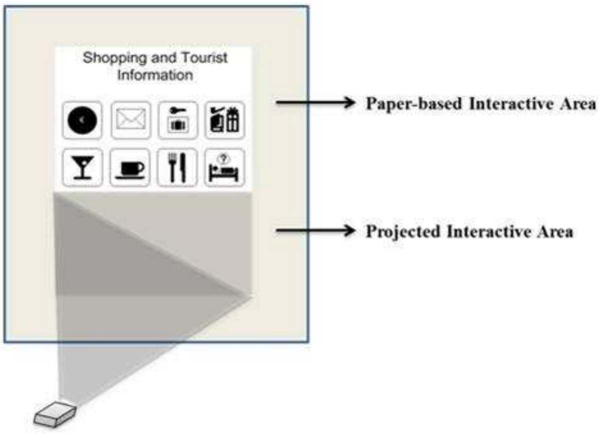


Figure 7.4 The mixed interface design.

The 2nd loop provides the user both with sound feedback and visual projected feedback. The quick sound response can reduce both the inconvenience of visual feedback delay and the anxiety caused by the waiting time due to hovering. Because the projected interface impropriates the arbitrary daily surface, the user can also obtain the tactile feedback from these daily surfaces. The interaction of this loop is constrained by the in-environment information from the physical interface; it is the extension or the evolvement of paper-based interaction.

During interaction, the physical paper-based interface and the digital projective interface are mutually exclusive. In other words, when the user enters the physical area and triggers the physical interactive items, the projected interactive items are locked until the user asks for entering into the projective area. Once the user starts to manipulate the projected interactive items, the physical interaction items are locked until the user quits navigation of the projected interface. The original start and end quit usually takes place from the physical paper-based interface.

7.4 Implementation

An overview of the mixed interface setup will be illustrated in this section (see Figure 7.5). It includes the paper augmented interface with the ARToolKit tags, the user with his/her finger augmented with the green marker, a camera-projector unit, as well as a tablet for calculating only. We discuss each item and their collaboration below.

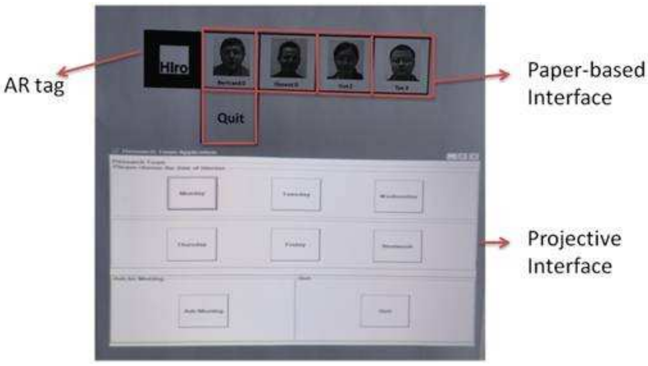


Figure 7.5 The mixed interface.

7.4.1 Wearable Configuration

Our wearable configuration contains a webcam, a pico-projector and a tablet for calculating. The Logitech webcam can obtain a 640×480 RGB frame. The pico-projector is Acer C120 with a resolution of 1280×800 maximum. Projector size is roughly 2.54cm×11.94cm×8.13cm with a weight 180g. In standard mode brightness can reach 100 lumens, while in low mode brightness maintains 75 lumens, able to support the indoor interaction. The tablet has a multi-touch screen, which can be carried on the back or in a messenger bag across the body. We fix the webcam and pico-projector together on the light cardboard support (see Figure7.6) and choose the chest as the worn point. The ribbon hanging on the neck is adjustable according to the stature of the user. The light camera-projector pendant is stabilized vertically along the chest, which is easy to wear on and remove from the neck.

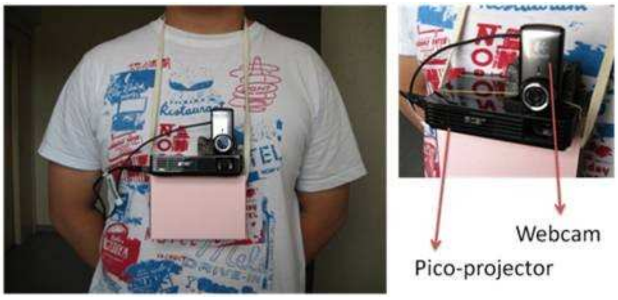


Figure 7.6 The wearable configuration with camera-projector unit.

7.4.2 Augmented Interface

The mixed interface is augmented with the ARToolKit tags, the number of which relies on the arrangement of the interface. In the current stage, we utilize only one tag for identifying the unique paper-based interface and locating the position of the interactive items. As shown in figure 7.7 (a), the ARToolKit tag on the left upside of the interface can be captured by the webcam and tracked by the ARToolKit software. The sound and visual feedback interface is written in C++ with the gtkmm library.

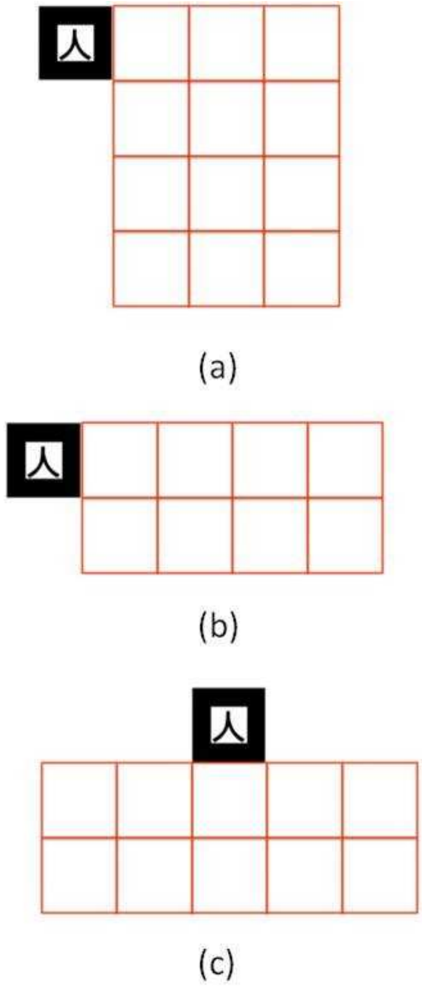


Figure 7.7 The arrangement of the physical-digital mixed interface.

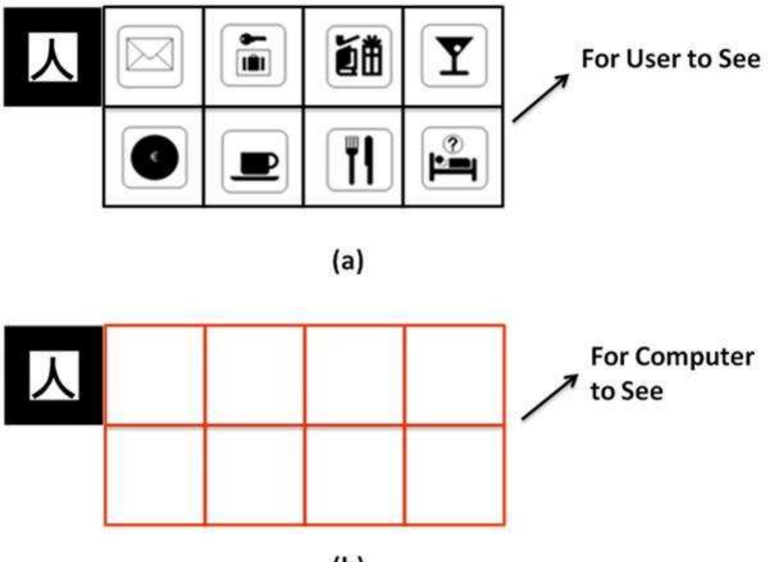
(a). The tag on the left upside position.

(b). The tag on the left position.

(c). The tag on the upside position.

We give some examples of tag arrangement. The predefined menu on the paper-based part is as follows. The main principle is that the tag should not be occluded by the finger or hands during interaction. Once the tag is hidden from camera view, its function will fail. We thus place the tags on left upside position (see Figure 7.7 (a)), left (see Figure 7.7 (b)), or upside (see Figure 7.7 (c)) for the right-handed user, while we place the tags in the right upside, right, or upside position for the left-handed user.

The projected interface is located below the physical interface. When the user switches to the projected interface, the information related to the physical interface is summoned. Figure 7.8 shows the interface seen from the view of the user and the view of the computer. The icon patterns within each interactive area are predefined according to the user's cognition, to cultural, social, and other aspects of psychology (see figure 7.8 (a)), meaningful for the user but meaningless for the computer. The computer actually sees the interface as shown in figure 7.8 (b); it sees a tag and its related zones. The red zone is defined as each interactive item area, with the same tag size. The size, number and relative location of red zones, as well as tag arrangement can be designed and defined to comply with the actual situations.



(b)

Figure 7.8 The paper-based interface part.

(a). The interface from the user's view.

(b). The interface from the computer's view.

7.4.3 Motion of Finger on Physical and Digital Interface

In our current work, we use the object tracking algorithm based on the Camshift algorithm (G. R. Bradski, 1998) by employing the OpenCV (G. Bradski, 2000) library. First, the captured frame is preprocessed. Then, we take a picture of the tracking object located on the user's finger in advance and extract the color feature from this image. Third, the back projection of the processed image is calculated and the Camshift tracks the distribution of target color feature based on the back projection. We can thus automatically track the color marker located on the index finger. Since the mixed interface has a physical part and a digital part, validation of selection of these two parts is different. For the former, validation of a pointing selection acts on the predefined zone as the same method stated in chapter 4, while for the latter, validation of the selection acts on the digital widgets like the buttons stated in chapters 5 and 6. Since the projective area is below the physical area, the user can switch vertically between these two areas.

7.5 Scenario of Research Team Interaction System on Mixed Interface

The mixed interface is applicable to both the EDI and EII, and can also realize the concepts of the EDI and EII with a variety of applications. The mixed interface for the EDI is dependent on the environment context, while the mixed interface for the EII is used for self-contextualization.

In the smart city, the IEI, EDI and EII can solve the same problems that the user encounters, as well as solve different problems respectively. To concretize these three interfaces, we implemented an application known as the Research Team Interaction System (RTIS), which varies slightly according to the interface. In this chapter, we continue to employ the RTIS according to the mixed interface for both EDI and EII. The RTIS scenario on a mixed interface for EDI is as follows. A team member wants to consult another member. When he/she arrives at the lab, he/she finds this person is out. He/she walks close to the door and starts to use the RTIS. He/she first selects and validates this member on the physical paper-based interface, which is pasted on door in advance. He/she then interacts with the summoned projective interface on the door, looks for an appropriate time and selects to check the schedule, all in the form of projected interactive items. After marking the decision, he/she asks for an appointment and obtains feedback from the system. When the user leaves the projection interactive area, he/she begins to connect with the physical area automatically, and can check other members' information from the start point. For the same scenario with EII, the user takes out a paper or a booklet that he/she holds in his hand to interact. This paper or booklet is printing with the appropriate tag and predefined menus in advance, and forms a reserved space for projection. In this way, the user can check the schedule and make the appointment via the paper on his/her hand as required.

7.6 Summary

In this chapter, we introduced our taxonomy of physical and digital interfaces and described the design and implementation of the mixed interface, as well as evaluation planning. The layouts of the mixed interface and two loops are also discussed in this chapter. To conclude, as shown in figure 7.9, in this chapter, we have investigated camera-projector device unit, two input techniques (hover and pinch gestures), the EDI and EII, the physical-digital mixed interface, and the chest point.

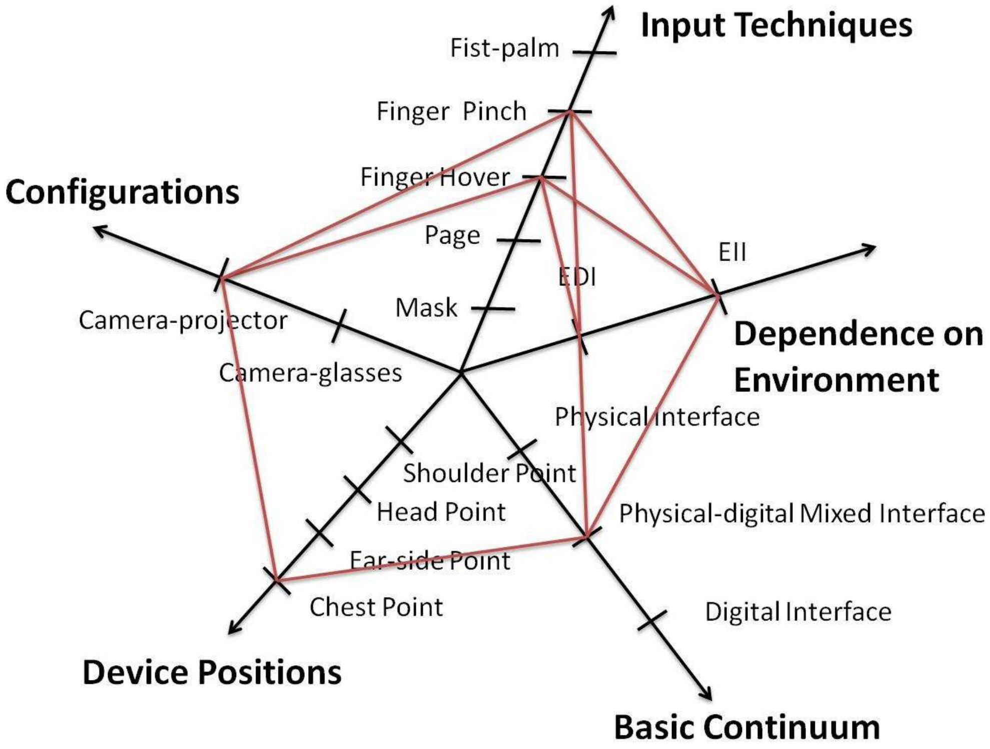


Figure 7.9 The spider figure of mixed interface for mobile interactions.

In the next chapter, we will draw the conclusions of our work. We will discuss the contributions and limitations as well as provide a prospective direction for future work.

8 Conclusions

8.1 Introduction

To enable interaction freely in an environment, one solution is to let the user wear the configurations that are made up of a set of interaction devices, and allow interaction with at least one hand free. Taking into account the location (physical, geographical or logical) and the target activities of the user, the interaction style and devices must be appropriate to the context. In this thesis, we first globally presented our design approach and our taxonomy of mobile user interfaces. Our research aims at investigating the innovative Environment Dependent Interface (EDI) and the Environment Independent Interface (EII). We described the design of these interfaces, their wearable configurations, real examples of use and a preliminary evaluation of input selection techniques to prove the feasibility of our prototypes. We then proposed a continuum from physical interface to digital interface in relation with EDI and EII. Based on the physical interface, MobilePaperAccess is a wearable camera-glasses system with a tangible user interface allowing access to digital information from a paper interface. The system is devised to validate our concepts of EDI and EII, which focus on enabling people to access their personal data as well as public resources at any time and in any place. The two interfaces (EDI and EII) and three input techniques (finger input, mask input and page input) have been evaluated in both the quantitative and qualitative methods. Third, along with the continuum, we continued to investigate the physical-digital mixed interface by a wearable configuration including webcam, pico-projector and tablet, with the goal of concretizing the concepts of EDI and EII. The mixed interface contains the marker-based part with ARToolKit tags and the projected digital part. Since the projection display can be an alternative method for providing the large image presentation without any external support of the device, the mixed interface is able to contain more information as well as offer more dynamic choices compared with the physical interface. To go one step further, the projected scalable interface has been studied on the basis of the PlayAllAround system, which is a wearable camera-projector system with a scalable interface allowing mobile interaction and providing both the nearer small-size interface and the farther large-size interface supporting both the private and public use. In addition, we proposed the design of a reference-cell as well as the principle from decomposition of application tasks to formation of a scalable interface. The hover gesture and the scalable interface have been evaluated globally. The results of our experiments have shown that our system performs well with the hover gesture and the scalable interface. Finally, besides investigation into the physical interface, the physical-digital mixed interface as well as the digital interface, we explored hand gestures including the hover gesture, the pinch gesture and the fist-palm gesture. We employed the pinch gesture and fist-palm gesture as the input for pointing, drag-drop action and painting. To satisfy the requirements of interaction in our sophisticated everyday life, we investigated how the user might interact with this system in both the stationary and mobile settings. We compared the interactions of the hover gesture and pinch gesture in three situations, namely standing, sitting and walking. Furthermore, the projected interface was evaluated, and satisfaction as to operation of the head worn configuration was discussed.

In this chapter, we describe briefly the contribution of the research conducted in my thesis, discuss the limitations that exist in the current work, and propose directions for future investigation.

8.2 Contributions

The first and most significant contribution of my work is the design of the Dependent Environment Interface (EDI) and the Independent Interface (EII), as well as the implementation of the prototypes using the EDI and EII as the fundamental stones. We have evaluated the feasibility and usability of the wearable system on the basis of the EDI and EII. Our wearable computing system lets the user move closer to the ubiquitous environment and interaction. The innovative input and output modality removes the restriction of the traditional desktop interaction and offers the advantages of wearable computing.

In the realization process, we defined the taxonomy of the mobile interfaces including the physical interface, the physical-digital mixed interface as well as the digital interface, which formally make up a continuum. Along this continuum, we studied paper-based interaction, the mixed interaction, and scalable interaction, as well as the related input techniques, the special characteristics and affordances of pico-projectors and the goggle with small screen. We designed, developed, and evaluated a rich set of interaction techniques and innovative interfaces for wearable computing. We also explored the usage scenarios enabled by our designs, involving various activities that satisfy the environment dependent and independent situations. First, the interaction modality of the EDI, achieved by the MobilePaperAccess system, supports privacy and avoids the problem of fat fingers and blocking of the focus point by shadow and the actual finger. The quantitative and qualitative results of the MobilePaperAccess system have shown that the main interaction error of this wearable paper-based interface is the locomotion error caused by jittering hands or body movement. Also, mask input with the EDI has the best performance due to its shortest interaction time, the shortest access time, no locomotion errors, and the best satisfaction. Second, the design of the mixed interface offers the possibility of enlarging interface space and capacity. This augmentation guides the future interaction designs for the pico-projector, which extends the modality combined with the tangible interface rather than the solo projection range. Third, by taking into account the special characteristic of the pico-projector scalability, we designed and developed the PlayAllAround system, based on our concepts of EII. We proposed the design of the reference-cell and the principle from decomposition of the application tasks to composition of the scalable interface, which is a pioneer work for the research of projection in mobility. This work reveals the potency of the pico-projector compared with the traditional fixed-size display. The results of our experiments have shown that the PlayAllAround system performs well with the scalable interface. Finally, we investigated the performance of the wearable camera-projector system with the EII in both the stationary and mobile settings. This work brings our research closer to the true ubiquitous computing environment. The results of our experiments have shown that the pinch gesture undergoes less influence than the hover gesture in the mobile settings. Mobility impacted on both these gestures. Also, the drag-drop action is more stable for interaction than the pointing action when the user is walking. The ear side position is a good position to display but we need to improve its stability and reduce its weight. The manual focus would influence interaction with the scalable interface.

For the wearable configurations, we have discussed the stabilizing positions of the camera-projector unit (the over-head position, the chest position, the ear-side position, etc.) by observation and qualitative evaluation, which would enlarge the research design space for wearable computing and devices.

To conclude, our current work has investigated many aspects of wearable interaction and the innovative user interface on the basis of the camera-glasses and camera-projector systems, and paves the way for future research in this field.

8.3 Limitations

This work still has certain limitations. These concern the restriction of existing devices, insufficient robustness in the true mobile environment as well as lack of study of the social interaction, which leave room for improvement and exploration.

First, existing devices restrain ubiquitous usage of our prototypes and systems. On the one hand, our display is based on the goggle attached small screen and the pico-projector. Even though the goggle attached small screen has the advantages of providing private information presentation, offering in-time glance feedback, and making more space for input, it is not totally perfect. The small screen cannot guarantee totally seamless overlaid information when augmenting reality. Instead, the transparent small display is more suitable for achieving the augmentation goal. Moreover, as regards wearing, it is not convenient to stabilize the small screen on the user‘s own corrected glasses. With respect to the pico-projector, current pico-projector products have low lumens; brightness varies from 15 lumens to 200 lumens, largely insufficient for the requirement to support interaction under normal illumination. Thus, our work with the pico-projector is performed in a darker environment. This limitation is likely to be alleviated in a few years, with the emergence of more mature technology, thus allowing camera-projector interaction to be performed in a true ubiquitous environment. On the other hand, the tablet on the back is still heavy to carry compared with pocket devices. However, normal existing handheld devices are unable to support real-time computer vision calculation.

Second, even if we endeavor to support interaction in a realistic mobile environment, the restriction still exists. The real background in our everyday life is multicolored, which will lead to mistaken recognition of the colored markers located on the fingers. Similarly, in a sophisticated background or a darker environment, efficiency of bare hand recognition will decease more or less in varying degrees.

Finally, when exploring the prototypes and usage scenarios, we focused on demonstrating the interaction concepts, the interaction techniques and related evaluations, but did not address issues on cooperation among multiple users as well as social applications.

8.4 Future Directions

We have explored a wide range of research issues related to wearable interaction. To come up with the solutions on the aforementioned issues, we discuss some aspects of the research directions for a more profound study in the future. In this section, we will discuss the future short-term directions and long-term directions.

8.4.1 Short-term Perspectives

We first discuss our short-term perspectives concerning evaluation of the mixed interface, study of the scalable interface, and improvement of wearable configurations.

8.4.11 Evaluation of Mixed Interface

In chapter 7, we discussed the design, implementation and scenario of physical-digital mixed interface. In this chapter, we provide the evaluation planning and preparation to test usability of the mixed interface.

We organize an evaluation and explore four main questions as follows to investigate the mixed interface.

- Question 1: Does the user find it easy to navigate with a floating projected interface on a paper-based interface in both stationery and mobile settings?
- Question 2: Is there any difference on interaction between these two settings?
- Question 3: In which situations do the users prefer to use EDI and EII?
- Question 4: In which location of the paper-based part would the user like to set the floating scalable projected part?

We plan to recruit 12 participants and discuss the main results of our evaluation. These results will help us propose the solutions for improvement and the perspectives in our future work.

8.4.12 Scalable Interface

The everyday surfaces in the true environment are sophisticated and have varying sizes, colors and textures, thus making the augmentation difficult on the projected interface. In our future work, besides planar surfaces, we also plan to take into account non-planar projective surfaces and the everyday colored surfaces, such as the cup surface, which is curved in the horizontal direction and has colors that are not just white. We also plan to consider using the palm or the forearm as the small and proximal projective surface. Thus, to detect these non-planar and more sophisticated surfaces, and distinguish the dominant interacting hand and the other parts of the body surface, we consider the depth-camera to be the input device.

In addition, projected augmented information requires perception of object surface and form in the context. In this thesis, we studied two scales of thresholds: the nearer interface and the farther interface. In the future, more thresholds scales will be investigated to achieve the goal of the arbitrary projected interface.

5.4.13 Improvement of Wearable Configurations

As regards configuration, we consider using a depth-sensing camera such as Microsoft Kinect (see Figure 8.1) or another custom depth-sensing camera instead of the webcam that we used in this thesis. With the depth information captured by depth-camera, we can detect the size, color and texture of the surface as stated in 8.4.12. Also, depth-information provides more space for recognition of dynamic gestures.

Furthermore, we plan to study quantitatively the influence of mobility on interaction when the wearable device units are fixed on different parts of the body. A leg speed recorder will be used to mark the state of the body movement.

8.4.2 Long-term Perspectives

We then discuss the long-term perspectives, including enlarged design space for EDI, enlarged design space for EII, and multi-user interaction.

8.4.21 Enlarged Design Space for EDI

Several improvement and extensions can be made to the current prototypes and systems based on the EDI. One possibility is to enlarge the design space of tangible input. Instead of relying on finger input, mask input and page input, the fridge magnet object with the specific markers could be used to make the multiple choices. The interface would be divided into several zones and the camera could recognize each zone. Each magnet object with a specific marker could be detected uniquely. Consequently, placement of the magnet object in the related zone would lead to selection even of multiple choices.

8.4.22 Enlarged Design Space for EII

Concerning the enlarged design space of EII, we will focus on improvement of input techniques. Besides the hover gesture, the pinch gesture and the fist-palm gesture, we will introduce more hand gestures for navigating our adaptive wearable interface. On the one hand, the static hand posture can be used to navigate specific events such as the calling gesture that can summon the telephone interface. Or the "OK" gesture can summon the validation function. In addition, the static hand posture can be used to navigate the physical interface. In this way, interactive items can be selected by hand postures other than the finger hover gesture or the pinch gesture. On the other hand, the dynamic hand gesture could allow more sophisticated interacted tasks. For example, the user can zoom in or zoom out the digital projected interface via hand gestures. Since the depth camera provides depth information of the objects, more accuracy will be provided for hand gesture recognition and detection. Furthermore, with depth information, the two hands can be distinguished in an efficient manner, in which one hand or another part of the body can be used as the projected surface.

8.4.23 Multi-user and Social Interaction

Our evaluation of the scalable interface reveals that users intend to share their interface with others. They are willing to share their large-size interaction experience with other people. We thus plan to investigate multi-person interaction and their collaboration on the tasks. Further interaction could be designed to support interaction of at least two people. We will also consider synchronous and asynchronous interactions in the future. In addition, both the EDI and EII have the capacity to support the multi-user. With the EDI, mobile users may work collaboratively on the large physical predefined interface. With the EII, the principal user could share his/her interface to the other user or let other interacts as the secondary role in a social manner. The roles of the actors and the relationship between them will be considered and defined.

Bibliography

Akaoka, E., Ginn, T., & Vertegaal, R. (2010). DisplayObjects: prototyping functional physical interfaces on 3d styrofoam, paper or cardboard models. In *TEI 2010: Proceedings of the 4th international Conference on Tangible and Embedded Interaction* (p . 49–56).

Albinsson, P. A., & Zhai, S. (2003). High precision touch screen interaction. In *CHI 2003: Proceedings of the SIGCHI conference on Human factors in computing systems* (p . 105–112).

Aleksy, M., Rissanen, M. J., Maczey, S., & Dix, M. (2011). Wearable Computing in Industrial Service Applications. *ANT 2011: Proceedings of the 2nd international conference on Ambient System, Networks, and Technologies*, *Volume 5*, p. 394–400.

Amft, O., & Lukowicz, P. (2009). From backpacks to smartphones: Past, present, and future of wearable computers. *Pervasive Computing, IEEE*, *Volume 8*(Issue 3), p. 8–13.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, *Volume 21*(Issue 6), p. 34–47.

Baillot, Y., Brown, D., & Julier, S. (2001). Authoring of physical models using mobile computers. In *ISWC 2001: Proceedings of the 5th international symposium on Wearable Computers* (p . 39–46).

Ballendat, T., Marquardt, N., & Greenberg, S. (2010). Proxemic interaction: designing for a proximity and orientation-aware environment. In *ITS 2010: ACM International Conference on Interactive Tabletops and Surfaces* (p . 121–130).

Baudhuin, E. S. (1996). Telemaintenance applications for the WearableTMPC. In *15th AIAA/IEEE: Proceedings of the Digital Avionics Systems Conference* (p . 407–413).

Beardsley, P., Forlines, C., Raskar, R., & VanBaar, J. (2005). Handheld projectors for mixing physical and digital textures. In *CVPR Workshops: Workshops of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (p . 112–112).

Beardsley, P., Van Baar, J., Raskar, R., & Forlines, C. (2005). Interaction using a handheld projector. *Computer Graphics and Applications, IEEE*, *Volume 25*(Issue 1), p. 39–43.

Benko, H., & Wigdor, D. (2010). Imprecision, Inaccuracy, and Frustration: The Tale of Touch Input. *Tabletops-Horizontal Interactive Displays*, p. 249–275. doi:10.1007/978-1-84996-113-4_11

Blask, G., Coriand, F., & Feiner, S. (2005). Exploring interaction with a simulated wrist-worn projection display. In *ISWC 2005: Proceedings of the 9th international symposium on Wearable Computers* (p . 2–9).

Bradski, G. (2000). The opencv library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, *Volume 25*(Issue 11), p. 120–126.

Bradski, G. R. (1998). Computer vision face tracking for use in a perceptual user interface. In *WACV 1998: Proceedings IEEE Workshop on Application of Computer Vision* (p . 214–219).

Buxton, W. (1990). A three-state model of graphical input. In *INTERACT 1990: Proceedings of 3rd IFIP international conference on Human-Computer Interaction* (p . 449–456).

Cakmakci, O., & Rolland, J. (2006). Head-worn displays: a review. *Journal of Display Technology*, *Volume 2*(Issue 3), p. 199–216.

Cao, X., Forlines, C., & Balakrishnan, R. (2007). Multi-user interaction using handheld projectors. In *UIST 2007: Proceedings of the 20th annual ACM symposium on User Interface Software and Technology* (p . 43–52).

Cao, X., Massimi, M., & Balakrishnan, R. (2008). Flashlight jigsaw: an exploratory study of an ad-hoc multi-player game on public displays. In *CSCW 2008: Proceedings of the 2008 ACM conference on Computer Supported Cooperative Work* (p . 77–86).

Cheverst, K., Davies, N., Mitchell, K., Friday, A., & Efstratiou, C. (2000). Developing a context-aware electronic tourist guide: some issues and experiences. In *CHI 2000: Proceedings of the SIGCHI conference on Human factors in computing systems* (p . 17–24).

David, B. T., & Chalon, R. (2007). IMERA: Experimentation Platform for Computer Augmented Environment for Mobile Actors. In *WiMOB 2007: 3rd IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, 2007* (p . 51).

David, B. T., Zhou, Y., Xu, T., & Chalon, R. (2011). Mobile user interfaces and their utilization in a Smart City. In *ICOMP'11: The 2011 International Conference on Internet Computing as part of WorldComp'2011 Conference, CSREA Press*.

Do-Lenh, S., Kaplan, F., Sharma, A., & Dillenbourg, P. (2009). Multi-finger interactions with papers on augmented tabletops. In *TEI 2009: Proceedings of the 3rd international Conference on Tangible and Embedded Interaction* (p . 267–274).

Dul, J., & Weerdmeester, B. A. (2008). *Ergonomics for beginners: a quick reference guide* (p. 12.). CRC Press.

Elias, J. G., Westerman, W. C., & Haggerty, M. M. (2010). Multi-touch gesture dictionary. Google Patents.

Erol, A., Bebis, G., Nicolescu, M., Boyle, R. D., & Twombly, X. (2007). Vision-based hand pose estimation: A review. *Computer Vision and Image Understanding*, *Volume 108*(Issue 1-2), p. 52–73.

Esenther, A., & Ryall, K. (2006). Fluid DTMouse: better mouse support for touch-based interactions. In *AVI 2006: Proceedings of the working conference on Advanced Visual Interfaces* (p . 112–115).

Feiner, S., Macintyre, B., & Seligmann, D. (1993). Knowledge-based augmented reality. *Communications of the ACM*, *Volume 36*(Issue 7), p. 53–62.

Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, *Volume 8*(Issue 5), p. 347–358.

Fuchs, H., Livingston, M., Raskar, R., Colucci, D., Keller, K., State, A., … Meyer, A. (1998). Augmented reality visualization for laparoscopic surgery. *Medical Image Computing and Computer-Assisted Interventation—MICCAI'98*, *Volume 1496/1988*, p. 934–943. doi:10.1007/BFb0056282

Furmanski, C., Azuma, R., & Daily, M. (2002). Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visible and obscured information. In *ISMAR 2002： Proceedings of International Symposium on Mixed and Augmented Reality* (p . 215–320).

Grant, D. A. (1948). The latin square principle in the design and analysis of psychological experiments. *Psychological bulletin*, *Volume 45*(Issue 5), p. 427.

Gratzel, C., Fong, T., Grange, S., & Baur, C. (2004). A non-contact mouse for surgeon-computer interaction. *Technology and Health Care-European Society for Engineering and Medicine*, *Volume 12*(Issue 3), p. 245–258.

Greaves, A., Akerman, P., Rukzio, E., Cheverst, K., & Hakkila, J. (2009). Exploring user reaction to personal projection when used in shared public places: A formative study. In *Mobile HCI 2009 Workshop: CAM3SN*.

Greaves, A., & Rukzio, E. (2008). Evaluation of picture browsing using a projector phone. In *MobileHCI 2008: Proceedings of the 10th international conference on Human Computer Interaction with Mobile Devices and Services* (p . 351–354).

Greaves, A., & Rukzio, E. (2009). View & share: supporting co-present viewing and sharing of media using personal projection. In *MobileHCI 2009: Proceedings of the 11th international conference on Human Computer Interaction with Mobile Devices and Services* (p . 44).

Grimes, G. J. (1983). Digital data entry glove interface device. Patent US4414537.

Gustafson, S., Bierwirth, D., & Baudisch, P. (2010). Imaginary interfaces: spatial interaction with empty hands and without visual feedback. In *UIST 2010: Proceedings of the 23th annual ACM symposium on User Interface Software and Technology* (p . 3–12).

Gustafson, S., Holz, C., & Baudisch, P. (2011). Imaginary phone: learning imaginary interfaces by transferring spatial memory from a familiar device. In *UIST 2011: Proceedings of the 24th annual ACM symposium on User Interface Software and Technology* (p . 283–292).

Ha, T., & Woo, W. (2006). Bare hand interface for interaction in the video see-through HMD based wearable AR environment. *Entertainment Computing-ICEC 2006*, *Volume 4161/2006*, p. 354–357. doi:DOI: 10.1007/11872320_48

Ha, Y., & Rolland, J. (2002). Optical assessment of head-mounted displays in visual space. *Applied optics*, *Volume 41*(Issue 25), p. 5282–5289.

Hang, A., Rukzio, E., & Greaves, A. (2008). Projector phone: a study of using mobile phones with integrated projector for interaction with maps. In *MobileHCI 2008: Proceedings of the 10th international conference on Human Computer Interaction with Mobile Devices and Services* (p . 207–216).

Harrison, C., Benko, H., & Wilson, A. D. (2011). OmniTouch: wearable multitouch interaction everywhere. In *UIST 2011: Proceedings of the 24th annual ACM symposium on User Interface Software and Technology* (p . 441–450).

Harrison, C., Ramamurthy, S., & Hudson, S. E. (2012). On-body interaction: armed and dangerous. In *TEI 2012: Proceedings of the 6th international Conference on Tangible and Embedded Interaction* (p . 69–76).

Harrison, C., Tan, D., & Morris, D. (2010). Skinput: appropriating the body as an input surface. In *CHI 2010: Proceedings of the 28th SIGCHI conference on Human factors in computing systems* (p . 453–462).

Höllerer, T., & Feiner, S. (2004). Mobile augmented reality. *Telegeoinformatics: Location-Based Computing and Services. CRC Press*.

Höllerer, T., Feiner, S., & Pavlik, J. (1999). Situated documentaries: Embedding multimedia presentations in the real world. In *ISWC 1999: Proceedings of the 3rd international symposium on Wearable Computers* (p . 79–86).

Holman, D., Vertegaal, R., Altosaar, M., Troje, N., & Johns, D. (2005). Paper windows: interaction techniques for digital paper. In *CHI 2005: Proceedings of the SIGCHI conference on Human factors in computing systems* (p . 591–599).

Hornecker, E., & Psik, T. (2005). Using ARToolKit markers to build tangible prototypes and simulate other technologies. *Human-Computer Interaction-INTERACT 2005, Volume 3585/2005*, p. 30–42. doi:10.1007/11555261_6

Hosoi, K., Dao, V. N., Mori, A., & Sugimoto, M. (2007). VisiCon: a robot control interface for visualizing manipulation using a handheld projector. In *ACE 2007: Proceedings of the international conference on Advances in Computer Entertainment Technology* (p . 99–106).

Höysniemi, J., Hämäläinen, P., Turkki, L., & Rouvi, T. (2005). Children's intuitive gestures in vision-based action games. *Communications of the ACM*, *Volume 48*(Issue 1), p. 44–50.

Hua, H., Gao, C., Biocca, F., & Rolland, J. P. (2001). An ultra-light and compact design and implementation of head-mounted projective displays. In *Proceedings IEEE Virtual Reality* (p . 175–182).

Huber, J., Liao, C., Steimle, J., & Liu, Q. (2011). Toward Bimanual Interactions with Mobile Projectors on Arbitrary Surfaces. In *Proceedings of Workshop on Mobile and Personal Projection* (p . 10–12).

Ishii, H. (2008). Tangible bits: beyond pixels. In *TEI 2008: Proceedings of the 2nd international conference on Tangible and Embedded Interaction* (p . xv–xxv).

Ishii, H., & Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. In *CHI 1997: Proceedings of the SIGCHI conference on Human factors in computing systems* (p . 234–241).

Kato, I. P. H., Billinghurst, M., & Poupyrev, I. (2000). Artoolkit user manual, version 2.33. Human Interface Technology Lab, University of Washington.

Kessler, G. D., Walker, N., & Hodges, L. F. (1995). Evaluation of the CyberGlove (TM) as a whole hand input device. Georgia Institute of Technology.

Klinker, G., Creighton, O., Dutoit, A. H., Kobylinski, R., Vilsmeier, C., & Brugge, B. (2001). Augmented maintenance of powerplants: A prototyping case study of a mobile AR system. In *Proceedings IEEE and ACM International Symposium on Augmented Reality* (p . 124–133).

Ko, J. C., Chan, L. W., & Hung, Y. P. (2010). Public issues on projected user interface. In *CHI EA 2010: Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems* (p . 2873–2882).

Kolsch, M., Turk, M., & Hollerer, T. (2004). Vision-based interfaces for mobility. In *MobiQuitous 2004: Proceedings of the first annual international conference on Mobile and Ubiquitous Systems: Networking and Services* (p . 86–94).

Konishi, T., Tajimi, K., Sakata, N., & Nishida, S. (2009). Projection stabilizing method for palm-top display with wearable projector. In *ISWC 2009: Proceedings of the 13th international symposium on Wearable Computers* (p . 13–20).

Kratz, S., & Rohs, M. (2009). HoverFlow: expanding the design space of around-device interaction. In *MobileHCI 2009: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (p . 1–8).

Krueger, M. W. (1991). *Artificial reality II*. Addison-Wesley Reading.

Krumm, J. (2009). *Ubiquitous computing fundamentals*. Chapman & Hall/CRC, (p.17).

Kubicki, S., Lepreux, S., & Kolski, C. (2012). RFID-driven situation awareness on TangiSense, a table interacting with tangible objects. *Personal and Ubiquitous Computing*, *Volume 16*(Issue 8), p. 1079–1094.

Kurata, T., Sakata13, N., Kourogi, M., Okuma, T., & Ohta, Y. (2008). Interaction using nearby-and-far projection surfaces with a body-worn ProCam system. In *EI 2008: Proceedings of the Engineering Reality of Virtual Reality 2008 in the 20th Annual IS&T/SPIE Symposium on Electronic Imaging* (Vol 6804–16).

Kuroda, T., Tabata, Y., Goto, A., Ikuta, H., & Murakami, M. (2004). Consumer price data-glove for sign language recognition. In *Proceedings of 5th International Conference Disability, Virtual Reality Assoc. Tech.* (p . 253–258). Oxford, UK.

Lamming, M., & Flynn, M. (1994). Forget-me-not: Intimate computing in support of human memory. In *Proceedings of FRIEND21: International Symposium on Next Generation Human Interfaces* (Vol 94, p . 2–4). Tokyo, Japan: Institute for Personalized Information Environment.

Lee, J. C. (2008). Hacking the nintendo wii remote. *Pervasive Computing, IEEE*, *Volume 7*(Issue 3), p. 39–45.

Lehmann, E. (2006). *Nonparametrics: statistical methods based on ranks (POD)*. Prentice-Hall: 1st edition(1975). Springer(Berlin): Revised edition(2006).

Liao, C., Tang, H., Liu, Q., Chiu, P., & Chen, F. (2010). FACT: fine-grained cross-media interaction with documents via a portable hybrid paper-laptop interface. In *MM 2010: Proceedings of the international conference on Multimedia* (p . 361–370).

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *Volume 22*(Issue 140), p. 1–55.

Livingston, M. A., Rosenblum, L. J., Brown, D. G., Schmidt, G. S., Julier, S. J., Baillot, Y., … Maassel, P. (2011). Military applications of augmented reality. *Handbook of Augmented Reality*, *Part 2*, p. 671–706. doi:10.1007/978-1-4614-0064-6_31

Loclair, C., Gustafson, S., & Baudisch, P. (2010). PinchWatch: a wearable device for one-handed microinteractions. In *MobileHCI 2010 Workshop: Proceedings of the 12th international conference on MobileHCI Workshop on Ensembles of On-body Devices*.

Lukowicz, P., Timm-Giel, A., Lawo, M., & Herzog, O. (2007). Wearit@ work: Toward real-world industrial wearable computing. *Pervasive Computing, IEEE*, *Volume 6*(Issue 4), p. 8–13.

Lyons, K., Starner, T., & Gane, B. (2006). Experimental evaluations of the twiddler one-handed chording mobile keyboard. *Human-Computer Interaction*, *Volume 21*(Issue 4), p. 343–392.

Lyons, K., Starner, T., Plaisted, D., Fusia, J., Lyons, A., Drew, A., & Looney, E. W. (2004). Twiddler typing: One-handed chording text entry for mobile phones. In *CHI 2004: Proceedings of the SIGCHI conference on Human factors in computing systems* (p . 671–678).

Matias, E., MacKenzie, I. S., & Buxton, W. (1994). Half-QWERTY: typing with one hand using your two-handed skills. In *CHI 1994: Proceedings of the SIGCHI conference on Human factors in computing systems* (p . 51–52).

Matsumoto, T., Watanabe, R., Eguchi, S., Aritake, H., Morihara, T., & Yamagishi, F. (1993). *Heads-up display*. Google Patents.

McFarlane, D. C., & Wilder, S. M. (2009). Interactive dirt: increasing mobile work performance with a wearable projector-camera system. In *Ubicomp 2009: Proceedings of the 11th international conference on Ubiquitous computing* (p . 205–214).

Milgram, P., Drascic, D., Grodski, J., Restogi, A., Zhai, S., & Zhou, C. (1995). Merging real and virtual worlds. *Proceedings of IMAGINA'95*, p. 1–3.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, *Volume 77*(Issue 12), p. 1321–1329.

Mistry, P., & Maes, P. (2008). Quickies: Intelligent sticky notes. In *IET 2008: 4th International Conference on Intelligent Environments* (p . 1–4).

Mistry, P., Maes, P., & Chang, L. (2009). WUW-wear Ur world: a wearable gestural interface. In *CHI EA 2009: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (p . 4111–4116).

Mizell, D. (2001). Boeing's wire bundle assembly project. *Fundamentals of wearable computers and augmented reality*, p. 447–467.

Mori, G., Paternò, F., & Santoro, C. (2002). CTTE: support for developing and analyzing task models for interactive system design. *Software Engineering, IEEE Transactions on*, *Volume 28*(Issue 8), p. 797–813.

Morris, D. (2010). Emerging Input Technologies for Always-Available Mobile Interaction. *Foundations and Trends® in Human–Computer Interaction*, *Volume 4*(Issue 4), p. 245–316.

Morris, M. R., Wobbrock, J. O., & Wilson, A. D. (2010). Understanding users' preferences for surface gestures. In *Proceedings of Graphics Interface 2010* (p . 261–268).

Müller-Tomfelde, C. (2007). Dwell-based pointing in applications of human computer interaction. *INTERACT 2007: Proceedings of 11th IFIP international conference on Human-Computer Interaction*, p. 560–573.

Ngai, G., Chan, S. C. F., Cheung, J. C. Y., & Lau, W. W. Y. (2010). Deploying a Wearable Computing Platform for Computing Education. *IEEE Transactions on Learning Technologies*, *Volume 3*(Issue 1), p. 45–55.

Ni, T., & Baudisch, P. (2009). Disappearing mobile devices. In *UIST 2009: Proceedings of the 22th annual ACM symposium on User Interface Software and Technology* (p . 101–110).

Ni, T., Karlson, A. K., & Wigdor, D. (2011). AnatOnMe: facilitating doctor-patient communication using a projection-based handheld device. In *CHI 2011: Pro-*

ceedings of the SIGCHI conference on Human factors in computing systems (p . 3333–3342).

Nickel, K., & Stiefelhagen, R. (2007). Visual recognition of pointing gestures for human-robot interaction. *Image and Vision Computing*, *Volume 25*(Issue 12), p. 1875–1884.

Nicolai, T., Sindt, T., Witt, H., Reimerdes, J., & Kenn, H. (2006). Wearable computing for aircraft maintenance: Simplifying the user interface. In *IFAWC 2006: Proceedings of 3rd International Forum on Applied Wearable Computing* (p . 1–12).

Nusser, S., Miller, L., Clarke, K., & Goodchild, M. (2003). Geospatial IT for mobile field data collection. *Communications of the ACM*, *Volume 46*(Issue 1), p. 45–46.

Ota, S., Takegawa, Y., Terada, T., & Tsukamoto, M. (2010). A method for wearable projector selection that considers the viewability of projected images. *Computers in Entertainment (CIE): Theoretical and Practical Computer Applications in Entertainment*, *Volume 8*(Issue 3), Article No. 17.

Podlaseck, M., Pinhanez, C., Alvarado, N., Chan, M., & Dejesus, E. (2003). On interfaces projected onto real-world objects. In *CHI EA 2003: Proceedings of the SIGCHI conference extended abstracts on Human factors in computing systems* (p . 802–803).

Potter, R. L., Weldon, L. J., & Shneiderman, B. (1988). Improving the accuracy of touch screens: an experimental evaluation of three strategies. In *CHI 1988: Proceedings of the SIGCHI conference on Human factors in computing systems* (p . 27–32).

Pryor, H. L., Furness, T. A., & Viirre, E. (1998). The virtual retinal display: A new display technology using scanned laser light. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol 42, p . 1570–1574).

Raskar, R., Beardsley, P., Van Baar, J., Wang, Y., Dietz, P., Lee, J., … Willwacher, T. (2004). RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors. In *ACM Transactions on Graphics (TOG)* (Vol 23, p . 406–415).

Raskar, R., Van Baar, J., Beardsley, P., Willwacher, T., Rao, S., & Forlines, C. (2003). iLamps: geometrically aware and self-configuring projectors. In *ACM Transactions on Graphics (TOG)* (Vol 22, p . 809–818).

Rekimoto, J., & Ayatsuka, Y. (2000). CyberCode: designing augmented reality environments with visual tags. In *DARE 2000: Proceedings of conference on Designing augmented reality environments* (p . 1–10).

Rhodes, B. J. (1997). The wearable remembrance agent: A system for augmented memory. *Personal and Ubiquitous Computing*, *Volume 1*(Issue 4), p. 218–224.

Rogers, M. S., Barr, A. B., Kasemsontitum, B., & Rempel, D. M. (2008). A three-dimensional anthropometric solid model of the hand based on landmark measurements. *Ergonomics*, *Volume 51*(Issue 4), p. 511–526.

Rolland, J. P., & Fuchs, H. (2000). Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoperators & Virtual Environments*, *Volume 9*(Issue 3), p. 287–309.

Roudaut, A., Lecolinet, E., & Guiard, Y. (2009). MicroRolls: expanding touch-screen input vocabulary by distinguishing rolls vs. slides of the thumb. In *CHI 2009: Proceedings of the SIGCHI conference on Human factors in computing systems* (p . 927–936).

Rouillard, J. (2008). Contextual QR codes. In *ICCGI'08: The Third International Multi-Conference on Computing in the Global Information Technology* (p . 50–55).

Rouillard, J., & Laroussi, M. (2008). PerZoovasive: contextual pervasive QR codes as tool to provide an adaptive learning support. In *CSTST 2008: Proceedings of the*

5th international conference on Soft computing as transdisciplinary science and technology (p . 542–548).

Rukzio, E., Holleis, P., & Gellersen, H. (2012). Personal projectors for pervasive computing. *Pervasive Computing, IEEE*, *Volume 11*(Issue 2), p. 30–37.

Salminen, S. (2001). Wearable Technology.

Schöning, J., Löchtefeld, M., Rohs, M., & Krüger, A. (2010). Projector Phones: a new class of interfaces for augmented reality. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, *Volume 2*(Issue 3), p. 1–14.

Shaer, O., & Hornecker, E. (2010). Tangible user interfaces: past, present, and future directions. *Foundations and Trends in Human-Computer Interaction*, *Volume 3*(Issue 1–2), p. 1–137.

Siek, K., Rogers, Y., & Connelly, K. (2005). Fat finger worries: How older and younger users physically interact with PDAs. *Human-Computer Interaction-INTERACT 2005*, *Volume 3585/2005*, p. 267–280. doi:DOI: 10.1007/11555261_24

Siewiorek, D., Smailagic, A., Bass, L., Siegel, J., Martin, R., & Bennington, B. (1998). Adtranz: a mobile computing system for maintenance and collaboration. In *Digest of Papers. Second International Symposium on Wearable Computers* (p . 25–32).

Smailagic, A., & Kenn, H. (2011). New Advances in Wearable Computing. *Pervasive Computing, IEEE*, *Volume 10*(Issue 4), p. 96–100.

Spitzer, M. B., Rensing, N., McClelland, R., & Aquilino, P. (1997). Eyeglass-based systems for wearable computing. In *Digest of Papers. First International Symposium on Wearable Computers* (p . 48–51).

Starner, T., Leibe, B., Singletary, B., & Pair, J. (2000). MIND-WARPING: towards creating a compelling collaborative augmented reality game. In *IUI 2000: Proceed-*

ings of the 5th international conference on Intelligent User Interfaces (p . 256–259).

Starner, T., Mann, S., Rhodes, B., Healey, J., Russell, K. B., Levine, J., & Pentland, A. (1995). Wearable computing and augmented reality. *The Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, MIT Media Lab Vision and Modeling Group Technical Report*, *Volume 355*.

Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., … Pentland, A. (1997). Augmented reality through wearable computing. *Presence: Teleoperators and Virtual Environments*, *Volume 6*(Issue 4), p. 386–398.

Sung, M., DeVaul, R., Jimenez, S., Gips, J., & Pentland, A. (2004). Shiver motion and core body temperature classification for wearable soldier health monitoring systems. In *ISWC 2004: Proceedings of the 8th international symposium on Wearable Computers* (Vol 1, p . 192–193).

Tajimi, K., Uemura, K., Kajiwara, Y., Sakata, N., & Nishida, S. (2010). Stabilization Method for Floor Projection with a Hip-Mounted Projector. In *ICAT 2010: Proceedings of 20th international conference on Artificial Reality and Telexistence* (Vol 10, p . 77–83).

Tamaki, E., Miyaki, T., & Rekimoto, J. (2009). Brainy hand: an ear-worn hand gesture interaction device. In *CHI EA 2009: Proceedings of the 27th of the international conference extended abstracts on Human factors in computing systems* (p . 4255–4260).

Van Dam, A. (1997). Post-WIMP user interfaces. *Communications of the ACM*, *Volume 40*(Issue 2), p. 63–67.

Vlahakis, V., Ioannidis, M., Karigiannis, J., Tsotros, M., Gounaris, M., Stricker, D., … Almeida, L. (2002). Archeoguide: An augmented reality guide for archaeological sites. *Computer Graphics and Applications, IEEE*, *Volume 22*(Issue 5), p. 52–60.

Vogel, D., & Baudisch, P. (2007). Shift: a technique for operating pen-based interfaces using touch. In *CHI 2007: Proceedings of the international conference on Human factors in computing systems* (Vol 1, p . 657–666).

Von Hardenberg, C., & Bérard, F. (2001). Bare-hand human-computer interaction. In *PUI 2001: Proceedings of the 2001 workshop on Perceptive user interfaces* (p . 1–8).

Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. *Communications of the ACM*, *Volume 54*(Issue 2), p. 60–71.

Wachs, J. P., Stern, H. I., Edan, Y., Gillam, M., Handler, J., Feied, C., & Smith, M. (2008). A gesture-based tool for sterile browsing of radiology images. *Journal of the American Medical Informatics Association*, *Volume 15*(Issue 3), p. 321–323.

Wang, R. Y., & Popović, J. (2009). Real-time hand-tracking with a color glove. *SIGGRAPH 2009: ACM Transactions on Graphics (TOG)*, *Volume 28*(Issue 3), Article No. 63.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, *Volume 265*(Issue 3), p. 94–104.

Wellner, P. (1993). Interacting with paper on the DigitalDesk. *Communications of the ACM*, *Volume 36*(Issue 7), p. 87–96.

Wellner, P., Mackay, W., & Gold, R. (1993). Back to the real world. *Communications of the ACM*, *Volume 36*(Issue 7), p. 24–26.

Wigdor, D., Forlines, C., Baudisch, P., Barnwell, J., & Shen, C. (2007). Lucid touch: a see-through mobile device. In *UIST 2007: Proceedings of the 20th annual ACM symposium on User Interface Software and Technology* (p . 269–278).

Wilson, A. D. (2005). PlayAnywhere: a compact interactive tabletop projection-vision system. In *UIST 2005: Proceedings of the 18th annual ACM symposium on User Interface Software and Technology* (p . 83–92).

Wilson, M. L., Robinson, S., Craggs, D., Brimble, K., & Jones, M. (2010). Pico-ing into the future of mobile projector phones. In *CHI EA 2010: Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems* (p . 3997–4002).

Zhang, X., Navab, N., & Liou, S. P. (2000). E-commerce direct marketing using augmented reality. In *ICME 2000: IEEE International Conference on Multimedia and Expo* (Vol 1, p . 88–91).

Zhou, Y., David, B., & Chalon, R. (2011). Innovative user interfaces for wearable computers in real augmented environment. In *HCI International 2011: Human-Computer Interaction. Interaction Techniques and Environments* (p . 500–509). Springer-Verlag Berlin/Heidelberg.

Publications

1. Zhou Y., David B T., Chalon R.: **Innovative User Interfaces for Wearable Computers in Real Augmented Environment.** HCI International 2011, 9-14 July 2011, Orlando, Florida, USA. In J.A. Jacko (Ed.): Human-Computer Interaction, Part II, HCII 2011, LNCS 6762, pp. 500-509, Springer-Verlag Berlin/Heidelberg (2011).
2. Zhou Y., David B T., Chalon R.: **MobilePaperAccess: Ubiquitous Paper-based Interfaces for Mobile Interactions**. The 3rd International Conference on Ambient Systems, Networks and Technologies (ANT' 2012), August 27-29, 2012, Niagara Falls, Ontario, Canda, Procedia Computer Science, Volume 10, Elsevier B.V., 2012, Pages 282-291. [On-line: http://dx.doi.org/10.1016/j.procs.2012.06.038] (2012)
3. Zhou Y., David B T., Chalon R.: **PlayAllAround: Wearable One-hand Gesture Input and Scalable Projected Interfaces**. ERGO-IHM 2012 Ergonomie et Interaction Homme-Machine, Bidart 16 octobre - Biarritz 17, 18, 19 octobre, ACM Digital Library (2012)
4. David B T., Zhou Y., Xu T., Chalon R.: **Mobile User Interfaces and their Utilization in a Smart City**, ICOMP'11 - The 2011 International Conference on Internet Computing as part of WorldComp'2011 Conference, Las Vegas 18-21 July 2011. (2011)
5. Xu T., David B T., Chalon R., Zhou Y**.: A Context-aware Middleware for Ambient Intelligence**. In Proceedings of 12th International Middleware Conference, Lisbon, Portugal. Middleware Posters'2011. (2011)
6. David B T., Yin C., Zhou Y., Xu T., Zhang B., Jin H., Chalon, R**., SMART-CITY: problematics, techniques and case studies**,

ICCM'2012 8th International Conference on Computing Technology and Information Management, April 24-26, 2012, Seoul, Korea (2012)

7. Zhang B., David B.T., Chalon R., Yin C., Zhou Y.: **Contextual Mobile Learning for professionals working in 'Smart City'**. Workshop SCiLearn at The 11th International Conference on Web-based Learning (ICWL'12), 2-4 September 2012, Sinaia, Romania, (2012)