

Innovative wearable interfaces: an exploratory analysis of paper-based interfaces with camera-glasses device unit

Yun Zhou · Tao Xu · Bertrand David · René Chalon

Received: 10 December 2012 / Accepted: 17 May 2013 / Published online: 3 July 2013
© Springer-Verlag London 2013

Abstract The new ubiquitous interaction methods change people's life 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. We found two solutions to enable ubiquitous interaction and put a stop to the limits imposed by the desktop mode: namely nomadism and mobility. Based on these two solutions, we have proposed three interfaces (Zhou et al. in HCI international 2011: human–computer interaction. Interaction techniques and environments, Springer, Berlin, pp 500–509, 2011): in-environment interface (IEI), environment dependent interface (EDI), and environment independent interface (EII). In this paper, we first discuss an overview of IEI, EDI, and EII, before excluding IEI and focusing on EDI and EII, their background, and distinct characteristics. We also propose a continuum from physical paper-based interface to digital projected interface in relation with EDI and EII. Then, to validate EDI and EII concepts, we design and implement a MobilePaperAccess system, which is a wearable camera-glasses system with paper-based interface and original input techniques allowing mobile interaction. Furthermore, we

discuss the evaluation of the MobilePaperAccess system; we compare two interfaces (EDI and EII) and three input techniques (finger input, mask input, and page input) to test the feasibility and usability of this system. Both the quantitative and qualitative results are reported and discussed. Finally, we provide the prospects and our future work for improving the current approaches.

Keywords Wearable interfaces · Input techniques · Augmented paper · Contextualization · Mobility

1 Introduction

With the emergence 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 [40], and innovative inputs and techniques are being increasingly considered. These new interaction methods change people's life 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 interaction problems of user input and output modalities to the more ethical problems of privacy, data protection, or even the 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

Y. Zhou (✉)
LIRIS lab., UMR 5205, CNRS, Université Lyon 1,
Bâtiment Blaise Pascal, 69621 Villeurbanne Cedex, France
e-mail: chouyun920@gmail.com

T. Xu · B. David (✉) · R. Chalon
LIRIS lab., UMR 5205, CNRS, Ecole Centrale de Lyon,
36 Avenue Guy de Collongue,
69134 Lyon-Ecully Cedex, France
e-mail: Bertrand.David@ec-lyon.fr

T. Xu
e-mail: daridxu@gmail.com

R. Chalon
e-mail: Rene.Chalon@ec-lyon.fr

modality. A sophisticated mobile environment requires a dedicated interface design, involving input and output techniques with new emerging features offering far more than the capacities of traditional techniques.

One of the available solutions 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 a webcam, a pico-projector, or other output displays. Mobile devices can include PDAs and smart mobile phones. Classical portable devices such as laptops cannot be included 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 take longer to access input compared with mobile phones. However, the tablet or the special laptop could form one part of a wearable configuration, contributing only 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 innovative interfaces based on the aforementioned solutions [46]: 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. In this situation, a fixed webcam and a wall video projector are appropriately located to allow in-environment interaction. The user uses his/her hands to interact with the public information available from a public wall, e.g., searching and browsing. The EDI and EII are both based on the user's wearable computer devices, allowing him/her to interact in mobility. We aim to provide the user with information that is decided by the environment, i.e., the environment provides users with information. In this way, the EDI refers to the strong relationship between the interface and the in-environment information. Going one step further, we propose the EII, which refers to the relationship between the interface and personal information. The actual users perform contextualization by showing the webcam appropriate predefined markers or menus. Users can thus contextualize their working environment by themselves without any contact with the environment.

In this paper, we first outline the concepts of IEI, EDI, and EII, and then focus on discussing EDI and EII, their background, and distinct characteristics. Next, to concretize the EDI and EII, we propose the MobilePaperAccess system, a ubiquitous paper-based interface for mobile interaction. We employ the following wearable configuration: a small screen attached to a goggle to provide visual information, a webcam to pick up the input signal, and a laptop as the calculating device. Our goal is to create a true

contextualization, based on the user's location or independent of it, which is more effective and adaptive to users' information needs by taking advantage of dynamic and physical environmental characteristics. Finally, we explain the evaluation with the aim of investigating input techniques as well as two interfaces: EDI and EII.

2 Related work

In this section, we outline the relevant research work that helped inspire our study on wearable interaction in relation with ubiquitous computing. Since ubiquitous computing covers a large number of aspects, we only address input and output techniques in the field of wearable computing and in the related research area: paper augmented interaction.

2.1 Wearable input techniques

The term "ubiquitous computing" was introduced by Weiser [42] 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 users to access 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, seamlessly immersed in the physical world with digital information. Early in 1993, Starner [38], one of the wearable computing pioneers from the MIT Media Laboratory, had attempted a heads-up display integrated with his glasses and a twiddler [26, 27] as the input device which can be located in the pocket. In recent years, miniaturization of mobile and wearable devices has made ubiquitous computing possible, and the search for mobile input and output modalities has become a research focal point. Input techniques fall into a wider range of approaches, including styluses [24], the digital glove [14], and mobile sensors to recognize hand gestures [17, 41] or objects [12]. The technology proposed by Skininput [18] employs the user's body as the interactive surface, such as the touch pad with bio-acoustic sensors and projector, which provides an always-available interface [31]. Minput [17] offers an input technique via gestures like flicking and twisting, which is carried out by two optical tracking sensors on the back of a small device. MotionBeam [44] is a novel interaction metaphor, based on the input via the projector movement: The user can navigate by changing the location and orientation of the projector. Besides the movement interaction of projector, researchers also focus on manipulating the dynamic

projection surface. OmniTouch [16] allows bare hand gestures as input, while SixSense [30] explores and proposes marked finger gestures as input. Both use the dynamic projection interface. Ni and Baudisch [32] investigated spatial interaction using the hand gesture as the input and the zero visual feedback as the output in disappearing mobile. They studied the limits of miniaturization of mobile devices, what the smallest future devices might be, as well as how the user would interact with these smallest devices.

2.2 Wearable output techniques

Compared with output modalities such as 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 rich interactive elements of GUIs. We focus on the visual output for feedback in this paper. 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 that is as large as possible. As wearable output visual displays, head-worn displays [15, 37], handheld mobile phones [3, 7, 43], and pico-projectors [35, 45] have been used to present information.

Small screen displays such as head-worn displays and mobile phones have several advantages. However, one drawback persists: These displays cannot avoid the limitation due to the small-size screen, in which visual output information content is restricted in a scale. 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 provide 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. In recent years, miniaturization of projectors has led to the emergence of mobile devices with embedded projector or standalone palm-size pico-projectors. The pico-projector, as a mobile display, has high full scalability and supports scalable interaction. In this way, the pico-projector can provide both small-size and large-size display experience. However, the properties of scalability and dependence on surrounded surfaces have given rise to the challenges for interaction with pico-projectors. It is challenging to project the interface in a high resolution on different surfaces of different colors, textures, and sizes, especially to provide the appropriate scalable interface. Besides, the insufficient lumens limit feasibility of use in daily light. Although researchers have investigated the wearable camera-projector system in many-

sided aspects, such as OmniTouch [16], SixthSense [30], and Brainy hand [39], the aforementioned problems have not yet all been solved.

2.3 Paper interaction

Ishii and Ullmer [22] have defined the tangible user interface (TUI) 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 if the terms related to TUIs vary, they share the same basic paradigm [11]: Users use their hands to manipulate some physical objects via physical gestures; a computer system detects this, alters its state, and gives feedback accordingly. Paper interaction is one of the tangible user interfaces [21]. Studies on paper interaction and paper interfaces [1, 19, 29] focused on augmented reality, and attempt to merge 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 [19] describes a projecting window prototype able to simulate 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 his fingers, hands, and stylus. The Quickies [29] system uses augmenting sticky notes as an I/O interface. The DisplayObjects [1] 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 and the desktop interaction—we choose to focus on paper interaction in the mobile situation.

In addition to the paper-based interface, tangible objects are themselves employed as tags and reminders, utilized to trigger 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 [36], or in the target area. RFID, ARToolKit markers, and QR codes are most often used for link tagging. In the TUI context, computer vision is often used to sense the 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 [20, 33, 34] that have made it possible to use contextual markers in a mobile environment. Furthermore, compared with other detection technologies such as RFID [2, 23], the ARToolKit tag (or QR code) is based on 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.

3 Overview of innovative user interfaces

As we stated above, one solution to enable ubiquitous interaction and put an end to the limits of the desktop mode is nomadism, in which the user is physically mobile and not equipped with any wearable or mobile device. Another possible solution to this problem is mobility, in 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 can be achieved by interacting with the EDI and EII.

In this section, we shall first explain the three innovative user interfaces (IEI, EDI, and EII) by discussing the relationships between three interfaces, three main elements, and contextualization styles. We then provide the scenarios for three interfaces. Next, we describe the principal and essential characteristics of EDI and EII. Finally, we propose a basic continuum that spans the range from physical interface to digital interface, based on the interaction techniques of our EDI and EII design.

3.1 Innovative user interfaces

Figure 1 represents the relationships between 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 that is dependent on the in-environment indication and information. As illustrated in Fig. 1, 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.), while the latter requires the environment, the wearable devices, and the user. Since both the IEI and EDI are dependent on the in-environment information, the contextualization style is environment-contextualization. Furthermore, the EII is independent from the environment, namely it relies neither on the in-environment information nor on the environment configurations. In this way, users

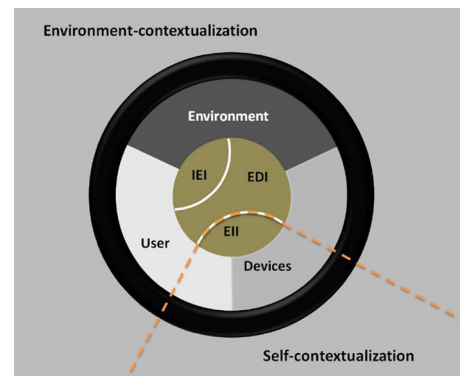


Fig. 1 An overview of IEI, EDI, and EII, with their elements and contextualization style

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 Fig. 1.

Consider the scenarios in the smart city [9] as follows:

Scenario 1 Li and Yan are research members, and they work in the same laboratory. One day, Li wants to discuss something with Yan, but when he knocks at Yan's door, he finds that Yan is out of the office. So Li walks to the public place outside the laboratory and browses Yan's public information via hand gestures (see Fig. 2a). He checks Yan's schedule, looks for an appropriate time, and sends him a date request. After obtaining feedback from the system, he returns to his office and continues to work.

Scenario 2 One day, Li wants to discuss with Yan, but Yan is not available. Outside Yan's door, Li sees a predefined paper interface pasted on the door, and he is wearing his wearable devices (see Fig. 2b). He then checks Yan's schedule, finds an available time, and sends a date request via a paper interface. After obtaining feedback, he returns to his office.

Scenario 3 One Saturday in a library, Li is looking through books when he suddenly remembers that he needs to discuss something with another new member John. So he opens his notebook and finds a predefined paper-based interface (see Fig. 2c). With this interface, he fixes an appointment with John. Or, in another way, he directly projects the interface on his table. After setting this digital appointment, he continues to look for books in the library.

The first scenario interprets the IEI, the second scenario explains the EDI, while the third one describes the EII. The IEI, EDI, and EII can solve the same problems that the user encounters, as well as solve distinct problems. In everyday life, it is essential to make appointments with people.

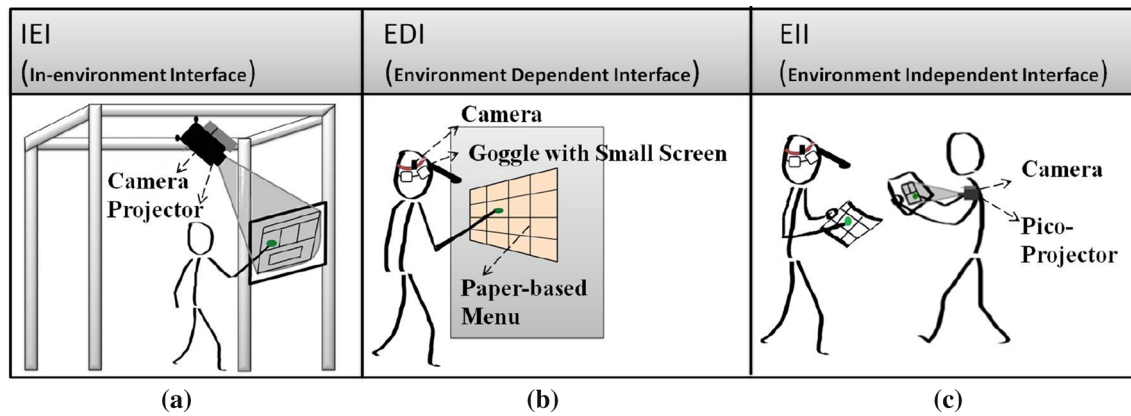


Fig. 2 Innovative user interfaces. a IEI. b EDI. c EII

Mobile innovative interfaces support the user in checking their schedule and making appointments, either dependent on or independent of the environment context.

In this paper, we exclude IEI, and mainly focus on the last two interfaces: EDI and environment independent interface. The EDI and EII are both based on users' wearable computing devices, allowing them to interact in mobility. With respect to configuration, the EDI and EII can use 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.

3.2 Environment dependent interface

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 the appliance, wall, book, door, and so on. In this way, public and professional guiding information can be used for contextualization.

We have studied the research field of augmented reality in relation to mobility for several years. The previous work can be characterized by two acronyms: MOCOCO (MObility, COntextualization and COoperation), and IMERA [8] (French acronym for mobile interaction with real augmented environment). Augmentation can be achieved in a conscious way, passively or actively, or in an unconscious way. 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., RFID) 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 paper, we are mainly concerned with the approach of conscious augmentation using passive markers. For the purpose of providing the user with in-environment information and interface with environment-contextualization, we investigate passive marker augmentation that 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 EDI is concretized via the passive marker augmentation method. The markers act as bridges linking the real environment and the digital information, and can be pasted on a wall, a notice board, an information board of a bus shelter, and appliances or a doorplate.

It is essential to define the distinct characteristics for EDI (see Fig. 3). The EDI must be closely related to in-environment information, which is dependent on the specific location. The location can be identified through either passive in-environment physical markers, or specific menus, or indications that 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 EDI.

3.3 Environment independent interface

Going one step further, we also explore both marker augmentation and non-marker augmentation to support and concretize the EII. The EII user can interact with projected

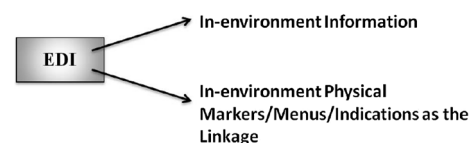


Fig. 3 The principal and essential characteristics of EDI

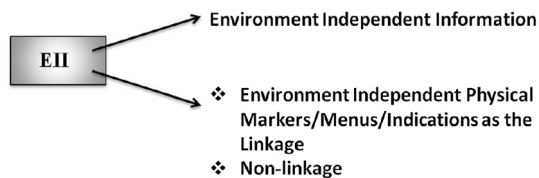


Fig. 4 The principal and essential characteristics of EII

dynamic information on the situation of non-marker augmentation. With respect to marker augmentation, when 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.

Environment independent information plays an important role in EII (see Fig. 4). Digital information in EII is summoned with no relation to the environment, and is not dependent on 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 from the location of the environment. Linkage for the EII is optional: non-linkage augmentation can be achieved by pure digital personal projection.

3.4 Continuum 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 the interaction techniques of our EDI and EII design (see Fig. 5). The physical interface surface is static and inflexible, usually unitary planar, considered as uni-planar. Since the elements in the interface are fixed and physical, these elements should all evolve in the uni-planar, rather than in the multilayer windows. In our study, we use the paper-based interface as the realization of the physical interface, namely all the interactive elements are predefined and printed on a piece of paper for interaction. The physical-digital interface incorporates the physical with the digital interface, in which the paper interface has been augmented with the projected interface in possession of the half-dynamics. Furthermore, the digital interface has full dynamics, which provides projected personal information for interaction. The last two interfaces are based on the multi-planar, by means of which interactive elements are organized logically in the dynamic multilayer windows.

In the EDI system, the interface relies closely on the environment and the context information, such as location. In other words, the presentation of the interface is not dependent on the individual's decision but rather on the environment. Based on this dependence, the EDI builds on the physical interface and the physical-digital interface,

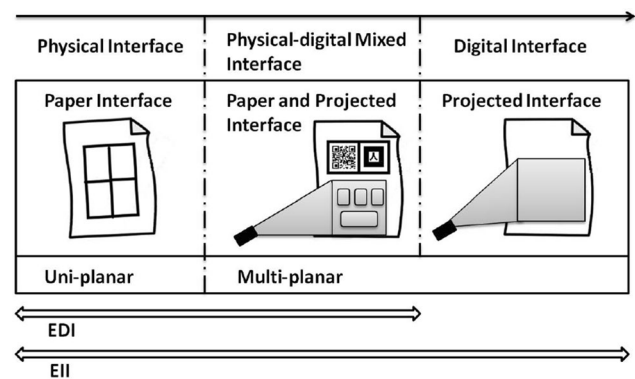


Fig. 5 The continuum from physical interface to digital interface for the EII and EDI

where the physical part is linked to the environment. In the EII system, the interface is determined by the actual individual, and can either be augmented with markers or predefined menus, or augmented by the required projected information. Thus, the EII entirely spans the physical to the digital interface.

4 Design of MobilePaperAccess

To implement our EDI and EII, we have designed and developed a ubiquitous paper-based system for mobile interaction, known as MobilePaperAccess. This is a wearable camera-glasses system with a paper-based interface allowing mobile interaction. We access in-environment digital information or environment independent information from the paper interface. In this section, we shall discuss the design of input techniques and paper surface.

4.1 Input techniques

We propose three input techniques as shown in Fig. 6: finger hover input (see Fig. 6a), mask input (see Fig. 6b), and page input (see Fig. 6c), all of which are used for selection.

One of the hand gesture solutions for users' selection input is to let the user hover for a second with his/her finger, while the selection signal can be generated via a span. When the user points at an interactive item such as a button, he/she needs to remain in the position of this item for a time interval. The interactive item is thus considered as selected and validated. Buxton specifies a three-state input model [6], which provides a conceptualization of some basic properties of input devices and interactive techniques. We utilize the three-state input model to explain the finger hover gesture, illustrated in Fig. 7. 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

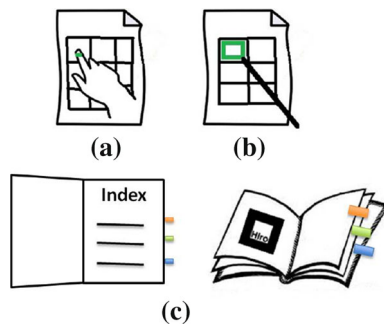


Fig. 6 Three input techniques. **a** Hover input technique. **b** Frame mask input technique. **c** Page input technique

vision, meaning that any movement of the finger has no effect on the system. As the finger enters the region of the webcam (state 1), the system starts to track: the tracking symbol is the tip of the user’s index finger. The two actions, “Hovering for a Second” and “Stop Hovering,” are closely linked, similar to the relationship between opening and closing a door. In this way, the “Stop Hovering” action is non-substitutable and closely linked to the preceding action. Thus, the return path from state 2 to state 1 is drawn in gray as shown in Fig. 7.

In addition to the finger selection technique, we propose a mask selection technique, which shares the same hovering method with the finger input. The mask consists of a rectangular green frame and a wand. The frame is in charge of selection, while the wand is held in the user’s hand for convenience. The real information printed on the paper can be read inside the frame.

For page input, we place only one marker on each page of a booklet. The user can show the webcam one marker at a time by flipping through the pages. We use a predefined booklet of several pages where each page contains an ARToolKit tag. The index in front of the booklet allows the user to access to the appropriate page. Also, the color

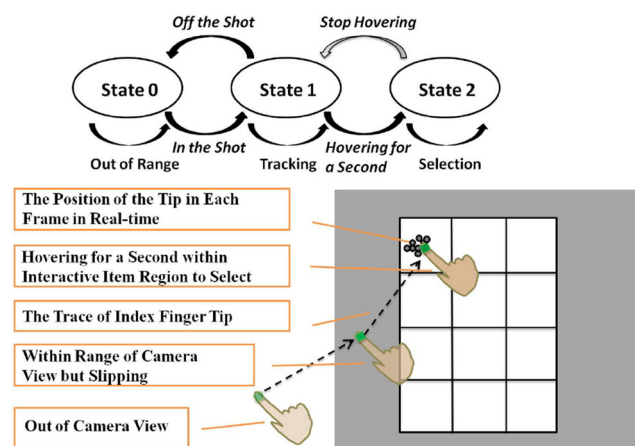


Fig. 7 The three-state model of the Hover gesture input and illustration

indicators related to each page on the side edge of booklet can facilitate the operation.

4.2 Paper surface

According to human factors (Fig. 8), the eye rolling angle is 15° comfort, and 35° maximum horizontally, and 30° up and 35° down vertically. The average forward grip reach is 74.3 cm [10]. The interactive surface held in hand should be less than 34.64 cm × 16.08 cm in size when reading distance is 30 cm. Thus, we select an A4 (29.7 cm × 21.0 cm) paper pasted on the wall as the EDI, an A4 paper held in the hand as the environment independent interface, and a predefined booklet held in the hand as the environment independent interface. We organize the layout in the comfortable range. The user thus does not need to move his/her head too much when reading the interface.

We segment the paper surface into several rectangular zones (see Fig. 9), and relate each zone with a unique event. The user can trigger the required action by selecting the relative zone. To ensure the real rectangular zones are recognized, we place the ARToolKit tags or color markers on the paper surface to assist augmentation. We also propose some examples of ARToolkit tag and color marker arrangement. What is most important is that fingers and hands should not occlude the ARToolKit tags or color markers during interaction. We place the tags in the left upside (see Fig. 9a) or upside position (see Fig. 9b) for right-handed users, while we place the tags in the right upside, right, or upside position for left-handed users. The two color markers are located at the ends of the diagonal lines: We place the tags left top and right bottom for left-handed users, and right top and left bottom for right-handed users (see Fig. 9c, d).

5 Implementation

The perspective of our MobilePaperAccess system includes the paper interactive surfaces augmented with the color

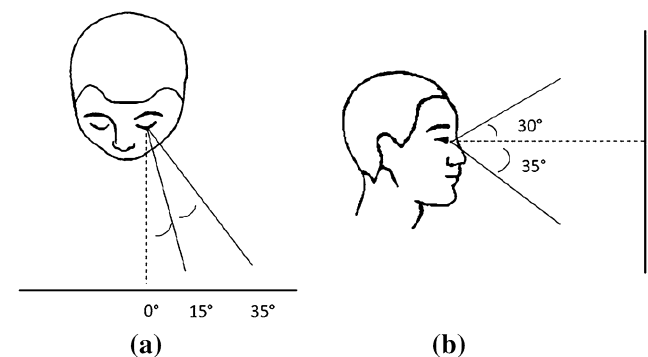


Fig. 8 The angle of eye rolling vertically (a) and horizontally (b)

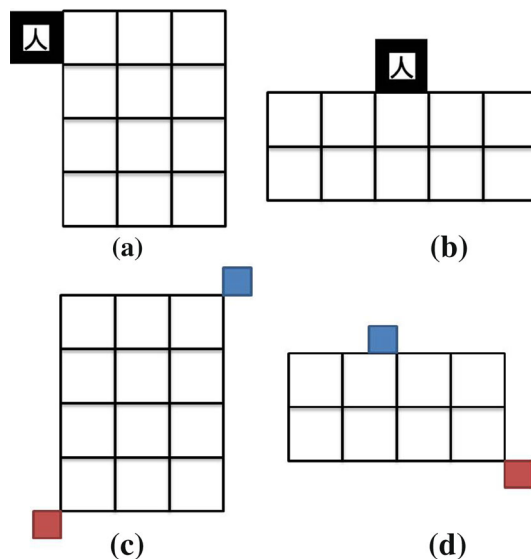


Fig. 9 The arrangement of the paper-based interface

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 capture the ARToolKit tags, the goggle with small screen to present the digital information, and a laptop for calculating. We will explain the implementation below with respect to wearable configuration, figure and mask motion, augmented paper and digital feedback, as well as applications.

5.1 Wearable configuration

Our wearable configuration consists of the camera-glasses device unit, and a laptop for calculating. The camera-glasses device unit described in this paper is made up of a RGB 640×480 webcam and a goggle attached with a small screen (see Fig. 10). We fix the webcam on a plastic hair band and let the user wear it on his/her forehead as shown in the figure below. The camera thus sees what the user sees as the user turns his/her head, and the small screen displays digital information precisely in the user's field of vision. As the user turns his/her head, the digital feedback follows the required direction. The viewer display is a Micro-Optical SV-6 PC viewer, with a resolution of 640×480 pixels. The laptop is equipped with a multi-touch screen, which can be used as a tablet and carried on the back or in a messenger bag along the body (see Fig. 10).

5.2 Finger and mask motion

Our three input techniques are based on computer vision techniques. In the current work, we use the object tracking algorithm based on the Camshift algorithm [5] by



Fig. 10 Wearable configurations

employing the OpenCV library [4]. 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. Thirdly, the back projection of the processed image is calculated, and the Camshift tracks distribution of the target color feature based on the back projection. We can thus automatically track the color marker located on the index finger. As shown in Fig. 11, we record the trace of the color marker by noting the x and y coordinates of the color marker in each frame. As shown in the figure, we count the number of tracking points in each interactive item area such as the gray 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.

5.3 Augmented paper and digital visual feedback

Implementation of output techniques includes augmented paper surface and digital visual feedback. Unlike the devices where the input takes place directly on the display surface, the digital display and the input of MobilePaper-Access are separated. Each paper-based interactive surface is either augmented with the color markers or with ARToolKit tags (see Fig. 12b). Take color markers as an example: Two color markers in a diagonal position (see Fig. 12a, d) shape a rectangle, which can be tracked by the webcam. As long as the webcam recognizes this rectangular shape, the grid within the shape is considered as icons and can be activated by pointing. The user is unaffected even if he/she rotates or moves the paper slightly during interaction. Besides, the booklet for interaction is augmented with ARToolKit tags, and each page has a tag as the identity (see Fig. 12c).

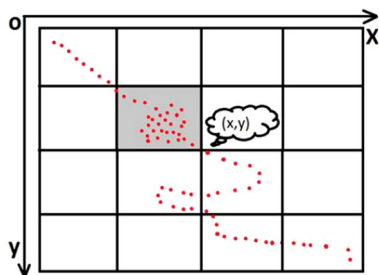


Fig. 11 The motion of the index finger

As regards digital feedback, the feedback information is presented on the small screen fixed on the right or left side of the goggle. Screen size limitation means that the display area is divided into two parts: the main display area and the auxiliary area (see Fig. 13). The main display area displays the information completely, while the auxiliary area displays the brief response of the information in the form of a keyword or tips, permitting a quick and just-in-time understanding of the user.

5.4 Applications

5.4.1 Research Team Management Application (RTMA)

To prove the concepts of EDI and EII, we developed an application known as the Research Team Management Application (RTMA) with the goal of managing research team members' exchanges. RTMA is based on the scenario with EDI and EII, as we stated above in the Sect. 3. With the same wearable configuration, the user consults a member's schedule using a predefined interface pasted in advance in the laboratory, or a customized paper, or a booklet held in his/her hand.

Fig. 12 The augmented paper with color markers (a), (d), with ARToolKit tag (b), augmented book

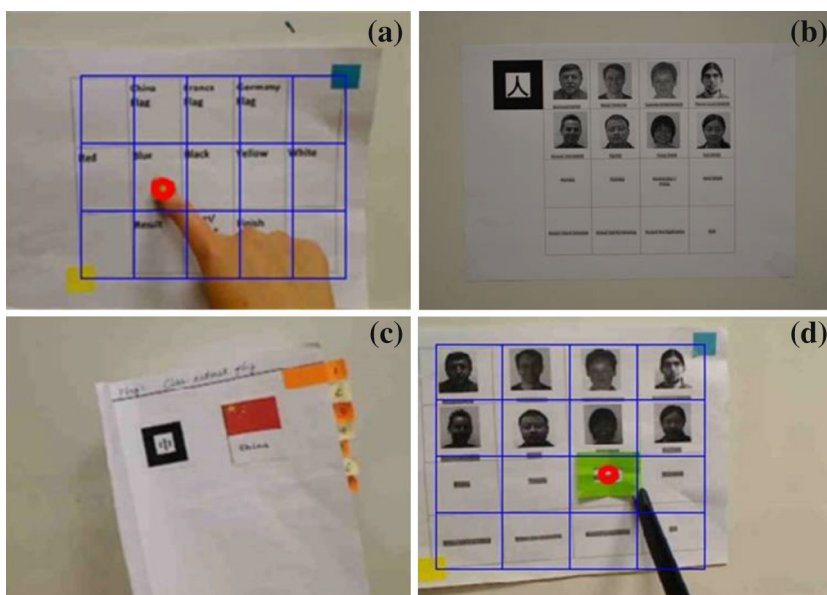


Fig. 13 The visual feedback in the small display

5.4.2 Flag application

We also propose a playful application called the Flag Application for the user to explore innovative interfaces with input and output techniques. From the Flag Application, the user can first select the name of his/her targeted national flag. He then starts to input the color composition of this flag. Finally, he can verify the result of his input composition. In the EDI scenario, the user interacts with a piece of predefined sheet pasted on the in-environment surface, while in the EII scenario, the user plays the Flag Application with a handheld predefined interface such as the surface of a notebook.

6 Evaluation and main results

To obtain a more thorough understanding of EDI and EII, as well as the creation of the MobilePaperAccess system

Table 1 Four test cases (Case A, B, C, and D)

Interfaces	Input techniques		
	Finger	Mask	Book/page
EDI	√ Case A	√ Case B	Not studied
EII	√ Case C	Not studied	√ Case D

and input modalities, we designed a structured evaluation comparing our three input techniques (finger input, mask input, and page input) and two interfaces (EDI and EII). To compare the three input techniques with two interfaces, we formed four cases as shown in Table 1 and provided the description of four cases in the figure. For Cases A and B, the participants stood, whereas for Cases C and D, the users sat or stood freely to simulate mobility. In this evaluation, we set and explored the three following research questions:

- *Question 1* Are three input techniques and our innovative input and output modality easy to learn or not?
- *Question 2* What is the performance of the four cases during the interaction?
- *Question 3* Does Fitts's law [28] have any influence on interaction time of wearable interfaces?

6.1 Participants and procedure

We recruited 12 student participants, 7 males and 5 females, aged between 19 and 29 with a mean age of 23.2. Their heights ranged from 157 to 188 cm, with an average of 171.8 cm. All participants had experience in using mobile devices. Only 6 of them had knowledge of human–computer interaction (HCI) such as reading the relevant books or taking classes in courses for introduction to HCI. All except one were right-handed.

We provided two types of program for each case: the toy application and the true application. The toy application was the Flag Application 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 input techniques and interfaces. They could play the Flag Application several times until they felt competent in the following true tasks. The true application was the RTMA, in each case of which we provided two tasks for the user to perform: task T1 and task T2. In this way, for each case, each participant had to perform one toy application with one learning task, and one true application with two tasks, i.e., in all four cases, each person had to perform 4 learning tasks plus $2 \times 4 = 8$ true tasks.

As the procedure shown in Fig. 14, the evaluation began with an explanation of the protocol by the text form, including the instruction and questionnaire. The questionnaire attached to the protocol contained two parts: The first part (pre-questionnaire) covered the background questions of individuals (age, gender, height, etc.) and questions on their familiarity with mobile devices and HCI, to be answered by the users before the test; the second part (post-questionnaire) provided questions mainly in Likert scale form [25] on their feelings and comments, to be completed by the users during and after the test. Next, we demonstrated how to interact with the MobilePaperAccess system in the process of learning. Besides demonstrating, we also guided the users by discussing with them. After practicing several times with the toy application, the users started to perform the RTMA. In this stage, we asked the participants to perform two tasks for each case. All the participants performed the tasks, respectively. They were instructed to check two different researchers' schedules and ask for an appointment with these two researchers as accurately and quickly as possible.

We employed a within-subjects design in this evaluation. The order of the four cases was counterbalanced with a 4×4 Latin square [13], while, inside, the order of two true tasks (T1 and T2) was counterbalanced with a 2×2 Latin square. The system only records automatically the log of performance with true tasks in four cases. For each case, each participant performed two tasks with 9 times of pointing for each task. This yielded 72 trials per participant ($2 \text{ tasks} \times 9 \text{ pointing trials} \times 4 \text{ cases} = 72 \text{ trials}$). Thus, the summary number was 864 trials ($12 \text{ subjects} \times 72 \text{ trials} = 864 \text{ trials}$).

We also set several variables for comparison. For the first group, the independent variables are input techniques (finger input, mask input, and page input) and interfaces (EDI and EII), which we grouped into four cases as stated above. The dependent variables were the interaction time sum, namely the sum of interaction time starting from the user's correct interaction to his/her correctly stopping each task in each case with the true application. For the second group, the independent variable was the interface layout, and the dependent variable was the interaction time of tasks T1 and T2. We also drew the access time and regarded it as the span from starting the application to the user's first interactive action in each case with the true application. We also recorded all the errors made in the true application of four cases, and marked the reasons for them. Each input step and the time cells were automatically logged by the system.

6.2 Main results

In terms of results, we obtained the interaction time sum, the interaction time of tasks T1 and T2, access time,

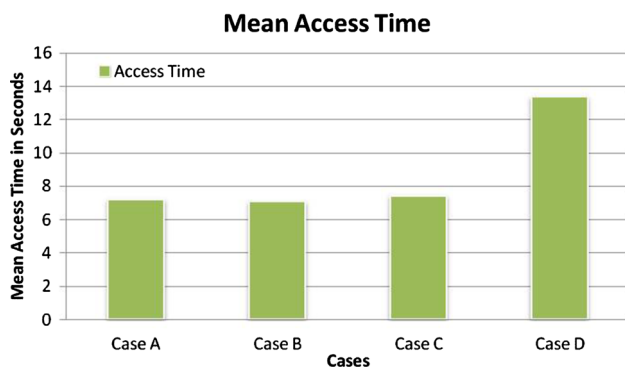


Fig. 17 The mean access time for each case

the same, with less than 8 s to access. Conversely, the access time in Case D was nearly twice as long as the other cases.

6.2.4 Interaction errors

Through observation and questionnaires, we found that the reasons of the errors are mainly due to user locomotion, misunderstandings of tasks, and attempts to do more than the tasks. These three errors were counted, respectively. Among these, the locomotion error is the interaction error. We counted the number of locomotion errors for each case of 12 participants. The number of interaction errors with EDI is less than with EII, and less with finger input than with page input as shown in Fig. 18.

6.2.5 User satisfaction on four cases

To obtain subjective opinions technically, we asked participants to respond to the Likert questionnaire items concerning easiness of learning for three input techniques with two interfaces. 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. Table 2 gives the average scores of four cases for the toy application and the true application. The scores showed that all participants thought it was not hard to learn and perform (the mean scores are all above 3).

6.2.6 User comments on four cases

Besides these scores, we also recorded users' comments. For Case A, i.e., the finger input with the EDI, four participants tired of lifting their arms after operating for a while, which led them to interact unsteadily with their fingers. Two participants said that the fixed position was efficient and convenient for the interaction. Furthermore, two participants commented on a physical chain reaction effect: When moving their arms and fingers, this resulted in

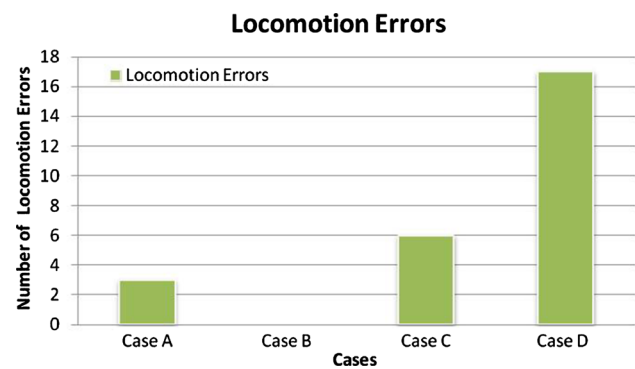


Fig. 18 The locomotion errors for each case

Table 2 Mean score of user satisfaction with the toy application and the true application in the four cases

Cases	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

a tiny movement of the camera fixed on their head. For Case B, one person said that the frame of the mask made it easy to choose and select items, while another user could not work properly with the frame's marker angle. Two people said their arm got tired. For Case C, more than half of the participants commented on the long time required for lifting their arms and unsteady fingers. They thought it was not easy to hold the interface steadily in their hands. Two participants experienced the chain reaction effect. For Case D, four participants found that the search for the right page to interact felt less easy when there were more pages in the booklet, and that returning to the index each time was not convenient. Only one participant mentioned feeling the chain reaction. One user preferred the marker interaction for a faster and more sensitive interactive experience. For the devices, six participants felt the screen was small to read, provoking a feeling of tiredness.

7 Discussion

The results of our exploring whether the three input techniques are easy to learn or not, is that that the finger, the mask and the page input techniques are all easy to learn. The average scores for easiness of learning and utilization with the three input techniques are all more than 3. However, not all learning has raised user satisfaction; after learning, the scores of true tasks in Cases A, B, and D are higher than toy learning scores. Besides, easiness of

learning varies slightly in the four cases. Out of these cases, interaction in Case B has the best score, which indicates it was the easiest and most convenient for learning compared with the others. From users' comments, we found that with EDI, more people reported a tired arm in Case A than in Case B. We thought that the mask stick played the role of an extended arm, leaving the arm in a more relaxed state, and reducing the effect of tiredness.

Furthermore, the answer to the second question stated above is that the sequence of performance from best to worst of the four cases is Case B, Case A, Case C, and Case D. Case B has the best overall performance, with the shortest interaction time, the shortest access time, no locomotion error, and the best satisfaction. Compared with Case A, B has fewer participants reporting a tired arm because the band with mask is more comfortable than lifting their hands. Case A performs better than Case C; they have virtually the same interaction time and access time except that A has a better score of satisfaction and fewer participants reporting a tired arm due to the fixed and stable interface. In turn, Case C performs better than Case D due to its shorter interaction time, shorter access time, fewer interaction errors, and better satisfaction score. Case D is most influenced by overall locomotion errors. From the users' comments, we found that the more pages there are, the harder the selection action is, even for the interaction time of tasks T1 and T2. The fact of searching for pages via a return to the index means that the input technique in Case D leads users to an unsteady interaction state. In a word, EDI performs better than EII, and the performance of input techniques from best to worst is mask, finger, and page. The best performance is awarded to the mask input technique with EDI.

This study also showed us the influence of Fitts's law on innovative wearable interfaces, which could answer our third question. From the ANOVA test, we found that the variable layout has no statistical significant influence on the interaction time of Cases A, B, and D. For Case D, the interface does not have the traditional layout, and

it is thus obvious that Fitts's law does not work on the interface in this case. In Cases A and B, the interaction time of T1 is shorter than that of T2 because pointing in T2 involves a longer distance than in T1. In Fig. 19, we can see that the blue points are related to task T1, while the red points are related to T2. The hand is usually located in the horizontal middle of the interface: It is quicker to reach the blue points than the red points (the transparent red dot cycle and the bottom-right red point illustrate the same distance as blue points). Besides, the variable layout has a statistical significant influence on the interaction time of Case C. Compared with EDI in Cases A and B, the locomotion amplifies the effect of Fitts's law with EII in Case C.

Finally, to reduce the locomotion errors and augment user experience in the wearable system with EDI and EII, we propose two solutions.

The first consists in increasing paper hardness and decreasing paper size. Users hold the paper with different degrees of strength that can result in its bending, thus reducing webcam recognition and leading to the same interaction problem as the locomotion errors. Paper hardness can compensate for this effect: We can choose cardboard as the paper interactive surface of the EII. Moreover, if we reduce paper size, the possibility of carelessly leaving part of the paper out of the webcam range will increase. 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, the number of interactive items in the paper-based interface also decreases.

To provide more interactive items and retain the link between information and physical indications, we propose another solution, namely the physical and digital mixed interface, which has been described in the continuum for EDI and EII in Sect. 4.2. With the aim of providing more information for the mixed interface and to add interactive items, we remove the configuration of the small-size display attached with goggle, and adopt the pico-projector as the output device. The projection display can be an alternative method for providing a larger visual presentation without any external device support. In this way, the mixed interface (see Fig. 20b) offers more dynamic interactive choices compared with the paper-based interface (see Fig. 20a). Since we also found that raising hands at eye level became tiring after a certain time and that the chain reaction reduced interaction efficiency, we propose changing the position of the webcam from the forehead to the chest to lower hand raising and ensure stability. We will fix the webcam and pico-projector together on the light cardboard support and choose the chest as the worn point for the mixed interface.

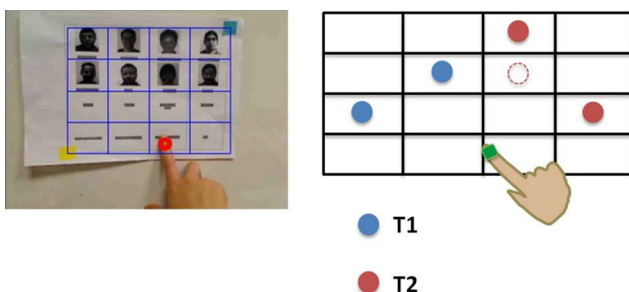
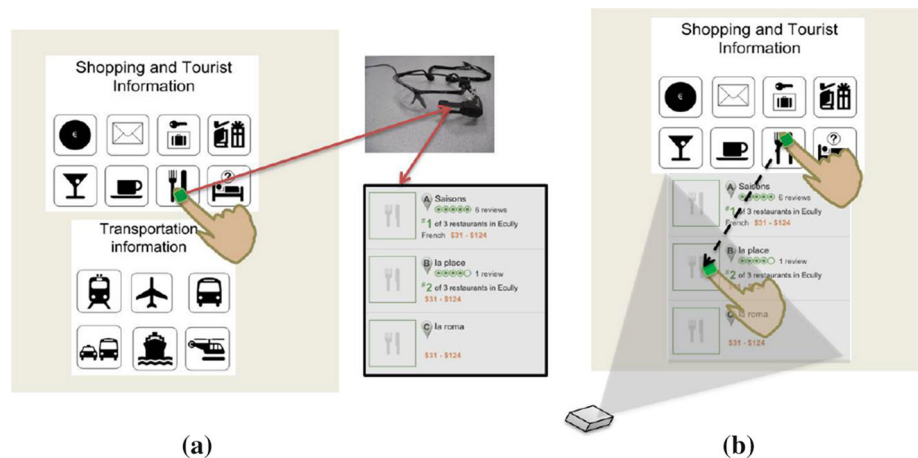


Fig. 19 The layout of RTMA

Fig. 20 From paper-based interface (a) to physical–digital mixed interface (b)



8 Conclusion and outlook

In this paper, we described our approach for exploring innovative user interfaces (IEI, EDI and environment independent interface), enabling the user to access in-environment information and environment independent information freely. We also explained the concepts of EDI and EII (environment independent interface), and our taxonomy of mobile user interfaces for EDI and EII. To realize EDI and EII, we proposed, designed, and implemented the Mobile-PaperAccess system, which is a wearable camera-glasses system, including the configuration: a webcam, a small screen attached to a goggle, and a laptop as the calculating device. Through this system, users can interact with the paper-based interface using finger, mask, and page input techniques. We organized an evaluation, and compared two interfaces (EDI and EII) and three input techniques (finger input, mask input, and page input). The quantitative and qualitative results showed the easiness of learning when interacting with EDI and EII, the performance of the three input techniques with two interfaces, and the influence of layout on interaction time with wearable interfaces.

For future work, we plan to investigate the physical and digital mixed interface with the camera-projector device unit containing the webcam, pico-projector, and a tablet, to perform the concepts of EDI and EII. Furthermore, more advanced input techniques of hand gestures such as the pinch gesture will be studied.

References

- 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, pp 49–56
- Asadzadeh P, Kulik L, Tanin E (2012) Gesture recognition using RFID technology. *Pers Ubiquit Comput* 16(3):225–234. doi: [10.1007/s00779-011-0395-z](https://doi.org/10.1007/s00779-011-0395-z)
- Ballagas R, Borchers J, Rohs M, Sheridan JG (2006) The smart phone: a ubiquitous input device. *IEEE Pervasive Comput* 5(1):70–77
- Bradski G (2000) The opencv library. *Dr Dobb's J* 25(11):120–126
- Bradski GR (1998) Computer vision face tracking for use in a perceptual user interface. In: WACV 1998: proceedings IEEE workshop on application of computer vision, pp 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, pp 449–456
- Choi J, Kim GJ (2013) Usability of one-handed interaction methods for handheld projection-based augmented reality. *Pers Ubiquit Comput* 17(2):399–409. doi: [10.1007/s00779-011-0502-1](https://doi.org/10.1007/s00779-011-0502-1)
- David BT, Chalon R (2007) IMERA: experimentation platform for computer augmented environment for mobile actors. In: Wi-MOB 2007: 3rd IEEE international conference on wireless and mobile computing, networking and communications, pp 51
- David BT, 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, Boca Raton, FL
- Dul J, Weerdmeester BA (2008) Ergonomics for beginners: a quick reference guide. CRC Press, Boca Raton, FL, pp 12
- Fishkin KP (2004) A taxonomy for and analysis of tangible interfaces. *Pers Ubiquit Comput* 8(5):347–358. doi: [10.1007/s00779-004-0297-4](https://doi.org/10.1007/s00779-004-0297-4)
- Fitzmaurice GW, Ishii H, Buxton WAS (1995) Bricks: laying the foundations for graspable user interfaces. In: CHI 1995: proceedings of the SIGCHI conference on human factors in computing systems, pp 442–449
- Grant DA (1948) The latin square principle in the design and analysis of psychological experiments. *Psychol Bull* 45(5):427
- Grimes GJ (1983) Digital data entry glove interface device. Patent US4414537
- Ha Y, Rolland J (2002) Optical assessment of head-mounted displays in visual space. *Appl Opt* 41(25):5282–5289
- Harrison C, Benko H, Wilson AD (2011) OmniTouch: wearable multitouch interaction everywhere. In: UIST 2011: proceedings of the 24th annual ACM symposium on user interface software and technology, pp 441–450
- Harrison C, Hudson SE (2010) Minput: enabling interaction on small mobile devices with high-precision, low-cost, multipoint

- optical tracking. In: Proceedings of the 28th international conference on human factors in computing systems, pp 1661–1664
18. 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, pp 453–462
 19. 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, pp 591–599
 20. Hornecker E, Psik T (2005) Using ARToolKit markers to build tangible prototypes and simulate other technologies. *Human-Computer Interaction-INTERACT 2005*, vol 3585, pp 30–42. doi: [10.1007/11555261_6](https://doi.org/10.1007/11555261_6)
 21. Ishii H (2008) Tangible bits: beyond pixels. In: TEI 2008: proceedings of the 2nd international conference on tangible and embedded interaction, pp xv–xxv
 22. 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, pp 234–241
 23. Kubicki S, Lepreux S, Kolski C (2012) RFID-driven situation awareness on TangiSense, a table interacting with tangible objects. *Pers Ubiquit Comput* 16(8):1079–1094. doi: [10.1007/s00779-011-0442-9](https://doi.org/10.1007/s00779-011-0442-9)
 24. 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, pp 361–370
 25. Likert R (1932) A technique for the measurement of attitudes. *Arch Psychol* 22(140):1–55
 26. Lyons K, Starner T, Gane B (2006) Experimental evaluations of the twiddler one-handed chording mobile keyboard. *Hum Comput Interact* 21(4):343–392
 27. Lyons K, Starner T, Plaisted D, Fusia J, Lyons A, Drew A, Looney EW (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, pp 671–678
 28. MacKenzie IS (1992) Fitts' law as a research and design tool in human-computer interaction. *Hum Comput Interact* 7(1):91–139
 29. Mistry P, Maes P (2008) Quickies: intelligent sticky notes. In: IET 2008: 4th international conference on intelligent environments, pp 1–4
 30. 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, pp 4111–4116
 31. Morris D (2010) Emerging input technologies for always-available mobile interaction. *Found Trends® Hum Comput Interact* 4(4):245–316
 32. Ni T, Baudisch P (2009) Disappearing mobile devices. In: UIST 2009: proceedings of the 22th annual ACM symposium on user interface software and technology, pp 101–110
 33. Rekimoto J, Ayatsuka Y (2000) CyberCode: designing augmented reality environments with visual tags. In: DARE 2000: proceedings of conference on designing augmented reality environments, pp 1–10
 34. Rouillard J (2008) Contextual QR codes. In: ICCGI'08: the third international multi-conference on computing in the global information technology, pp 50–55
 35. Rukzio E, Holleis P, Gellersen H (2012) Personal projectors for pervasive computing. *IEEE Pervasive Comput* 11(2):30–37. doi: [10.1109/MPRV.2011.17](https://doi.org/10.1109/MPRV.2011.17)
 36. Shaer O, Hornecker E (2010) Tangible user interfaces: past, present, and future directions. *Found Trends Hum Comput Interact* 3(1–2):1–137
 37. Spitzer MB, Rensing N, McClelland R, Aquilino P (1997) Eye-glass-based systems for wearable computing. In: Digest of papers. First international symposium on wearable computers, pp 48–51
 38. Starner T, Mann S, Rhodes B, Healey J, Russell KB, 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, vol 355
 39. 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, pp 4255–4260
 40. Vandam A (1997) Post-WIMP user interfaces. *Commun CM* 40(2):63–67
 41. Wang RY, Popović J (2009) Real-time hand-tracking with a color glove. In: SIGGRAPH 2009: ACM transactions on graphics (TOG), vol 28(3), Article No. 63
 42. Weiser M (1991) The computer for the 21st century. *Sci Am* 265(3):94–104
 43. Willis KD (2012) A pre-history of handheld projector-based interaction. *Pers Ubiquit Comput* 16(1):5–15. doi: [10.1007/s00779-011-0373-5](https://doi.org/10.1007/s00779-011-0373-5)
 44. Willis KDD, Poupyrev I, Shiratori T (2011) Motionbeam: a metaphor for character interaction with handheld projectors. In: CHI 2011: proceedings of the SIGCHI conference on human factors in computing systems, pp 1031–1040
 45. Wilson ML, Craggs D, Robinson S, Jones M, Brimble K (2012) Pico-ing into the future of mobile projection and contexts. *Pers Ubiquit Comput* 16(1):39–52. doi: [10.1007/s00779-011-0376-2](https://doi.org/10.1007/s00779-011-0376-2)
 46. 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, Springer, Berlin, pp 500–509