HANDHELD PROJECTOR INTERACTION

by

Xiang Cao

A thesis submitted in conformity with the requirements

for the degree of Doctor of Philosophy

Graduate Department of Computer Science

University of Toronto

Abstract

Handheld Projector Interaction

Xiang Cao

Doctor of Philosophy

Graduate Department of Computer Science

University of Toronto

2009

The recent trend towards miniaturization of projection technology indicates that handheld devices will soon have the ability to project information onto any surface, thus enabling interaction and applications that are not possible with current handheld devices. This opens up an emerging research area on interaction using handheld projectors. With the ability to project information, a handheld device can surmount the limitations of its small internal screen by creating a larger information display on an external surface. By doing so, the display and interaction space can be expanded to cover almost an entire physical environment. Large amounts of data can be displayed, a rich interaction vocabulary can be supported, and multiple co-located people can share the viewing experience at the same time.

In this thesis, I investigate research issues involved in the design, implementation, and user performance and behaviors regarding the usage of interactive handheld projectors. I create a handheld projector interaction prototype platform, and explore interaction concepts and techniques to support both single and multi-user interaction using one or several handheld projectors. I also empirically investigate the user behaviors related to handheld projector usage, in terms of both quantitative interaction performance with pointing tasks, and qualitative social behaviors that emerge from a game application.

This work is a multi-faceted investigation on handheld projector interaction, and will provide the groundwork for future research and development of interactive handheld projectors.

Acknowledgements

Over the years in the making of this thesis, I was indebted to numerous people who kindly helped and made my work and life possible.

During the course of my doctoral study, I had the precious opportunities to learn from many great researchers who served as my advisors, mentors, and collaborators. Firstly, I have been extremely fortunate as a student to work with my doctoral supervisor Ravin Balakrishnan, who always provided me with insightful comments, demonstrated how research should be done, and most importantly constantly challenged me to do better. My supervisory committee, Khai Truong and Karan Singh always made invaluable suggestions, asked the right questions over the years, and helped me to approach my work with different perspectives and methodologies. I can never thank them enough. I would also like to heartily thank Eyal de Lara, my latest committee member, for reading my thesis in a short notice and giving excellent feedback, and my external examiner Brad Myers for thoroughly going over my thesis and providing his insightful advice. I have also been greatly privileged to work as a research intern with Shumin Zhai at IBM Almaden Research Center, and Andy Wilson at Microsoft Research. They not only directed me on exciting research projects, but also educated me with their unique insight and experience outside the school environment. In addition, it was a highly rewarding experience to collaborate with many other outstanding researchers throughout the various research projects I worked on, such as Clifton Forlines, Michael Massimi, Ken Hinckley, and Scott Hudson, as well as other colleagues at University of Toronto, IBM and Microsoft. Last but not least, the co-supervisor during my Master's study Allan Jepson prepared me with valuable background in computer vision that helped me greatly with my thesis research. It has been a true privilege to work with these brilliant researchers and exceptionally kind people.

Being a member of the DGP lab in University of Toronto has always been an amazing experience. Every person in the lab has supported me greatly during my study and life. Only to name a few, I would like to first thank our lab administrator John Hancock, who professionally and patiently helped me throughout all my time in the lab. My research and study were not possible without his support. Gonzalo Ramos, Michael McGuffin, Shahzad Malik, Anastasia Bezerianos, Tovi Grossman, Daniel Wigdor, and Shengdong Zhao, who graduated before me and shared with me their valuable experience and advice. Abhishek Ranjan, who acted as a great officemate and accompanied me during working days and nights. Noah Lockwood, Tomer Moscovich, Gerry Chu, Michael Jurka, Ryan Schmidt, Xiaojun Bi, Jack Wang, and Seok-Hyung Bae who contributed their figures and voices in my research videos and pictures. Daniel Vogel and Joe Lazlo, who provided unique artistic and creative comments. And above all, the discussions, suggestions, and friendship from all the lab mates. My school life has been a great pleasure thanks to all of them.

I would also like to heartily thank my research assistants, Jacky Li, Serena Kao, Clarissa Mak, and Janek Lasocki-Biczysko, who invested their time and hard work and made substantial contributions to my research. It was a great learning experience on my part as well to have worked with these bright, creative, and responsible young students.

Finally, I cannot emphasize enough all the support and encouragement from my dear family and friends in Toronto, in China, and all over the world. I owe my sincere gratefulness to them, without whom I could not have made the progress I have had.

Refereed Publications from the Thesis Work

(content appears in this thesis with permission from ACM)

Xiang Cao, Ravin Balakrishnan. (2006). Interacting with dynamically defined information spaces using a handheld projector and a pen. *Proceedings of UIST 2006, ACM Symposium on User Interface Software and Technology*. p. 225-234. (http://doi.acm.org/10.1145/1166253.1166289)

Xiang Cao, Clifton Forlines, Ravin Balakrishnan. (2007). Multi-user interaction using handheld projectors. *Proceedings of UIST 2007, ACM Symposium on User Interface Software and Technology*. p. 43-52. (http://doi.acm.org/10.1145/1294211.1294220)

Xiang Cao, Jacky Jie Li, Ravin Balakrishnan. (2008). Peephole pointing: Modeling acquisition of dynamically revealed targets. *Proceedings of CHI 2008, ACM Conference on Human Factors in Computing Systems*. p. 1699-1708. CHI 2008 Best Paper Honorable Mention. (http://doi.acm.org/10.1145/1357054.1357320)

Xiang Cao, Michael Massimi, Ravin Balakrishnan. (2008). *Flashlight Jigsaw*: An Exploratory Study of an Ad-Hoc Multi-Player Game on Public Displays. *Proceedings of CSCW 2008, ACM Conference on Computer Supported Cooperative Work*. p. 77-86. (http://doi.acm.org/10.1145/1460563.1460577)

.

Contents

List of Figures

Chapter 1

Introduction

Handheld devices are becoming increasingly ubiquitous in our daily life: PDAs and cell phones help us manage personal information, access information resources, communicate with others, and entertain ourselves almost anywhere. Their rapidly improving computation, storage, and communication capabilities enable many tasks and applications traditionally reserved for desktop computers to shift onto these handheld devices, which can be easily carried and used in various environments. However, the small form factor that makes them so appealing is also a significant limitation in that the resulting small sizes of the embedded screens make it difficult to display large amounts of information, enable rich interactions, or support multiple co-located users viewing the information. As more and more advanced activities emerge on handheld devices, the shortage of display and interaction space also becomes more remarkable.

A possible solution to this small screen limitation of handheld devices may lie in recent advances in projection technology, which have seen projectors become smaller, lighter, cheaper, and require less power (Figure 1.1). Given this trend, it is reasonable to expect that within a few years projectors would be carried in a pocket or embedded in other mobile devices such as cell phones and PDAs. With the ability to project information, a handheld device can surmount the limitations of its small internal screen by creating a larger information display on an external surface (Figure 1.2). By doing so, the display and interaction space can be expanded to cover almost an entire physical environment. Large amounts of data can be displayed, a rich interaction vocabulary can be supported, and multiple co-located people can share the viewing experience at the same time. In addition, the ability to project information in the physical environment may create an experience that blends the virtual and the physical worlds. Therefore, handheld projectors could more easily support advanced activities that are difficult on traditional handheld devices, as well as creating new applications beyond those on current handhelds and desktops.

Figure 1.1: Miniaturized projectors.

(images from www.aboutprojectors.com, www.symbol.com, www.electronics-lab.com)

Figure 1.2: Envisioned handheld projector usage. (image from www.microvision.com)

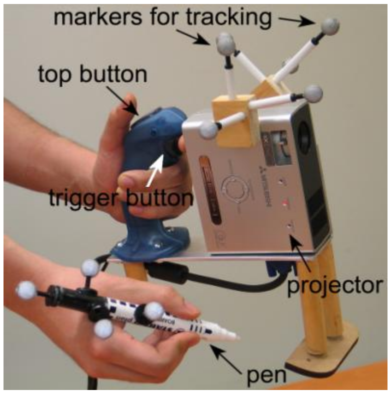
As a promising new technology that has the potential to become prevalent in the near future, handheld projectors open up an emerging research area on how people may perform interactions using them. Given the new input/output affordances and the possible mobile usage scenarios enabled by handheld projectors, traditional techniques designed for desktop computer interaction are unlikely to suit them. New paradigms and techniques need to be designed and evaluated to support handheld projector interaction, both for a single user and for multiple users located in the same physical environment. On the other hand, it is important to investigate how the usage of handheld projectors will affect users' ability and behavior while interacting with mobile devices and communicating with other people in a social setting. These findings would guide the design and development of future handheld projector applications.

However, having only started to attract researchers in recent years, handheld projector interaction has very much remained an uninvestigated area previously. This thesis aims at an exploration of research issues involved in the design, implementation, and user performance and behaviors regarding interactive handheld projector usage. To achieve this goal, I employed a multi-faceted research methodology that integrates theoretical analysis, creative design, technical development, and both quantitative and qualitative user studies. My current work is a multi-faceted investigation on handheld projector interaction, and will provide the groundwork for future research and development of interactive handheld projectors.

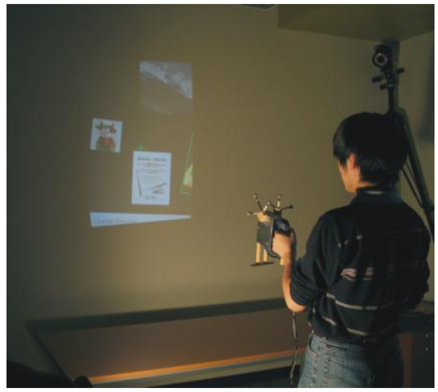
1.1 Outline

In Chapter 2, I first review previous relevant research that helps guide my exploration on handheld projector interaction. I start my research with the design and implementation of a general handheld projector interaction prototype system used as my research platform (Figure 1.3), reported in Chapter 3. By changing and warping the projected image according to the projector's movement tracked by a Vicon (www.vicon.com) motion capture system, a "Flashlight" metaphor is created, in which the projected image reveals a portion of a larger virtual workspace that is stationary on the projection surface.

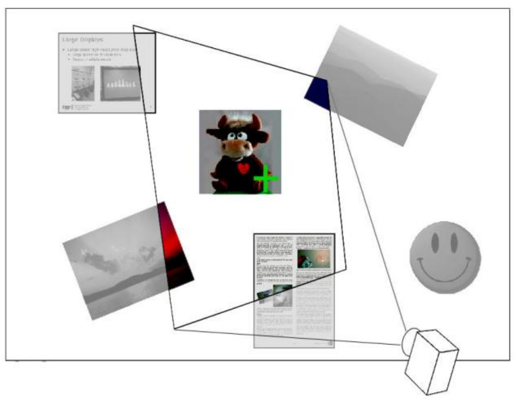
The user can use the cursor in the center of the projection image and/or a passive pen to provide input to the system. I also develop techniques for the user to interactively define workspaces in a physical environment. The platform is general and flexible enough to experiment a large variety of interaction designs and applications, and all research in this thesis is based on it.



(a) Handheld projector.



(b) System in use.



(c) Flashlight metaphor.

Figure 1.3: Handheld projector prototype platform.

The basic case of handheld projector interaction is when a single projector is used for interaction. In Chapter 4, I develop a set of interaction techniques to support interaction using a single handheld projector. These techniques cover general operations useful for various handheld projector applications (*e.g.*, object manipulation and parameter adjustment), and can be combined to establish higher-level dialogs that are suited to the context. These designs take account of the special characteristics and affordances of handheld projectors (such as the dynamic image resolution determined by the projection distance), and provide a generic interaction vocabulary that could be widely applied by future researchers and designers. I also explore pen-based and bimanual interaction along with a handheld projector. Finally, I investigate several usage scenarios that can be supported by using a single handheld projector at a time, including single-person usage, synchronous and asynchronous multi-person usage, and game usage.

Given handheld projectors' unique affordances and interaction styles, it is important to quantitatively investigate users' performance with generic interaction tasks using them. As a first step in this direction, in Chapter 5 I experimentally study target pointing tasks under the Flashlight metaphor, in which the workspace is partially and dynamically revealed by a moving display window (Flashlight). The Flashlight metaphor is employed throughout my handheld projector interaction design and plays a central role in the interaction style. In order to focus my investigation on the influence of the metaphor itself rather than the performance of the prototype system, the experiment is conducted on a desktop computer simulating the flashlight metaphor with abstract 1D pointing tasks. Results reveal that pointing time is inversely correlated with the size of the display window, in addition to being affected by target size and distance. I propose and validate a new quantitative model to predict targeting pointing time, extending Fitts' law [39] to incorporate the display window size. This study has important implications to the design of handheld projector interfaces, as well as to other interfaces that utilize the Flashlight metaphor, such as Peephole interaction [130] on spatially aware displays.

With the vision of handheld projectors embedded in every mobile device, it naturally leads to how they interact with each other when co-located. In Chapter 6, I explore the design space of multi-user interaction using multiple handheld projectors. I expand the single-user interaction designs in Chapter 6 to support multiple users working in a shared physical space, each using their own handheld projector. A set of interaction concepts and techniques are developed to specifically support interaction between users such as ownership & access control, file exchange, and linkage between objects. Special emphasis is on ad hoc composition of multiple projections, opening up new possibilities such as the blending of multiple personalized views. I also investigate techniques and concepts to support privacy and independent work, which is crucial when handheld projectors are to be used in public spaces. With these designs, daily interpersonal communication tasks such as exchanging contacts and scheduling meetings can be largely simplified. This work unleashes the power of handheld projectors to support co-located interpersonal interaction, in contrast to traditional handheld devices which are almost exclusively single-user devices.

Handheld projectors create a considerably different experience comparing with current handheld devices. In particular, the multi-user characteristics described in Chapter 6 and the semi-public displays created by projectors are likely to influence people's social behaviors around mobile devices. In Chapter 7, I investigate social interaction patterns emerging from handheld projector usage with a specific case study – public games. Public social games played on mobile devices or public displays have become popular in recent years both as entertainment and as a research topic. I design and deploy a lightweight ad-hoc multi-player game, *Flashlight Jigsaw*, which leverages the affordances of handheld projectors. Such a game might encourage co-located multi-player gaming in public spaces, enable tighter relationships between the game and the physical world, and offer an interesting experience for spectators. A simulated setup of the game was deployed in a shared lab space and a public atrium for two weeks in total. Through interviews supported by observations and system logs I explored the experiences and behaviors of players and spectators. I also investigated the interrelationship between the game and the spaces it is deployed in. The research resulted in findings regarding game play, communication, social interaction, spectatorship, and space and location around such a game. I use these findings to develop design implications for future handheld projector games, which could provide insights for other social applications of handheld projectors, and also public social games in general.

In summary, this thesis explores several aspects of handheld projector interaction within a systematic research framework. I contribute the groundwork for handheld projector interaction research including technical development, interaction design, user performance modeling, as well as empirical study and evaluation of such a novel interaction technology. A prototype system provides the platform for exploring various interactions and applications for handheld projectors. For both single-user and multi-user usages, generic interaction techniques are created as the basic vocabulary for handheld projector interfaces suited to a wide range of applications; and higher-level usage scenarios are explored, which demonstrated how the affordances of handheld projectors might support better user experiences with various activities. The quantitative model for pointing tasks under the Flashlight interaction metaphor supplies a deeper understanding of user performance using handheld projectors as a new interaction style, and the deployment and qualitative study of handheld-projector-supported public games opens up insights on how such a novel interaction technology may affect social behaviors between people in real life. In Chapter 8, I provide a more detailed summarization of these contributions, and discuss limitations of the work as well as directions for future exploration.

Chapter 2

Related Work

In this chapter, we review previous relevant research that helps guide our exploration on handheld projector interaction. The chapter is structured as follows: In Section 2.1, we first present an overview of various technologies that enable miniaturization of projectors, followed by image processing techniques to compensate for projector movement. These advances make the development of interactive handheld projectors possible. Section 2.2 introduces interaction metaphors and input techniques previously explored for handheld projectors. Section 2.3 looks at applications and usage scenarios experimented or envisioned by previous researchers. In Section 2.4, we summarize other related research areas that have implications on our research, such as handheld device interaction, projector-based interaction, augmented reality, and public social games.

2.1 Supporting Technologies

In this section, we first provide a background of the projection technologies that enable miniaturized projection devices. Then we discuss the approaches used to correct the projected image when a handheld projector is moved by the user.

2.1.1 Projection Miniaturization

Three technical approaches have been employed to miniaturize projection devices: utilizing LED light sources with conventional projection technology; steering laser beams to produce the projected image pixels sequentially; and generating laser projection by diffraction through hologram patterns.

LED Light Source Projector

The most mature technology up to date that can reduce the size of projection devices is by using Light-Emitting Diodes (LED) as the light source to replace traditional lamps (typically fluorescent bulbs) used by current mainstream projectors. The recent rapid improvement in LED performance has made LED-based illumination possible [56]. The small size of LEDs immediately enables compact projector designs that were not possible with traditional light sources. In addition, LED light sources provide advantages such as immediate projection readiness, long running time, and swift switch-off. Examples of current products employing this technology include the Mitsubishi Pocket Projector and the Samsung Mini Projector (Figure 2.1). These projectors typically have a size that can fit in people’s palm, and weight around one pound. Kanayama *et al.* [66] also explored using a specially designed collimator to improve the brightness of LED projectors.

(a) Mitsubishi Pocket Projector. (b) Samsung Mini-Projector.

Figure 2.1: LED light source projectors.

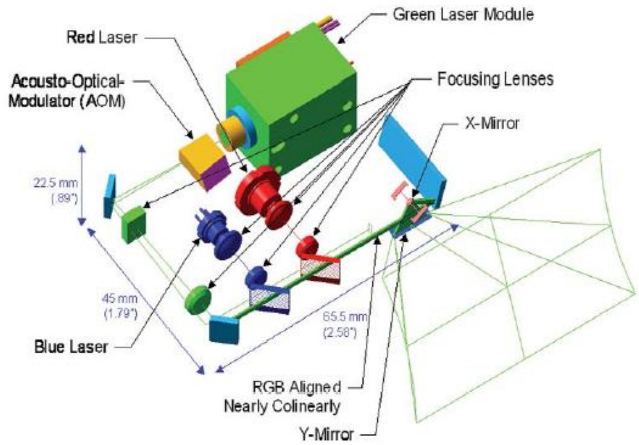
(images from www.masternewmedia.org, uk.gizmodo.com)

Although the light source component can be effectively downsized, conventional projectors that typically utilize Digital Light Processing (DLP) technology include another central optical component - a micro-mirror array that comprises of typically millions of mirrors. By rapid repositioning of the mirrors, different intensities for each color channel are created. The large number of moving parts involved in the mirror array makes further miniaturization difficult. Keeping the projection in focus is also a challenge when the projector is constantly moving as in the handheld scenario.

Laser-Steering Projector

Instead of generating all the pixels at one time as in DLP projectors, people have also explored steering a single laser beam (or a one-dimensional laser beam array) to sequentially generate the pixels one by one (or line by line). This technology is analogous to the function of the Cathode Ray Tube (CRT) technology used in desktop monitors, replacing the CRT electron beams by visible laser beams. The number of moving parts needed to steer the laser beam is minimal compared with DLP micro-mirror arrays, therefore the optical components of the projector can be significantly downsized. Because of the high convergence of laser beams, the projected image stays sharp regardless of the projection distance. This technology is becoming increasingly popular in the recent years.

Based on this technology, the Laser Projection Display (LPD) prototype from Symbol Technologies [112] uses two orthogonally moving mirrors to steer the laser beam. The size of the LPD engine is approximately 6.6cm x 4.6cm x 2.3 cm. Similarly, the 7mm-thick Microvision SHOW handheld projector (www.microvision.com) is capable of producing a 100-inch image at DVD resolution. (Figure 2.2) Other research in this aspect includes Hernandez [53], who used a DC motor to steer a one-dimensional laser beam array; Zalevsky *et al.* [134], who used spatial light modulation (SLM) to steer the laser beam through amplitude modulation.



(a) Optical Layout.

(b) Symbol LPD Prototype (right).

(c) Microvision SHOW handheld projector.

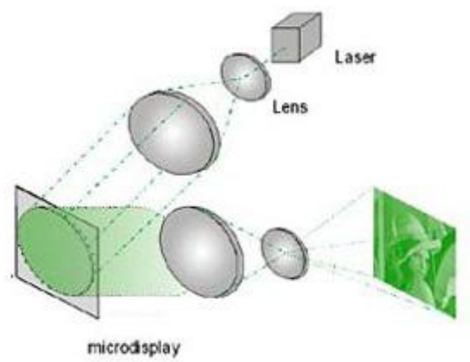
Figure 2.2: Laser-Steering Projectors.

(images from www.symbol.com, www.itechnews.net)

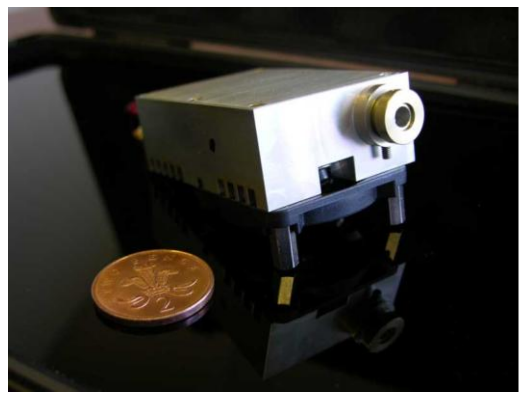
Hologram Laser Projector

An alternative way of generating projection images with laser is through hologram. The hologram laser projector uses laser light sources in conjunction with a phase modulating micro-display on which a hologram pattern, rather than the desired image, is displayed. The patterns are calculated such that, when the micro-display is illuminated by coherent laser light, the diffraction of the laser results in the formation of the image (Figure 2.3a). Since the image is formed entirely by diffraction, the system does not involve any moving parts such as the moving mirrors used in DLP or laser-steering projectors. Therefore its size can be even further reduced.

The PVPro diffractive technology (www.lightblueoptics.com) developed by Light Blue Optics enables a single-color hologram laser projector prototype with the size of a matchbox, and can be readily embedded into mobile devices (Figure 2.3b). One of the key challenges to produce a hologram laser projector is the real-time calculation of the hologram pattern to be displayed. See Cable *et al.* [26] for an approach to reduce the computational complexity of the hologram generation algorithm.



(a) Projection mechanism.



(b) PVPro prototype.

Figure 2.3: Hologram laser projector. (images from www.lightblueoptics.com)

2.1.2 Image Correction

With the above projection technologies available, it is possible to make handheld projection devices that can create large display spaces virtually anywhere. However, holding and moving the projector by hand also means the relative angle and position between the projector and the projection surface are constantly changing, which will result in unfavorable keystone distortion. On the other hand, it is impossible to use the projector itself to point to something in the projected content, since the image is simultaneously moving with the projector. Therefore the first key issue is to produce a stabilized and distortion-free projection image that is stationary relative to the projection surface, regardless of the projector's movement.

Sukthankar *et al.* [89, 111] present the first attempt in using computer vision technology to automatically correct the keystone distortion of a static projector. The key idea is to pre-warp the image sent to the projector in such a way that the distortions induced by the arbitrary projector-screen geometry are precisely negated. The perspective transform between two arbitrary planes (in this case, the projector image plane and the projection screen plane) can be expressed by the equations below (in homogeneous coordinates) [45]:

$$\begin{pmatrix} xw \\ yw \\ w \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 \\ p_7 & p_8 & p_9 \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = \mathbf{H} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

by which point(X, Y) on the projector plane is mapped to point (x, y) on the screen plane (w is an arbitrary multiplier). The transformation matrix **H** (subject to an arbitrary scale factor), called the homography, uniquely defines the transform between the two planes. The homography matrix **H** can be determined by as few as four pairs of point correspondence. By pre-warping the projection image by $\mathbf{H}^{-1}$, the perspective distortion can be effectively negated (Figure 2.4).

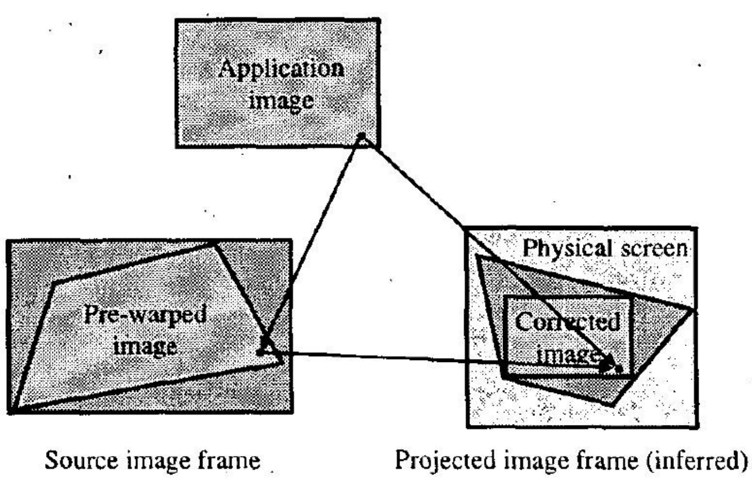


Figure 2.4: Image correction using projector-screen homography. (image from [89])

Although this work was aimed at correcting distortion for a static projector in the application of projected presentation, the spirit of pre-warping the projection image by the projector-screen homography remains the same in all subsequent works (such as [92]), including those which correct the image dynamically for handheld projectors [13]. Depending on the application scenarios and supporting technologies, researchers employed different approaches to extract this homography.

One way to recover the projector-screen homography is to use a camera to automatically detect the projected image frame and the projection surface frame, using image processing techniques. The homography can then be calculated from the correspondences between the vertices of the two frames. Sukthankar *et al.* [111] used a static camera to detect the boundary of the projection screen, which is assumed to be an uniformly lit object with visible edges. *Beardsley et al.* [13] attached a camera to their handheld projector prototype, and detected visual markers attached on the projection surface. This approach relies on either special visual characteristics of the projection surface, or attaching visual markers to the surface, therefore cannot easily support ubiquitous use in arbitrary environments.

Raskar *et al.* [93] also explored embedding wireless tags in the environment and in physical objects, which transmit both their identities and geometries to a handheld location-aware projector through radio frequency signals. The homography can then be computed from the geometry of the tags. This approach provides high-precision image correction, and can also handle projecting on objects with complex geometries (non-planar or multi-planar). However, the dependence on pre-embedded tags again precludes it from ubiquitous use.

If the projector has been fully calibrated, the projector-screen homography can also be determined from tracking the position and orientation of the projector relative to the projection surface. This approach does not require pre-implementing the environment with markers or tags, therefore might be more suited to ubiquitous use. For example, Rapp *et al.* [90] embedded orientation sensors in their handheld projector to recover the transformation and correct the image accordingly. With the rapid improvement of portable and lightweight location sensing technologies such as indoor GPS [117] or TrackSense (computer vision approach based on structured light patterns) [68], it is reasonable to anticipate this tracking-based approach will become ubiquitously deployable in the near future.

Other applications of these image correction mechanisms include supporting projection on curved or multi-planar surfaces [4, 6, 100], using multiple projectors to create a seamless display on arbitrary surfaces [6, 91], and eliminating shadows or blinding lights for the speaker in projected presentations [28, 113].

In summary, although the projector miniaturization and image correction technologies are not technically mature yet, they have provided sufficient support for developing a research prototype, with which we can start to explore and experiment with handheld projector interaction techniques and scenarios.

2.2 Interaction Metaphor and Input Techniques

In this section, we introduce some interaction metaphors and input techniques that have been previously explored for interactive handheld projectors. Because handheld projector interaction research is still in the early stage, these metaphors and techniques are all of exploratory nature, and leave much space for more systematic investigations.

2.2.1 Interaction Metaphors

Floating Window

The simplest metaphor is to use the handheld projector to directly project the content typically displayed on the embedded screen on handheld devices, without any attempt to compensate the projector movement. The user sees a floating window wherever the projector is pointing at, which is essentially an enlarged version of the traditional embedded screen. The Hotaru system [109] (Figure 2.5) implemented this metaphor by using a static mounted projector to simulate displays projected from PDAs tracked by a stereo camera. The system also allows users to annotate, rotate and transfer files by touching the projected displays with fingers.

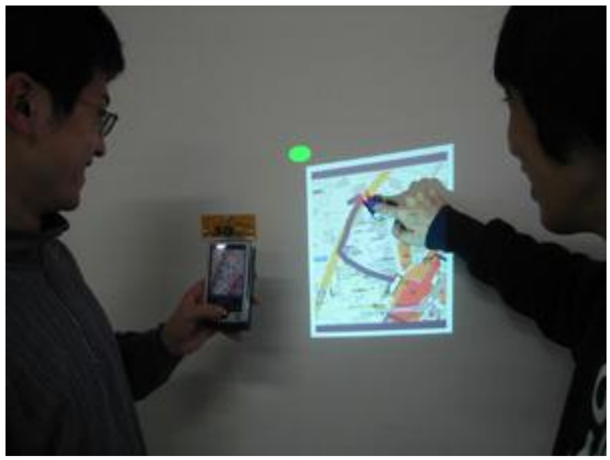


Figure 2.5: Floating window metaphor (Hotaru system). (image from [109])

The floating window metaphor is independent of the physical surface to be projected on, therefore lightweight to implement. Conventional handheld applications can be directly projected to create a better viewing experience. However, the effective interaction space is still limited by the projector's resolution. On the other hand, the fact that the floating window is clutched with the projector makes it impossible to use the projector itself as a pointer for interaction. Therefore, the system has to depend on additional input for pointing operations, either by direct touch on the projection [109], difficult to control when the projection is constantly moving; or using a pointing device (e.g. stylus) on the handheld device, which divides the user's attention from the display.

Projected Desktop

Leveraging the image correction techniques introduced in Section 2.1.2, the handheld projector can also be used to display a stable and distortion-free virtual desktop inscribed within the projection image. Conventional desktop applications can then be directly migrated onto the projected desktop. By displaying a pointer at the center of the projector's image plane, movement of the projector can directly control the pointer's movement across the stabilized desktop image. With direct pointer control and a button on the projector, all standard mouse interactions in a WIMP interface are possible. Figure 2.6 shows the projected desktop metaphor illustrated in [13, 93], in which the user uses the projected cursor to perform mouse operations with a web browser.

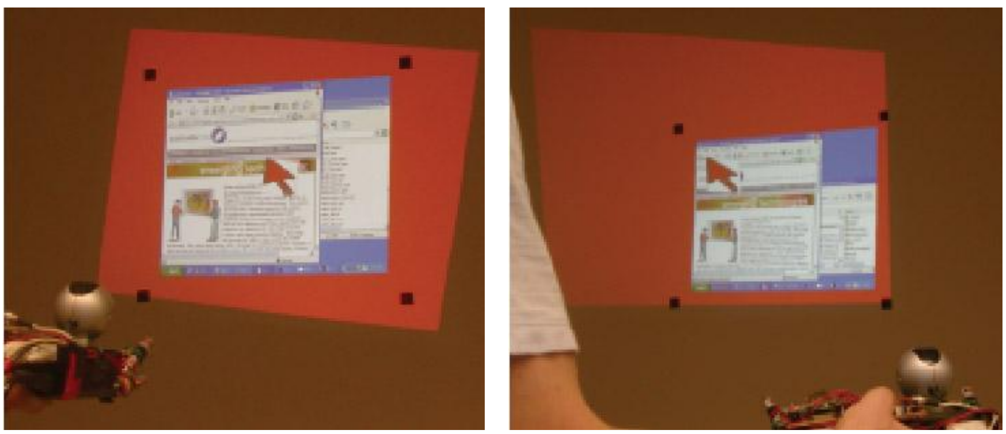


Figure 2.6: Projected desktop metaphor. (images from [13])

Since this interaction metaphor is directly compatible with desktop GUI, the learning efforts required are minimal to users. However, in order to render the entire desktop within the projected image, lots of image pixels are wasted, as shown by the red regions in Figure 2.6. This pixel wastage is aggravated if we require the pointer (displayed at the center of the projection image) to be able to traverse the entire desktop, which means the desktop cannot exceed a quarter of the projected image. This compromises the merit of the large display space created by the handheld projector. Forlines et al. [44] alleviated this problem (and the jittery input problem) by designing the Zoom-and-Pick widget, which supports fluid local zooming. On the other hand, the projected desktop metaphor restricts the projector's movement within a small space and angle, therefore negates the desirable affordances provided by the handheld projector as a mobile device that can be freely moved in 3D space.

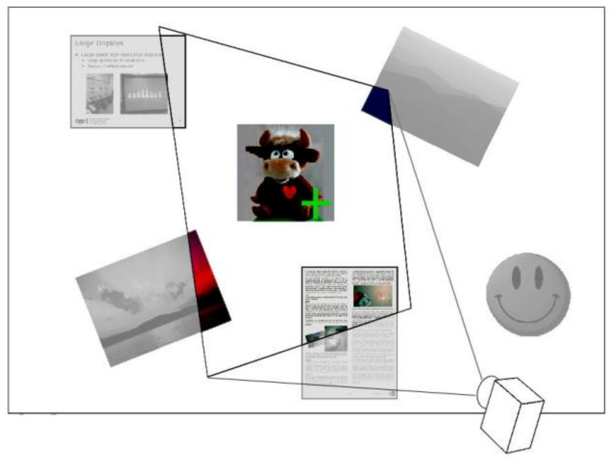


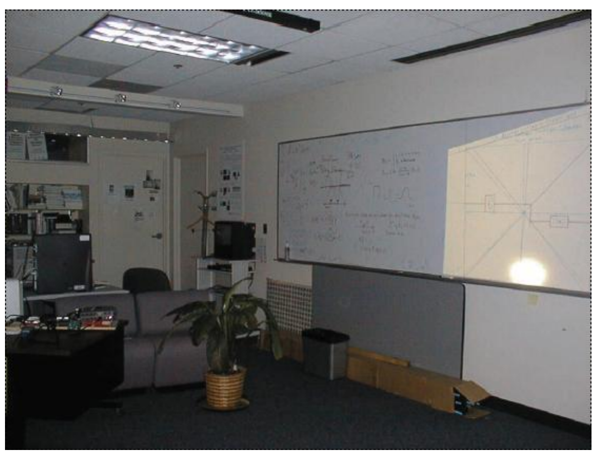
Figure 2.7: Flashlight metaphor.

Flashlight

Going one step further from the projected desktop metaphor, the flashlight (also known as spotlight) metaphor expands the interaction space of the handheld projector to the entire physical environment. The projected image reveals a portion of a large virtual workspace that is situated on the projection surface (Figure 2.7). When the projector is moved, the projected image content changes accordingly to reflect the change of the projected region relative to the workspace. By doing so, an illusion is created as if the user is exploring a stationary workspace attached to the physical surface. Especially, if multiple workspaces are associated with multiple physical surfaces, the result is a user experience of looking around in a dark environment with a flashlight. This metaphor is conceptually similar to the various researches on peephole displays [41, 61, 116, 130] work, which support both browsing and interacting with a large virtual workspace using spatially aware displays. Directly projecting on physical surfaces results in a tighter coupling between the workspace and the physical environment, in contrast to displaying on handheld screens used in peephole display research.

Teller et al. [115] first implemented a *software flashlight* prototype as one example application of their pose-aware devices. They used a portable projection system augmented with position listeners to overlay meta-information onto surfaces or objects of interest (Figure 2.8). The system was developed purely for viewing information, and did not provide a way for user interaction.

(a) Prototype.



(b) Information overlay.

Figure 2.8: Software flashlight. (images from [115])

Rapp et al. [90] present *Spotlight Navigation*, using a handheld projector embedded with orientation sensors to produce a circular-shaped spotlight to explore and interact with projected data (Figure 2.9). Simple interactions such as drag-and-drop and semantic zooming are supported. The user can also use the projected cursor to draw virtual inks on the wall. They developed a tremor cancellation module to smooth the inks drawn by a jittery hand.

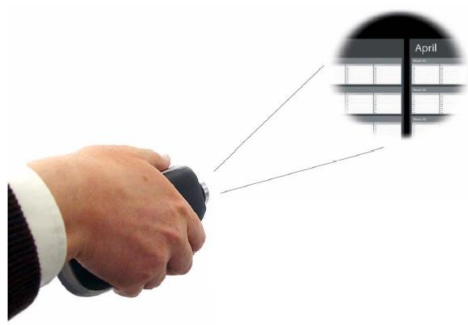


Figure 2.9: Spotlight Navigation. (image from [90])

Blaskó et al. [20] used a statically mounted projector to simulate a wrist-worn projection display that supports a *spotlight technique* to reveal different portions of information by wrist movements (Figure 2.10).

Figure 2.10: Spotlight Technique. (image from [20])

The flashlight metaphor makes full use of the projector's display capability, and creates much larger display spaces than what is possible with the projected desktop metaphor. On the other hand, it breaks the boundary between the virtual information and the physical environment, and results in a new design space. However, because the user only sees a portion of the workspace at one time, it brings up the possibility that the interaction context is lost at some point. It also raises challenges for the users to efficiently navigate through the workspaces and interact with currently off-screen content.

My exploration in this thesis is based on the Flashlight metaphor, which is extended to involve the entire physical environment. I also create interaction techniques to alleviate the disadvantages of the Flashlight metaphor, and experimentally investigate its influence on user performance.

2.2.2 Input Techniques

Different input techniques have been explored to interact with handheld projectors. These techniques have different pros and cons, and are suited to different applications and scenarios.

Pointing using the Projector

For the Projected Desktop and the Flashlight metaphor, the most straightforward technique is to emulate mouse operations (mostly pointing and clicking) utilizing the cursor displayed in the projected image center and a pushbutton on the projector. Beardsley [13] and Raskar et al. [93] used the point-and-click input to operate standard desktop applications such as a web browser (Figure 2.6). Rapp et al. [90] also support basic drag-and-drop operations using point-and-click in their system.

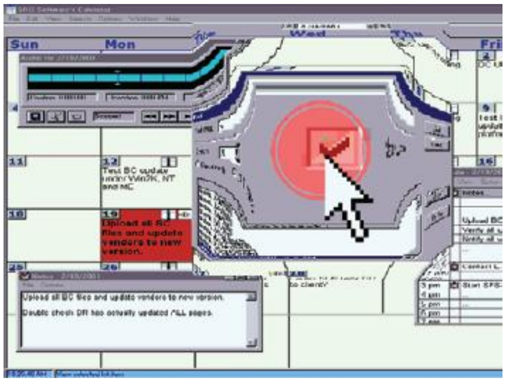


Figure 2.11: Zoom-and-Pick widget. (image from [44])

These pointing-based techniques are directly compatible with traditional mouse input, therefore require least learning effort for users. However, precise freehand pointing is a difficult task due to hand jitter, as reported by Myers et al. [81] in their empirical study comparing a set of freehand pointing devices, and as observed in many laser pointer input systems [75, 82]. To address this problem (and the problem of pixel wastage as discussed in section 2.2.1), Forlines et al. [44] designed an interactive widget called Zoom-and-Pick (Figure 2.11). By introducing a “dead zone” within which the projector movement does not affect the cursor position, and allowing the user to zoom in on areas of interest, accurate target selections can be made.

Direct Touch

Another way to interact with the projected information is by direct touching on the projection surface, as if on a touch-screen. The Hotaru system [109] tracks the user’s finger by attaching a LED marker to it. The user can use the finger on top of the projection surface to annotate documents, rotate objects, and transfer files between devices (Figure 2.12).

Figure 2.12: Finger touch input on projection surface. (images from [109])

Direct touch input can achieve a much higher precision than what is possible with freehand pointing. It also enables operations that are not necessarily inside the projection region. Finally it offers a tighter connection between the interface and the physical world. However given the constraint of reachability, direct touch is only applicable to small-scale local operations on surfaces that are near the user. In the floating window metaphor as used in [109], it also causes confusions about the reference frame when the window is moving.

On-Device Input

Researchers have also explored augmenting the handheld projection devices with additional input channels to provide more interaction possibilities. Blaskó et al [20] embedded their simulated wrist-worn projection display with a touch-sensitive pad to support a set of cursorless widgets [21] for functions such as zooming and panning. Rapp et al. [90] used a wheel on the projector to perform zooming operations. Other possibilities include adding a joystick to control the cursor, or using the stylus input currently available on PDAs. When augmenting the projector with additional input channel, special attention needs to be paid not to distract the user's attention from the projected information to the device itself.

My exploration in this thesis combines pointing using the projector and direct touch using a passive pen, and creates a rich interaction vocabulary based on them for both single- and multi-user scenarios.

2.3 Applications and Scenarios

As a device that is both mobile and sharable, the handheld projector can support applications and scenarios that are difficult or impossible to achieve using traditional mobile or situated devices. In this section we discuss some applications and scenarios that have been experimented or envisioned by researchers. Most of these works are in the stage of proof-of-concept prototypes rather than functional systems.

2.3.1 Personal Information Processing

The most straightforward application scenario is to use the handheld projector as an extension of traditional PDAs to view and process personal information. The user can temporarily convert the physical environment around into workspaces to manage personal information almost anywhere. The *Spotlight Navigation* prototype [90] supported basic personal applications such as a contact book, a calendar and a notebook. Blaskó et al. [20] implemented a stock information browser application as a test-bed for their wrist-worn projection interaction techniques. The Hotaru system [109] supports viewing and annotation of digital maps (Figure 2.12).

2.3.2 Interacting with the Physical World

The ability to project information in arbitrary physical spaces blurs the boundary between virtual information and the physical world. The handheld projector provides new possibilities for the user to interact with the physical environment around.

The projector can be used to overlay auxiliary information on physical objects or regions, and create a user experience similar to that created by augmented reality (AR) systems [5, 8, 14, 25]. However, compared with traditional AR systems that display information on see-through head-mounted displays (HMD) [57, 122, 131] or handheld displays [83, 121] that are designed for single-person experience, the "projected augmented reality" [13] supported by handheld projectors display information that can be viewed by multiple people, therefore can better support collaborative applications. On the other hand, seeing overlaid information directly on the physical surfaces instead of through display instruments results in a more natural experience, and avoids attention division for the user. Finally, utilizing direct touch input techniques discussed in section 2.2.2, the user might directly add and edit the overlaid information (e.g. annotations of physical objects) on the physical surfaces, instead of relying on indirect techniques such as in Patel et al. [83].

Teller et al. [115] used their *software flashlight* to overlay metadata in the form of text or geometric information (Figure 2.8) in the environment. They described the scenario that construction workers use the projected information to depict hidden infrastructure (e.g. electric mains or plumbing inside walls), access instructions, or display diagrams of planned construction. Beardsley et al. [13] also discussed applications such as projecting a map of the environment served by a fuse box onto the fuse box itself, and letting the user to select a location on the projected map to highlight the appropriate fuse that serves the location (Figure 2.13). Raskar et al. [93] depicted a warehouse scenario, in which the user locates items about to expire (red circles) and annotates some of them (larger white circles). These annotations can also be accessed later by other users (Figure 2.14).

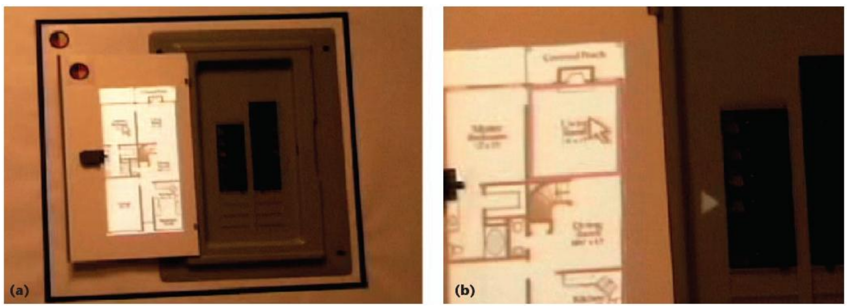


Figure 2.13: Augmentation of a fuse box. (images from [13])

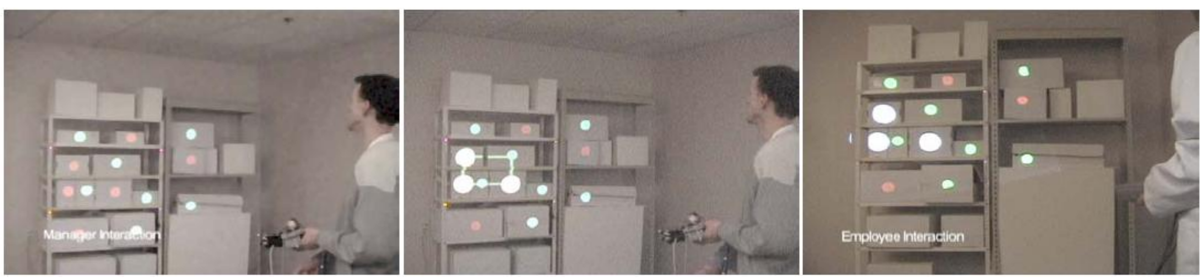


Figure 2.14: Warehouse scenario. (images from [93])

In addition to outputting information into the physical environment, researchers have also explored taking information from the environment as input.

Beardsley et al. [13] present using the projected cursor to select a physical region of interest similar to the hold-and-drag region selection operation using a mouse in desktop applications (Figure 2.15). The selected region can then be used as input to the system. For example, an attached camera can capture the visual appearance of the region and perform computer vision processing such as object recognition or text translation.

Raskar et al. [93] used a camera attached to the handheld projector to copy image texture from one physical surface and paste it onto another surface (Figure 2.16). Although they developed this technique merely to work around the technical limitation that their system cannot project on arbitrary-shaped surfaces, this conceptually provides the possibility to seamlessly transform information between the physical and the virtual form, and transfer physical information between places. For example, the user may take a snapshot of a physical poster, and later on post a virtual version of it in another place.

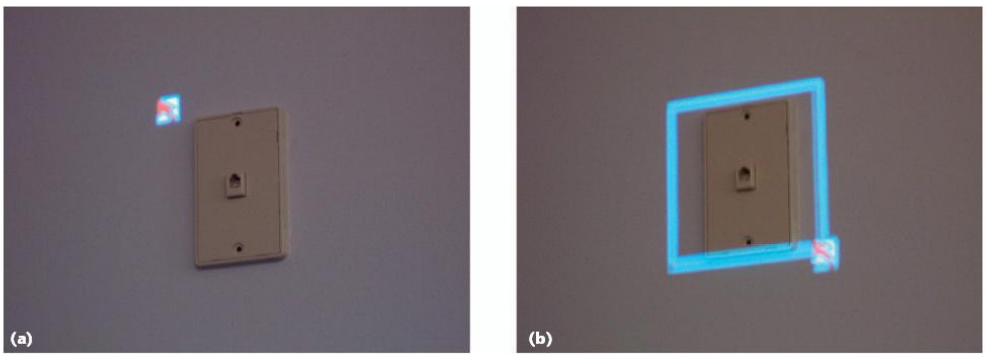


Figure 2.15: Selecting a physical region of interest. (images from [13])

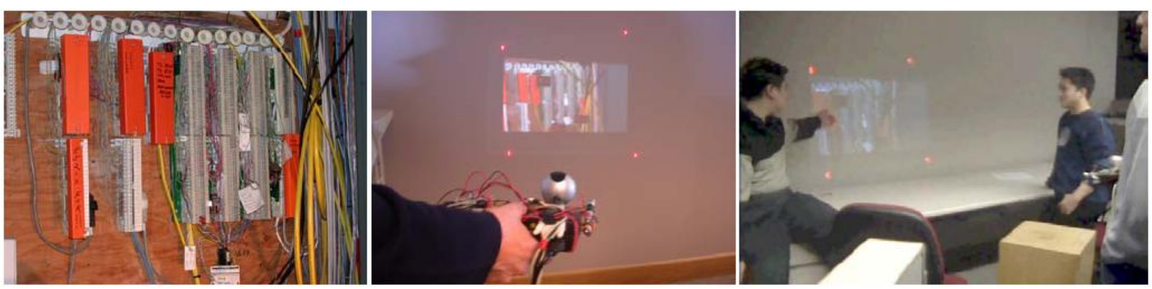
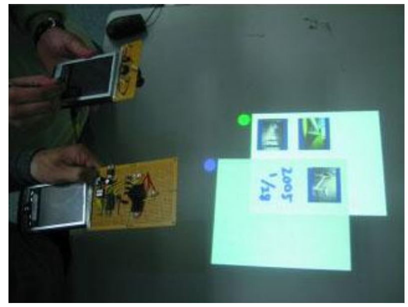


Figure 2.16: Copying and pasting image textures between surfaces. (images from [93])

2.3.3 Interpersonal Information Exchange

Given the sharability of the projection display created by handheld projectors, they are naturally suited to support interpersonal communications, which is difficult to achieve with traditional handheld devices. The Hotaru system [109] supports simple interaction between multiple projectors. File transfer between devices can be made by either finger touch or overlapping the projection images (Figure 2.17).

(a) By finger touch.



(b) By overlapping projections.

Figure 2.17: Information exchange between projectors. (images from [109])

On the other hand, user-authored information may be left in the physical environment to facilitate asynchronous interpersonal communication. Raskar et al. [93] also described in their warehouse scenario how annotations are used to communicate information between a manager and an employee (Figure 2.14).

In this thesis, I explore a variety of applications and scenarios of handheld projector usage, which include but are not limited to the three areas above.

2.4 Related Research Areas

In this Section, we briefly summarize other research areas that are related to or have implications on handheld projector interaction.

2.4.1 Handheld Device Interaction

Given the increasing popularity of handheld devices in recent years, extensive research has been conducted on various techniques to interact with them. Researchers have thoroughly investigated gesture-based interaction techniques [56, 59, 60, 76], text entry techniques [79, 90, 110, 135], multimodal interactions [92, 96, 113, 131], and so on.

Specifically, several visualization and interaction techniques were explored to overcome the shortcoming of the small-sized display and interaction space on handheld devices. ZoneZoom [98] and City Lights [135] supported zooming visualizations and interfaces to maximize space utilization. Halo [11] is an efficient technique to visualize the locations of off-screen objects. Collapse-to-Zoom [12] allows the user collapse irrelevant content to make space for important information. In particular, several researchers [41, 61, 116, 130] have explored using a position-tracked handheld device to provide a window to a larger virtual workspace, similar to the Flashlight metaphor we employed in our research. However, all these solutions can only alleviate but not eliminate the limitation of small handheld displays. By embedding projectors in handheld devices, the display and interaction space will no longer be limited by the size of the device itself.

2.4.2 Projection-based Interaction

In addition to handheld projector interaction, researchers have also explored rich interaction in a physical environment enabled by situated or steerable projectors. Pinhanez and colleagues present the Everywhere Displays project [85, 86], which used a mounted steerable projector combined with a computer vision system to transform any surface in a room into an interactive interface (Figure 2.18). They have applied their system in a retail store scenario [110].

Figure 2.18: Everywhere Displays. (images from www.research.ibm.com/ed)

Dietz et al. [33] use multiple projectors to create interactive public displays targeted at persuading people. They also supported implicit interaction depending on the distance between the user and the projected display. Flagg and Rehg [42] use multiple projectors to alter the appearance of an artist's canvas and guides the artist to construct the painting. Lee et al. [69] present a technique for projecting content onto movable surfaces that adapts to the motion and location of the surface using a mounted projector.

Empirical studies were also conducted to investigate the user experience and performance with projection-based interaction. Voida et al. [120] reported a Wizard-of-Oz study on users' preference of object manipulation techniques in a projector-based augmented environment. Podlaseck et al. [87] studied people's reactions to user interfaces projected onto different real-world objects.

The use of handheld projector provides similar affordances like those provided by the above projection-based systems, but avoids instrumenting the environment, therefore is more suited to ubiquitous usage.

2.4.3 Collaborative Work

Computer supported cooperative work (CSCW) has been extensively explored by researchers. These collaborations between people can happen either in a co-located setting, or across time and location intermediated by computing systems.

Co-located Collaboration

Co-located collaborative groupware has been widely investigated in other settings, especially with shared displays such as walls [22, 54, 62, 106] or tabletops [114, 136, 138]. Huang and Mynatt [58] used semi-public displays to support co-located group members to maintain awareness and collaborate. Morris et al. explored different strategies of control mechanism in applications on shared tabletop displays [79], and also multi-user coordination policies for co-located groupware in general [80]. Simon [106] studied the first-person experience and usability of co-located interaction in a projection-based virtual environment. Shoemaker and Inkpen [105] also explored techniques to present private information in a co-located single display groupware (SDG) application.

Inkpen and colleagues studied co-located collaborative entertainment for adults [74] and children [59].

Similarly, we may achieve co-located collaboration and communication between people by sharing a handheld projection display, or by interaction between several handheld projectors, each owned by a user.

Asynchronous Collaboration

Researchers have also explored collaboration between people at different times. Edwards et al. [36] present Bayou, an infrastructure to support the construction of asynchronous collaborative applications. Weng and Gennari [124] studied the use of annotations in supporting asynchronous collaborative writing, and Cadiz et al. [27] studied the role of web annotations in asynchronous collaboration around documents. Preguiça et al. [88] present research on data management support for asynchronous groupware. Ganoe et al. [48] used collaborative public displays to support activity awareness in asynchronous distributed work. Sakamoto and Kuwana [100] also explored integrating support for both synchronous and asynchronous communication in cooperative work.

Similarly, with the ability to leave information and annotations in the physical environment, handheld projectors can effectively transform the environment into a shared conduit to support asynchronous communication and collaboration.

2.4.4 Augmented Reality

Augmented reality (AR) [5] is a variation of Virtual Reality (VR) [96]. Instead of completely immersing the user inside a synthetic environment as in virtual reality, augmented reality superimposes virtual objects and information upon or composite with the real world.

Most augmented reality systems make use of see-through head mounted displays (HMD) [57, 122, 131] to overlay information while the user is observing the physical world through them (Figure 2.19). Researchers [83, 121] have also experimented using camera-enhanced handheld devices as displays to create a lightweight indirect augmented reality experience (Figure 2.20).

Figure 2.19: See-through HMD. (images from [5])

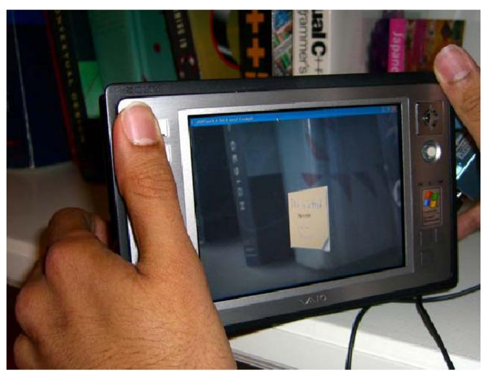


Figure 2.20: Handheld augmented reality. (image from [83])

Applications of augmented reality includes assisted medical surgeries [7, 108], instructions for manufacturing and repair [38, 107], annotations of physical objects and environments [37, 97, 99], robot teleoperation [34, 67] and entertainment [30, 72, 76].

The handheld projector can also be used to overlay information in the physical environment, as discussed in section 2.3.2. The sharable display created by handheld projectors supports collaborative applications better, compared with conventional AR devices that are designed typically for single-person use.

2.4.5 Public and Social Games

Computer games that involve social interaction between players have attracted researchers in recent years as an artifact to investigate social behaviors mediated by computing. Researchers have studied social games both played on the internet [35] and in the real world [17, 24], and those that involve both [43].

Specifically, Social games played on mobile devices have become popular recently as entertainment and as a research topic. A notable example is Mogi [63], a location-based game for mobile phones launched in Japan. The players collect virtual items depending on their locations, and trade them with other players to complete collections while they are moving outside. Researchers have studied similar location-based mobile games [9, 15] that people play while doing something else, and observed many interesting social interaction patterns from the players [29, 31]. Benford et al. [16, 18] also investigated how uncertainty and error in location can influence the game experience and players' strategies. These games are played exclusively on mobile devices, hence the game experience is only understood by the players but not surrounding spectators. In particular, Bell et al. [15] reported that player behaviors appeared strange to other people around and drew unwanted attention.

On the other hand, a few researchers have also experimented with games and interactive entertainment using situated large displays in public spaces. Games played on these public displays are visible to all people in the space, and might seamlessly blend into public space experiences. Schminky [95] is a multi-player game played on PDAs. The game was deployed in a café for one week, with a public display showing the social network that resulted from game playing. MobiLenin [101] is an entertainment system that allows people to use mobile phones to vote for music video clips to be played on a public display. FishPong [132] is a ball-and-paddle style game played on a tabletop display using augmented coffee cups, designed as an "icebreaker" game. Public display games have also seen their emergence as commercial products, especially those using projected displays that people can interact using their bodies, such as the GestureFX™ system (www.gesturetek.com/gesturefx/introduction.php).

In addition to those discussed here, many other games that aim at pervasive experiences in the real world have been explored. Magerkurth *et al.* provides a more detailed review [73].

Given the special affordances of handheld projectors, game applications supported by them could combine the attributes and advantages of both mobile social games and public display games, and create a unique game experience for players and spectators.

Chapter 3

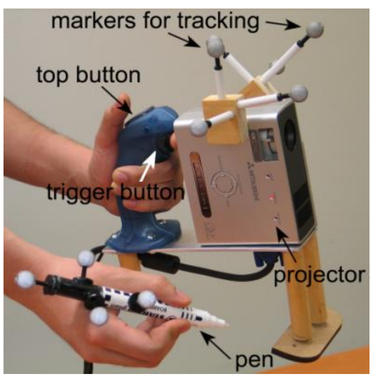
Prototype Platform and Mechanism

With the ability to project information, a handheld device can surmount the limitations of its small internal screen by creating a larger display on an external surface. Furthermore, if the projector's position and pose information is available, the system could change the displayed information accordingly, and create an illusion of exploring large workspaces embedded in the physical environment as if using a flashlight (Figure 1.3b, c). By doing so, the display and interaction space can be expanded to cover almost an entire physical environment, and support interaction concepts not possible on traditional desktop or handheld devices.

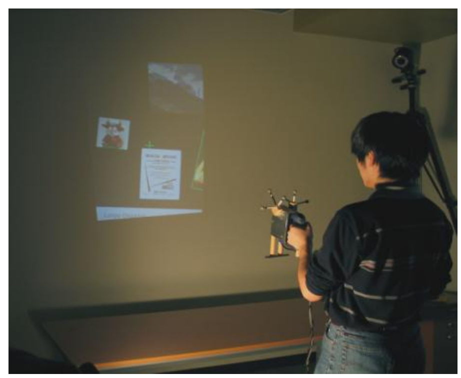
Building on these unique affordances, we develop a handheld projector interaction prototype as the research platform for all our following work. The system employs a Flashlight metaphor, in which the projected image reveals a portion of a larger virtual workspace that is stationary on the projection surface. Multiple workspaces can be embedded in the physical environment. To enrich the interaction possibilities and leverage human ability to perform bimanual tasks, a passive pen is included to write/draw on the surfaces. In this chapter, we describe the concepts and implementations of the platform system, as well as techniques to interactively define workspaces in a physical environment.

3.1 Basic Concepts

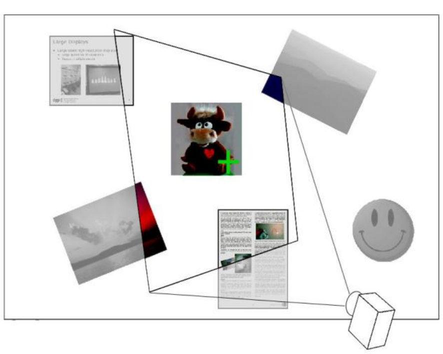
The central concept of our interaction scheme is a flashlight metaphor. The image projected on a physical surface by the handheld projector reveals a portion of a larger virtual workspace that is situated on the surface (Figure 3.1c). When the projector is moved, the projected image content changes accordingly to reflect the change of the projected region relative to the workspace. By doing so, we create an illusion of exploring a stationary workspace relative to the physical surface. When multiple workspaces correspond to different physical surfaces (Figure 3.1d), this results in a user experience of looking around in a dark room with a flashlight. For the sake of simplicity, we refer to the physical projection surfaces corresponding to these virtual workspaces as "surfaces" hereafter.



(a) Handheld projector.



(b) System in use.



(c) Flashlight metaphor.

(d) Multiple workspaces.

Figure 3.1: Prototype platform and basic concepts.

Virtual objects such as pictures, documents and folders can be contained in the workspaces. To indicate the attention point of interaction, a cross-shaped cursor is positioned at the center of the projected image. The cursor can be freely moved across different workspaces by pointing the projector. The cursor size scales according to the distance between the projector and the surface, so as to maintain a relatively constant visual size for the user. Using the cursor and two buttons attached on the projector, the user can interact with the workspace and the virtual objects. A passive pen is used to draw and perform local interactions. The system responds when the pen tip is touching a surface.

3.2 Hardware Platform

We use a Mitsubishi™ PK10 Pocket Projector to prototype our system (Figure 3.1a). The projector weighs about 1 pound, and has a resolution 800×600 pixels. We augmented the projector with a handle and two buttons (the primary “trigger” button and the secondary “top” button). An integrated stand enables placing the projector on the table in a comfortable projection angle. The passive pen is made from a whiteboard marker with the tip replaced by plastic, and without any electronics embedded. Passive reflective markers attached to the projector and the pen enable tracking by a Vicon camera-based tracking system (www.vicon.com) which provides 6-dof (position + pose) information for both the projector and the pen at millimeter precision at up to 120Hz. The projector is connected to a 2.4GHz Pentium 4 PC, which produces the projection image and handles the interaction.

The Vicon tracking system enables us to track the projector and the pen with high precision and low latency in a relatively small region. However, we anticipate upcoming wireless location tracking systems such as such as indoor GPS [117] or TrackSense[68], possibly combined with on-projector sensors such as tilt sensors, will soon enable such tracking more cheaply and ubiquitously, so as to allow our designs to be widely deployed in the near future.

3.3 Image Generation and Correction

The projector's optical model is represented by a projection matrix that describes the transformation between a 2D point on the projector's image plane and its projection in the 3D world. Each workspace is represented as a 3D plane that describes the near-planar surface it corresponds to, and the valid region of the workspace (represented as a rectangle) that information and interaction must reside within. At each frame, using the projector model, the workspace models, and the current position and pose of the projector acquired from the tracking system, a linear system is solved to calculate which workspace the projector is pointing at, and the exact position and shape of the projected image region. The system then decides what information to project and warps the image accordingly to compensate for the distortion caused by non-orthogonal projection (Figure 3.2).

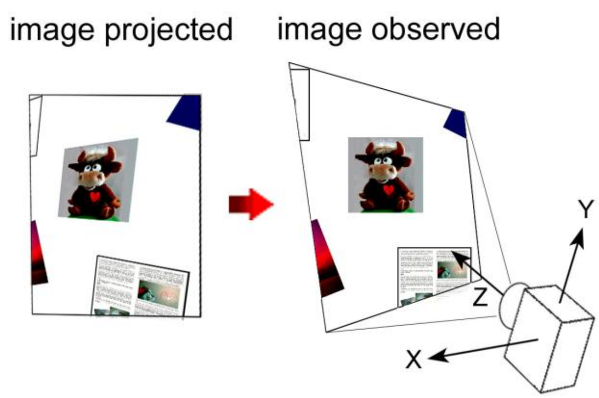


Figure 3.2: Display mechanism.

For clarity, we define the local X, Y and Z axes of the projector as in Figure 3.2. Certain interaction techniques are associated with rotating the projector along some axes.

The projector is calibrated beforehand by detecting correspondences between 2D image points and 3D projection points using standard computer vision techniques [45], while the workspaces can be defined (calibrated) interactively during use, as described later in Section 3.4. To decide whether the pen is being used, the system checks the distance between the pen tip and the surfaces. The pen is active when its tip is less than 3mm from a surface and falls within a valid workspace.

We use the dynamic recursive low pass filter described by Vogel and Balakrishnan [119] to reduce the input jitter caused by hand movement and tracking noise without noticeable lag.

3.4 Workspace Definition Techniques

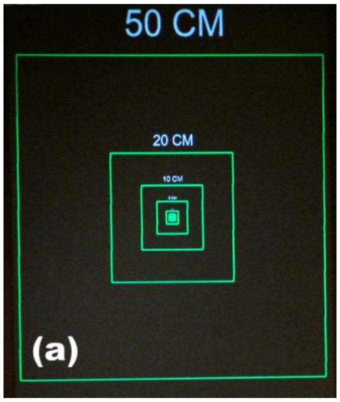
When entering a new environment, the user first defines workspaces corresponding to the physical surfaces in that environment. Any near-planar surface can be potentially defined into a workspace, e.g. walls, tables, boards, or multiple surfaces of a physical object. To define a workspace, the system needs to collect two pieces of information: (a) the 3D plane that approximates the surface; (b) the valid region of the workspace.

Depending on the reachability of the surface and the precision needed for the workspace, the user may choose to define a workspace either using the projector or using the pen.

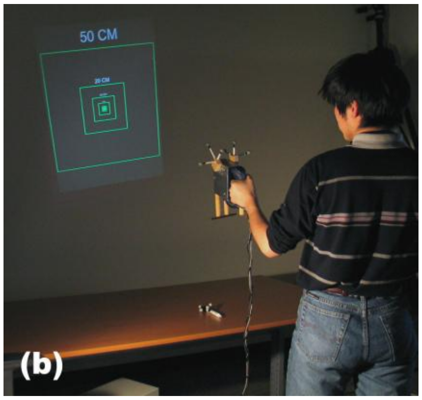
3.4.1 Defining a Workspace using the Projector

To define a new workspace, the projector projects a set of concentric quadrangles, each notated with text indicating the desired side length of it, ranging from 1mm to 100m (Figure 3.3a). These quadrangles stay static relative to and distort along with the projected image region, just like being projected by an ordinary non-rectified projector. The user's task is to point to the surface to be defined, and adjust the projector pose to make the quadrangles appear as upright squares. At the same time, the user can enlarge the squares by pressing the secondary button, or shrink them by pressing the primary button in order to roughly match the displayed sizes of the squares to those suggested by the notations. Alternatively, the user may also move the projector towards or away from the surface to finely adjust the displayed sizes. Once satisfied, the user pressed the two buttons together, and the surface plane of the new workspace is defined accordingly.

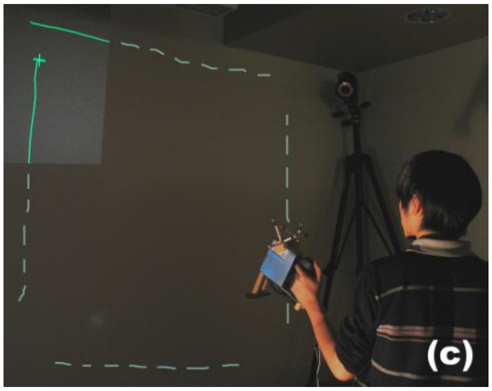
Then the user defines the valid region by sketching the four border lines of the rectangle. Pressing down and holding the primary button, the user starts to sketch the top border line using the cursor. Releasing the button finishes the line. Similarly, the user then sketches the remaining border lines in the order of right, bottom and left, and the valid region of the workspace is defined accordingly. The sketched border line segments do not have to form a closed rectangle.



(a) Projected pattern.



(b) Defining the surface.



(c) Defining valid region (dotted lines added for illustration purposes only).

Figure 3.3: Defining a workspace using the projector.

We now explain the internal mechanism for this definition procedure. In the surface plane definition stage, the system assumes that the surface plane is always perpendicular to the Z axis of the projector (namely the optical axis) with a fixed distance D from the projector's optical center (Figure 3.4). Virtual squares are projected onto this imaginary plane with physical sizes indicated by the text notations, and centered at the point that the projector's Z axis intersects the plane. Since the position and orientation of the imaginary plane is updated along with the projector, the projected squares stay static relative to the projection image on the real surface, and distort along with the projection image. By making the displayed quadrangles squares, the user effectively makes the imaginary plane parallel to the real surface to be defined. Pressing the primary/secondary button increases/decreases D, *i.e.*, moves the imaginary plane along the projector's Z axis, and results in changes in the sizes of the projected squares on the real surface. By the time the displayed size matches that indicated by the notation, the imaginary plane matches the real surface exactly. Pressing the two buttons together locks the position of the imaginary plane, and defines it to be a workspace.

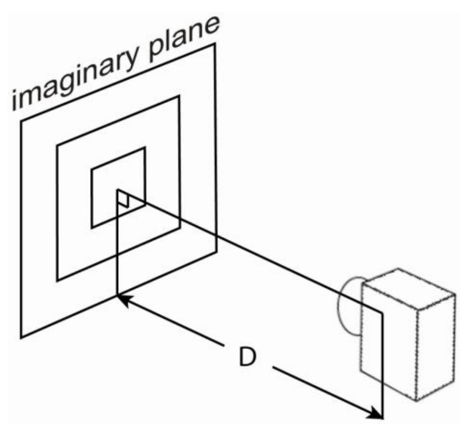


Figure 3.4: Internal mechanism for defining workspace using the projector

After the workspace is defined, the rectangle of the valid region is calculated by fitting a straight line to each of the border lines the user sketches.

This procedure is applicable to virtually all near-planar surfaces present in the environment, except the rare cases when the user cannot possibly point the projector perpendicularly to it, such as a screen hanging in the air above the user. The accuracy of this procedure is limited by the user's subjective estimation of the squares' shapes and sizes. However, the definition error is negligible if the user only needs to interact with the workspace using the projector, especially when the surface is far from the user. In order to use the pen on the surface, a more precise definition is required, which can be achieved using the pen itself.

3.4.2 Defining a Workspace using the Pen

When the surface is within physical reach, the user can define the workspace using the pen. To do so, the user sketches the four border lines of the workspace (Figure 3.5), similar to the region definition stage using the projector. Instead of using the projector cursor, the user sketches the border lines using the pen tip touching the surface. Again, the user uses the primary button to indicate start and end points of each border line. After border lines are sketched, both the surface plane and the valid region are defined accordingly in a single step.

Since the pen tip is tracked in 3D, the internal mechanism is simply fitting a plane to all the pen tip points collected during the sketching. The valid region calculation is the same as that in the definition procedure using the projector.

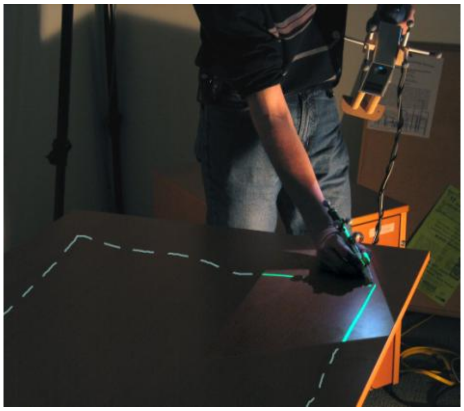


Figure 3.5: Defining a workspace using the pen.

(dotted lines added for illustration purposes only)

These two definition techniques can be used in combination to efficiently define multiple workspaces, and accommodate different surface characteristics. For example, to precisely define a large wall, the user may first define the wall plane using the pen, and then use the projector to define the border lines that are not reachable by the pen. These techniques can also be used to revise an existing workspace, by simply re-specifying the relevant parts. These definition techniques are available to the user in a special workspace definition mode, as opposed to the normal interactions described in Chapter 4. The workspace definition mode is automatically in effect when a user enters a new environment, or can be triggered using a crossing menu (described in Chapter 4) at any time by the user.

Enabled by the *Flashlight* interaction metaphor and the ability to dynamically define workspaces in various physical environments, our prototype platform can support a large variety of interaction techniques and applications. Building on this platform, we explore interaction techniques and usage scenarios using a single handheld projector in the next chapter.

Chapter 4

Interaction using a Single Handheld Projector

Although some previous research exists on interactive handheld projectors [13, 44, 90, 93], there has not been a systematic exploration of the design space of handheld projector interaction techniques. Most of the existing techniques are either application specific, or still relying on conventional mouse-like operations, which may not be suited to the characteristics and affordances of handheld projectors. In order for the handheld projector to be usable as a general and flexible interaction medium like the desktop computer, a vocabulary of generic interaction techniques needs to be developed, just as the WIMP (window, icon, menu, and pointing device) scheme was evolved into the standard for desktop GUI. As higher-level applications are to be explored on handheld projectors, these basic techniques would serve as building blocks that could be readily used in combination to create fluid user interfaces. To achieve this goal, and building upon the platform described in Chapter 3, in this chapter we explore the design space of interaction techniques using a single handheld projector, aiming at supporting a large variety of potential applications, some of which demonstrated in our exploration of usage scenarios. While designing these techniques, we deliberately consider and exploit the special characteristics and affordances of the handheld projector. On the other hand, the ability to project and annotate in a physical environment offers the possibility to augment physical objects with overlaying information. Although these techniques are designed for using a single projector at one time, several multi-person scenarios can also be supported by sharing the viewing experience, or communicating asynchronously using the physical environment as a conduit.

4.1 Design Principles

In our exploration, we are guided by the following goals:

Generic interaction schemes

Instead of designing for specific applications, we aim for a set of generic techniques that would be useful for any handheld projector application. Depending on the scenarios, users could combine these low-level interaction techniques to establish higher-level dialogs that are suited to the context.

Applying real-world protocols

Since the interaction spaces are integrated into the physical environment, we adopt protocols that people already use to interact with the physical world and other people where applicable, such as moving towards an object to see details, and using one's body to block other people from seeing private information.

Division and integration between projector and pen

To leverage the different affordances of the projector and the pen, we deliberately assign tasks that are more global and coarse to the projector, and local and precise tasks to the pen. However, where appropriate we also exploit concurrent bimanual interaction with both devices.

Supporting multiple people

Although we focus on the case where only one user operates the projector at a time, the sharability of projected information naturally encourages multiple people to view and work with it. Our designs attempt to accommodate multiple people where possible.

4.2 Interaction Techniques

In the following sections, we describe techniques for interacting with workspaces after they are defined.

4.2.1 Projector-Only Interactions

Using the projector alone, the user can perform tasks that focus on browsing and organizing information.

Virtual Object Manipulation

The basic manipulation of virtual objects is similar to that in a desktop GUI. A click of the primary button selects the object at the cursor position. Holding the primary button, the user can move the object around by moving the cursor. Releasing the primary button releases the object. The object can be seamlessly moved either within the same workspace, or between different workspaces.

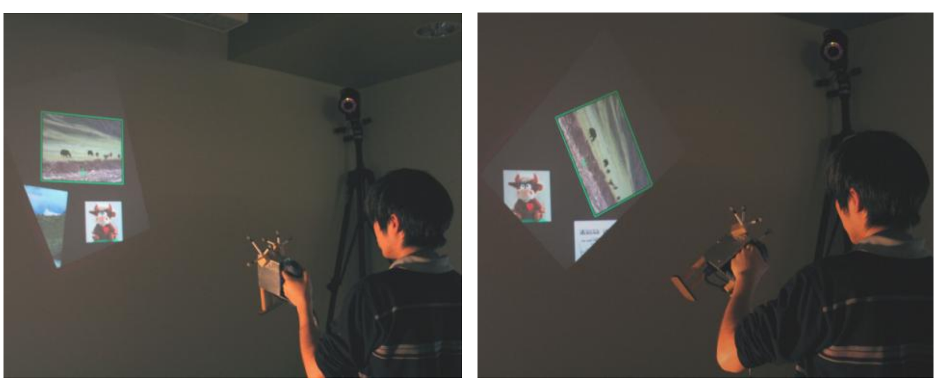
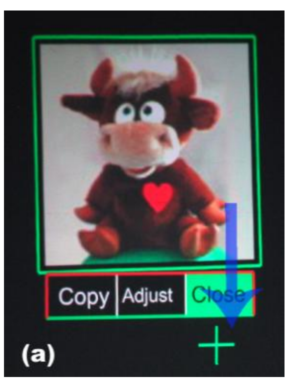


Figure 4.1: Moving and rotating objects with projector.

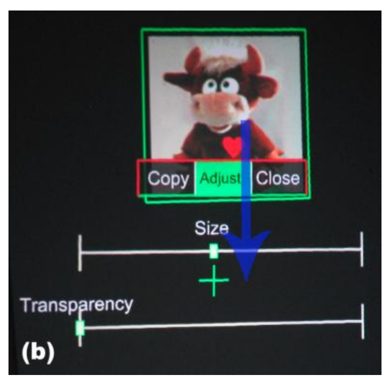
In addition, the user can rotate the projector along its Z-axis (effectively by rotating the wrist) to rotate the captured object. Moving and rotating can be performed simultaneously (Figure 4.1). Since the orientation of the object is consistently mapped to that of the projector, the object stays aligned with the user's view angle, even when it is moved across different workspaces, or when the user walks around the workspace. On the other hand, the user can efficiently position and rotate the object to accommodate other people's perspectives.

Crossing Widgets

Interactive widgets are used to trigger commands and adjust parameters. To provide robustness against the relative lack of precision of free-hand pointing with the projector, which could be worsened by a physical button-click action, we aim at a set of widgets that alleviate the requirement of precise cursor positioning and button-clicks. We adopted the concept of crossing-based widgets [2, 4] in our projector interaction to achieve facile interaction.



(a) Crossing menu.



(b) Crossing sliders

Figure 4.2: Crossing widgets (arrows show the crossing path).

A context-sensitive menu is triggered by pressing the secondary button. Holding the button, the user uses the cursor to cross a menu item to activate it, and crossing the menu in the opposite direction deactivates the item, as suggested by the green and red edges of the menu (Figure 4.2a). Releasing the button executes the currently active item, and dismisses the menu. Hierarchical menus are also supported by crossing the items in different layers sequentially.

Crossing based sliders are used to adjust continuous parameters. Crossing the slider activates it, and then the sliding block moves to match the cursor movement parallel to the slider. The parameter value changes accordingly. Note that instead of grabbing the slider block, the user can cross anywhere on the slider to directly dial the block to that position. Crossing the slider in the opposite direction locks the slider block. Using sliders arranged in a row, multiple parameters can be adjusted in a continuous manner. In this case, the previously active slider is locked once another slider is activated. If the user does not want to change some of the parameters, s/he simply moves the cursor around the corresponding sliders to avoid activating them. Figure 4.2b illustrates using crossing sliders to adjust the size and transparency of an object. Incidentally, crossing sliders are also used as scrollbars for scrollable objects such as documents or presentation slides.

By associating different sliders (or groups of sliders) with different menu items, we provide a fluid way to adjust many parameters in a single continuous sequence of crossing actions without need for numerous button-clicks.

These interactive widgets automatically move and scale, so that they always stay inside the projected image. Hence the user can operate the widgets using local movements, while reposition them within/between workspaces using more global movements, without need for a mode switch. Tracking menus [40] provide similar affordances to operate and reposition a widget with a single pointer. However, since our design takes into account the projected image region, the result is a more implicit and less intrusive mechanism. In addition, it smoothly accommodates the user's movement in the physical environment. Because the size of the projected image is proportional to the distance between the user (hence the projector) and the surface, scaling the widgets accordingly results in approximately the same visual size and the same amount of hand movement to operate the widgets.

When invoked, the widget also adjusts its orientation (0°, 90°, 180° or 270°) to accommodate the user's view angle, according to the orientation of the projector at that moment.

Resolution Gradation and Information Granularities

A unique characteristic of using a handheld projector is that the local image resolution changes depending on the distance between the projector and the surface. When the projector moves closer to the surface, the same amount of image pixels is distributed in a smaller region, resulting in a higher local resolution. This naturally matches user's experience of viewing physical spaces. From afar the user has a large but low-resolution display region to overview the workspace, while coming closer, the user focuses on a small but high-resolution region to acquire details. This resolution gradation spontaneously happens in an implicit continuous manner.

Figure 4.3: Proximity adjusts information granularity.

In addition to image resolution gradation, dynamic resolution can be applied to the semantics of information. A virtual object may switch between different information granularities when the distance between the user and the surface changes. High-level information is displayed when the user is far from the surface, and detailed information is displayed when the user comes closer. This is realized by calculating the area ratio between the virtual object itself and the projected image region, which effectively measures how many image pixels are distributed to the object, given the total amount of pixels of the projector is fixed. This also results in a switch of information granularities while scaling the object. Figure 4.3 illustrates a city map switching between granularities when the user moves.

Workspace Overview

Since the projector usually displays only a portion of the workspace at one time, the user may occasionally have difficulties in navigating the workspace and locating a virtual object. To assist navigation, pressing both buttons together triggers an overview panel showing a miniature of the current workspace (Figure 4.4). Like the interactive widgets, the panel moves and scales with the projected image region. When moved into a different workspace, the content of the panel switches to overview the newly entered workspace.

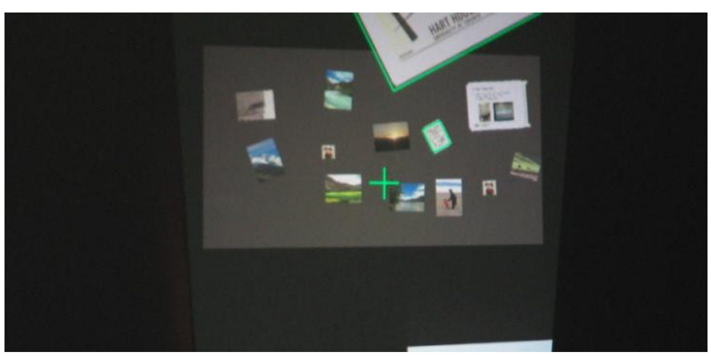
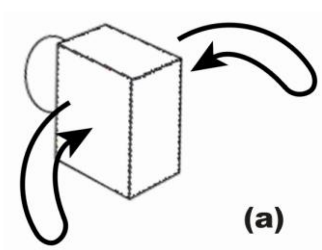


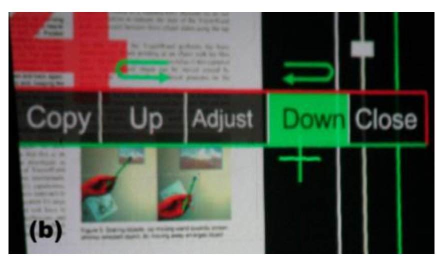
Figure 4.4: Workspace overview panel.

Flick Gesture

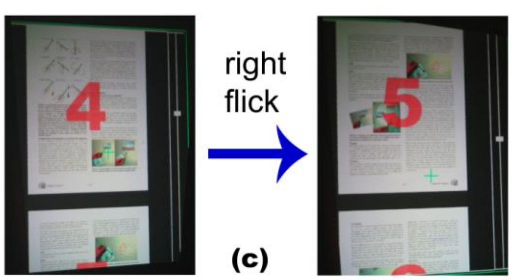
Depending on the currently selected object, a “flick” gesture to the left or the right acts as a shortcut to frequently used menu commands related to it. To make the flick gesture, the user quickly rotates the projector along its Y-axis to the left or right and then back again (Figure 4.5a). To inform the user about the shortcut, the corresponding menu item is marked with a gesture icon (Figure 4.5b). Figure 4.5c illustrates using a flick gesture to page up/down in a document, and Figure 4.5d illustrates using it to switch between functions (magnifying, increasing contrast, and querying information) for a multi-functional magic lens. Incidentally, in the information query mode, the information text rotates along with the magic lens, providing a way to adjust the text orientation to accommodate different view angles, either for different people, or for the same user at different times. We choose to only employ this flick gesture because of its simplicity, and do not further pursue other possible gestural interactions, which are beyond the focus of this research.



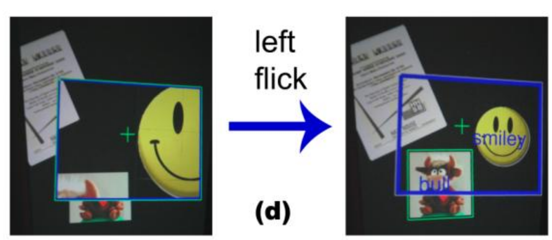
(a) Gesture trajectory.



(b) Gesture icons in a menu.



(c) Turning pages.



(d) Switching lenses.

Figure 4.5: Flick gesture.

Personal Folder

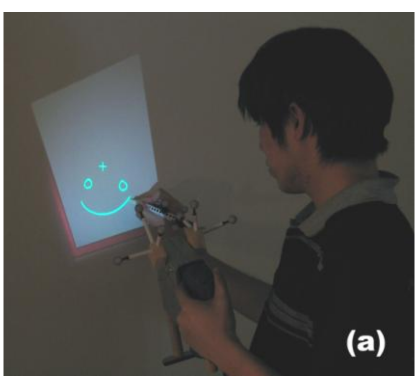
A personal folder contains the virtual objects with which the user may want to interact. Through a menu command, this personal folder can be accessed in any workspace, and the objects inside it can be dragged into the workspaces.

Depending on the usage scenario, these virtual objects might be stored in the handheld device (personal data), in devices in the environment (context data), or both.

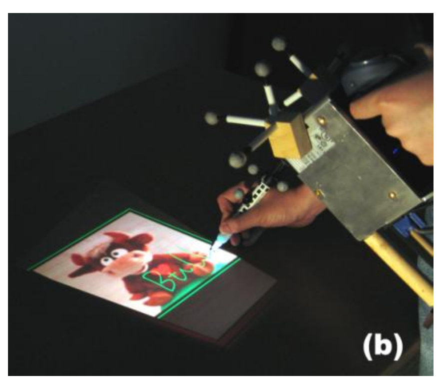
4.2.2 Pen and Projector Interactions

Using the passive pen along with the handheld projector, the user can draw pen inks in workspaces, annotate on virtual or physical objects, and perform local interactions more precisely.

Although most of the pen techniques described below do not necessarily require holding the projector (the projector could be put down on a table, or not used at all in an "eyes-free" scenario), we suspect in most cases users will hold the projector with the non-dominant hand to set the display/interaction context, and use the dominant hand to perform pen interactions. Therefore these pen-based techniques are described as bimanual interaction techniques.



(a) Drawing on a surface.



(b) Annotating a virtual object.

Figure 4.6: Pen interaction on surfaces.

When the pen tip touches the surface, a pen stroke can be drawn in the workspace or on a virtual object (Figure 4.6). Neighboring strokes are grouped into ink. Ink can be moved, rotated, scaled (using a crossing-slider) and closed (using a crossing-menu) just like other virtual objects. The pen does not have to reside inside the projected image region to draw, hence the user may make “blind” notes while viewing other portion of the workspaces, or make “secret notes” when s/he does not want other people to see what is being written.

In order to annotate on a remote virtual object that is out of physical reach, the user can use the projector cursor to capture and drag the object to a closer location while holding the primary button, and then annotate on it using the pen. To avoid jittery inking caused by the movement of the hand holding the projector, the object’s position is frozen once the user starts writing. Incidentally, when the object is too large to be displayed completely by the projector, this enables the user to temporarily pin down the object with the pen and move the projector to browse it. Once the user finishes the annotation, s/he releases the primary button, and the annotated object flies back to its original position, as if it is spring-loaded. This provides an efficient way to annotate virtual objects scattered around a large physical environment, without the need to walk around or rearrange the objects.

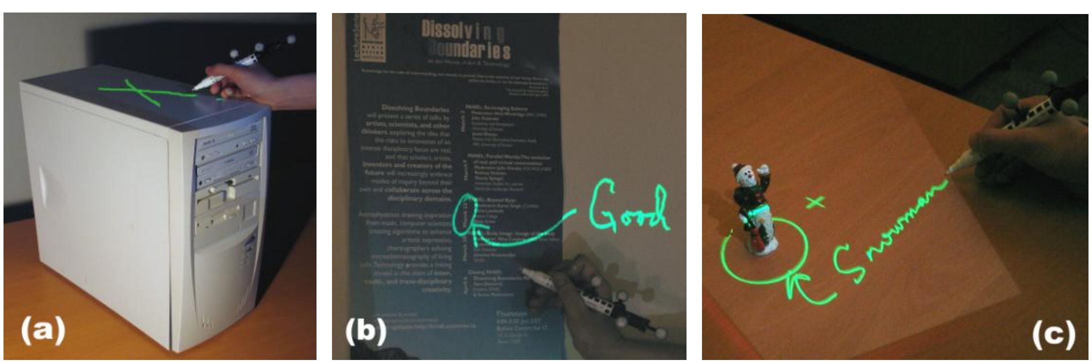


Figure 4.7: Annotating physical objects.

Utilizing the functionality of drawing in workspaces, the user can also annotate static physical objects in the environment. Depending on the relationship between the physical object and the workspace, the object may be annotated in three different ways: (a) on the surface of the object, when the surface has been defined into an workspace. This applies to relatively large objects with near-planar surfaces, such as a box or a computer. (b) Over the object, when it forms part of a workspace. This applies to flat objects attached to other surfaces, such as a poster or a brand. (c) Around the object, when it protrudes from a workspace. This applies to relatively small objects that stay on top of other surfaces, such as a toy on the table, or a switch on the wall. Figure 4.7 illustrates.

The pen can also be used to perform local interactions with a virtual object, which enables precise control that is difficult to achieve with the projector cursor. Figure 4.8 illustrates using the pen to operate the crossing-based scrollbar of a document. Note that the pen does not have to act inside the projected image. Once the user activates the scrollbar, s/he can focus the projector on the document and scroll it using the pen, without paying visual attention to the widget any more. Complete eyes-free operation of the widget is also possible if the user has become sufficiently familiar with its position. This enables the user to focus on the information content rather than the interface.

Figure 4.8: Using the pen to scroll a document.

Although we describe these pen-based techniques as bimanual interactions, they may also be used for multi-user interaction where one user controls the projector to set the context, and another uses the pen to perform the actions.

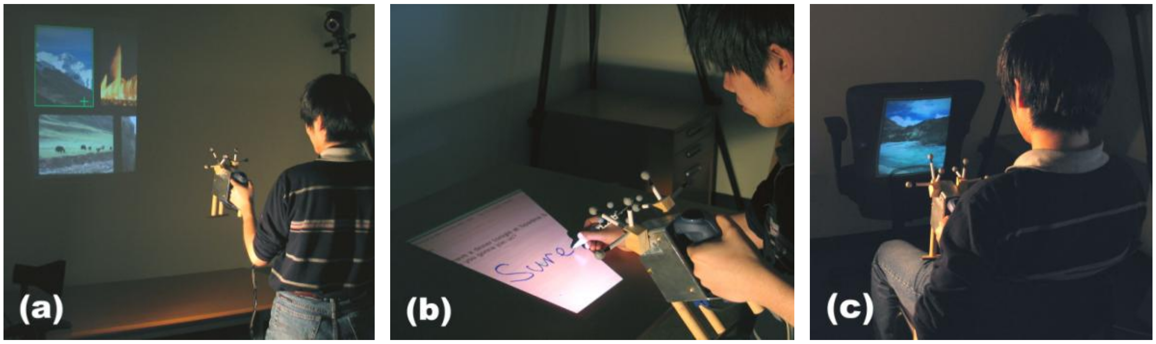
4.3 Usage Scenarios

Utilizing the interaction techniques described, various user scenarios can be supported. Instead of focusing on specific applications, we categorize these scenarios by the generic interaction patterns, both between the user and the system, and between different people.

4.3.1 Single-Person Usage

The capability to easily and quickly define workspaces anywhere enables the user to temporarily convert the physical environment around into workspaces to manage personal information (Figure 4.9). Depending on the environment, these workspaces can take various forms, such as walls, tables, floors, seatbacks, or even the user's lap. The dynamic image size and resolution of the handheld projector readily accommodates interacting on surfaces of different sizes and distances. The possible large scale of the workspaces supports tasks that are difficult on a traditional handheld device with limited display and interaction space, such as managing a personal album, browsing a map, or annotating a large document. When the user leaves the environment, s/he closes the workspaces using a menu command, and takes away the processed information along with the device.

Spatial memory can help the user to efficiently organize information, especially when multiple workspaces are available. For example, in an office room the user may place different groups of information on different walls, and designate the ceiling as a recycle bin. By doing so, the user can easily locate the information, and use coarse hand movement (possibly eyes-free) to organize the information between walls. Taking this particular situation as an example, moving a virtual object upward onto the ceiling means "send to recycle bin" to the user, and almost acts as a gesture-like command.



(a) Managing photo album. (b) Replying to an Email. (c) Interacting on the back of a seat.

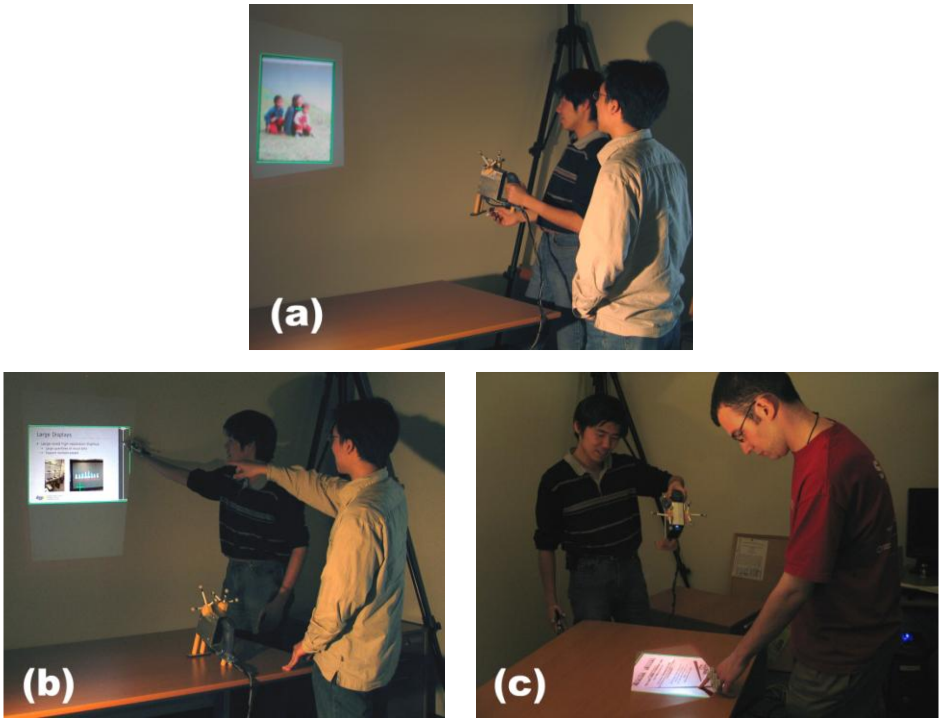
Figure 4.9: Single user scenarios.

4.3.2 Synchronous Multi-Person Usage

Our techniques can also support communication between multiple people. Here we focus on the case where one main user controls the projector and shows the information, while other people look at the information and use the pen to communicate. Scenarios that involve use of multiple projectors together are explored in Chapter 6.

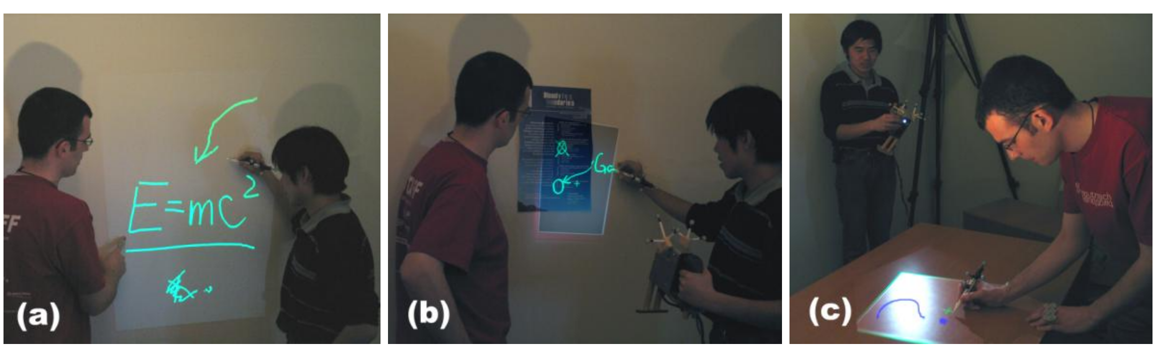
Given the mobility and lightweight interactions of the handheld projector, it is particularly suited for casual ad hoc collaborations. Similar to the single-user scenario, people can temporarily annex their physical environment into workspaces to support their communication. Even a simple task like showing information to others can utilize different interactions depending on context (Figure 4.10). Note that privacy concerns may arise when the main user needs to process personal information that s/he does not wish other people to see. In this case, moving close to the surface results in a small display region, which can be blocked from others using the user's own body.

Multiple people can also use the pen to facilitate communications. A blank wall can be turned into a virtual whiteboard to support brainstorming (Figure 4.11a). People can also collaboratively annotate a virtual object (e.g. editing a document) or a physical object. Figure 4.11b illustrates two people discussing about a physical poster. When the main user is far from others (e.g. a teacher in a classroom), s/he can use the projector to pass on a virtual writing pad (or any other virtual objects) to them and let them write on it (Figure 4.11c). Once the main user releases the button, the writing pad flies back to him/her with the notes made by other people. This provides an efficient way to collect comments from people in a large environment.



(a) Main user holds the projector, showing a picture for both people. (b) Showing presentation slides, with the projector on the table and using the pen to switch slides. (c) Main user positions and orients the document to accommodate the other person.

Figure 4.10: Synchronous multi-user scenarios.



(a) Virtual white board. (b) Discussing a poster. (c) Passing a writing pad.

Figure 4.11: Pen usage in synchronous multi-user scenarios.

4.3.3 Asynchronous Multi-Person Usage

Instead of closing the workspaces when the user leaves, s/he may also let the information stay in the physical environment to facilitate asynchronous communication. When another user enters the environment, s/he can use the projector to retrieve information left by others. Thus, the handheld projector serves as an information conduit between people across time.

Provided the ability to define any near-planar surface into a workspace, the user can leave pen inks and paste information almost anywhere in the physical environment. When another user comes into the environment later, s/he can use the projector to retrieve the information left in the environment by the other people (Figure 4.12). The relationships between physical objects in the environment and the virtual information can be two-way. The virtual information may serve as annotations and auxiliary information for a physical object (Figure 4.12c), or the physical object may serve as a proxy to attract attention to the virtual information (Figure 4.12a, b).

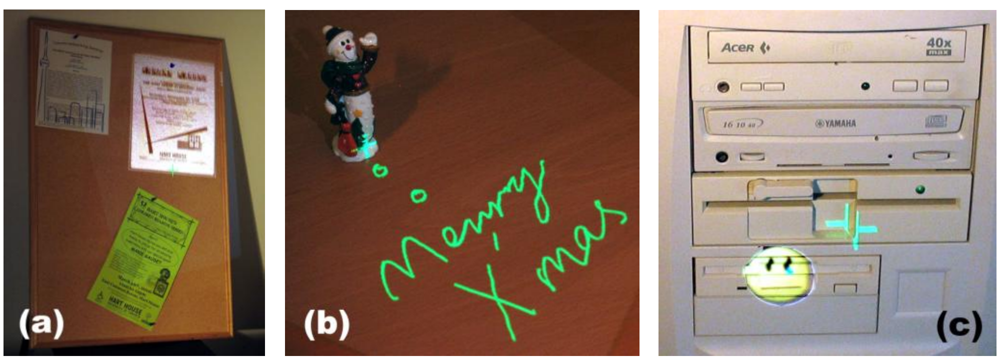


Figure 4.12: Information left in the environment.

(a) Virtual flyer on a physical bulletin board.

(b) Note written on a desk. (c) Error icon on broken device.

The concept of overlaying information in physical environments has been explored in augmented reality applications [8, 14, 25]. However, our ability to easily define workspaces and author the overlaying information on the fly may result in casual asynchronous communication between users. Multiple people may also create a rich information environment through asynchronous collaboration, similar to collaborative information sharing on public displays [22, 49]. This scenario also includes the case that a single user leaves reminders to him/herself for future use.

4.3.4 Game Usage

The affordances of the system also open up design options for games. Figure 4.13 illustrates a mockup of a shooting game. With the guidance of stereo sound effects suggesting the positions of virtual enemies, the player uses the projector to spot and shoot at them. Other game scenarios are possible, such as using the pen as a gestural input device to cast spells, as if using a magic wand. The player can dynamically define the physical game environment, and exploit the physical layout of the environment to diversify the game experience. The result is an alternate game setting that can be deployed virtually anywhere.

Figure 4.13: Shooting game.

4.4 User Feedback

In order to collect some early informal user feedback on our designs, we asked five male graduate students, aged between 20~30, to try the prototype system. Each participant was shown the interaction and workspace definition techniques, and asked to freely and extensively try out the techniques and functionalities. Given our current focus on evaluating the basic interaction techniques rather than the various usage scenarios, the participants attended individually and mostly under the single user scenario. Each trial session lasted 40-50 minutes. We observed participants' behaviors, and informally interviewed them about their opinions and suggestions, and in what scenarios they would use the system.

All participants easily understood the flashlight metaphor and the presence of multiple workspaces. Their first impressions of the system were unanimously "very cool". They did not show or express any difficulties in learning the interaction techniques. The overview widget was found to be especially helpful. Though one participant raised the concern of possible fatigue resulting from moving the projector all the time, nobody actually reported fatigue during the trial periods. Although no participants had prior experience with crossing interfaces, the crossing-based widgets were broadly welcomed. They felt that these widgets not only greatly reduced the problem of jittery input, but also saved time in performing complex operations. All participants were enthusiastic with the ability to leave virtual annotations in the physical environment. They were particularly excited about the idea of leaving "secret" notes, since this reminded them of "magic holographs" in fairy tales.

The main complaints relate to the technical limitations of the system, such as display jitter, having to manually adjust the projector's focus, and the limited working area of the tracking system. Another concern is that because only part of the workspace is displayed, sometimes the user may lose the interaction context, such as when browsing a menu with deep hierarchy, or performing a flick gesture. One possible solution is to let the context temporarily move with the projector in certain situations.

The participants made suggestions for improvements. Two participants expressed the desire to incorporate handwriting recognition to input text. Some participants suggested adding more input channels on the projector, such as a small keyboard or joystick, as present on cell phones. One user asked for the ability to write in the air using the pen.

The participants also conceived many scenarios in which they would use the system, e.g. taking notes, browsing and editing large documents, giving presentations, collaboratively working with others, looking up for context-related instructions, and entertainment.

4.5 Discussion

Although we did not take a formal iterative user centered design process, we carefully considered the particular characteristics of our system when making design choices. For example, we believe that two physical buttons on the projector provide sufficient input bandwidth for the interface. Given that the handheld projector is to be constantly moved in the 3D space, and that the user's visual attention is focused on the projection instead of the projector itself, we suspect that additional buttons could add unnecessary complexity from both physical ergonomics and ease of comprehension perspectives. The adoption of crossing interfaces instead of point-and-click interfaces was motivated by the observations that freehand pointing in 3D space is somewhat imprecise.

In the scenario of asynchronous communication between people, the user needs to discover the information posted in the environment previously. This may be supported in various ways. When the user is seeking information about a physical object, s/he would actively browse around it with the projector. On the other hand, users may leverage physical objects of special meanings to attract people's attention for information update, e.g. a bulletin board with both physical and virtual posters on it, or a small object deliberately left on the desk. Automatic reminders provided by the system may also help, e.g. when the user enters the environment, visual arrows or audio notifications could guide him/her to check for information updates. When the projector can be uniquely identified by the system, e.g. with embedded RFID, user-specific information can also be accessed without needing to log in.

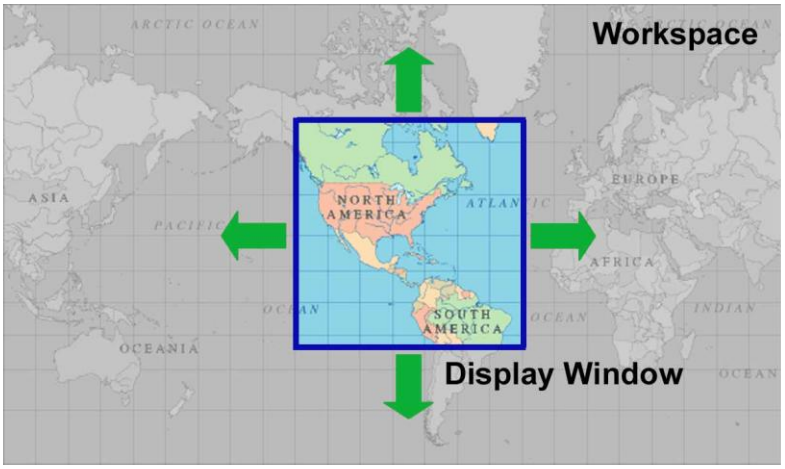
Our system provides some affordances similar to augmented reality (AR) systems (e.g., [83]), such as annotating physical objects. However, the use of a handheld projector makes our system more appropriate for collaborative applications, compared with AR devices which are designed typically for single-person use. Our system also eliminates the need for the system's prior knowledge of the environment, which is essential for similar systems using situated projectors such as Everywhere Displays [85, 86].

The Flashlight metaphor underlying our interaction designs means that the user only sees a portion of the workspace at one time. This interaction style creates both pros and cons. On the advantage side, the ability to choose what and where to display provides the user control over the workspaces. Therefore we can avoid the information overload problem well-known in ubiquitous computing [60], when information is all over the place and overwhelms the user's attention. Furthermore, when multiple people are present, this also avoids disturbing other people with irrelevant information, and supports privacy for personal information. These can be further facilitated by filtering the information to be displayed according to the specific user of the projector. However, not seeing information outside the flashlight (display window) may create navigation difficulties for the user, although alleviated by the workspace overview feature we designed. The loss of interaction context surrounding the current display window may also result in temporary confusions. For lower-level tasks such as target pointing, the user performance may degrade when the user does not know the exact position of the target before it is revealed by the display window. In order to gain a deeper understanding of how the Flashlight metaphor affects user performances on generic interaction tasks, in the next chapter we investigate and create a predictive model for target pointing tasks under such a metaphor, *i.e.*, “Peephole Pointing”.

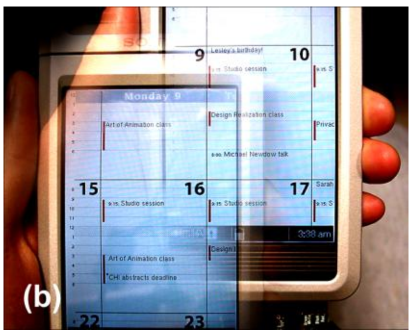
Chapter 5

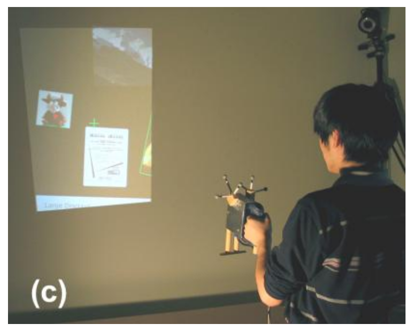
Modeling Pointing Tasks under Flashlight Metaphor

As illustrated in Chapter 3 and 4, the Flashlight metaphor plays an essential role throughout our handheld projector interaction design. By moving the projector (hence the flashlight), the user is able to view and interact with workspaces that are larger than which could be displayed by the projector at once. Similarly, researchers have explored using spatially tracked physical displays as a “peephole” to navigate through larger workspaces [41, 61, 130], therefore overcoming the inherent limitation of the small display size embedded in handheld devices. In general, the Flashlight metaphor refers to physically moving a spatially aware display that acts as a display window to reveal different parts of the virtual workspace, which is stationary relative to the physical world (Figure 5.1). Other terms used by researchers to refer to this interaction style include “Peephole” [130] and “Spotlight” [90] . Given the value of this metaphor for interactive handheld projectors and other handheld devices in general, it is important to understand how it affects users’ ability to perform interaction tasks, especially when information is initially residing outside the Flashlight (or peephole) region. These understandings will provide valuable insights for future designs utilizing the Flashlight metaphor.



(a) Concept.





(b) Using a spatially aware PDA. (image from [130]) (c) Using a handheld projector.

Figure 5.1: Generalized Flashlight (Peephole) metaphor.

As a first step in this direction, in this chapter we study user performance in pointing tasks – a fundamental building block for higher-level interactions – where the flashlight/peephole display needs to be moved to reveal the target, noted as "Peephole Pointing" hereafter. Based on theoretical analysis we propose a quantitative model for peephole pointing, and experimentally validate it. Our work considers how user performance is affected by the size of the display and the presence or absence of prior knowledge of the target's location.

5.1 Background: Fitts' Law

Fitts' law [39] is broadly accepted to model aimed pointing tasks in general. The movement time *T* needed to point to a target of width *W* and at distance (or amplitude) *A* can be expressed as:

$$T = a + b \log_2\left(\frac{A}{W} + c\right) \tag{1}$$

where *a* and *b* are empirically determined constants, and *c* is 0, 0.5 or 1 (details discussed in MacKenzie [70]). The logarithmic term $\log_2(A/W + c)$ is referred to as the Index of Difficulty (*ID*) (measured in "bits") of the task. Kabbash and Buxton [64] show that Fitts' law also applies to the case that the cursor is an area while the target is a point, with *W* set as the width of the area cursor. Researchers have also expanded Fitts' law to model two- and three-dimensional targets [3, 50, 51, 71].

Although Fitts' law has been proved highly successful in modeling conventional pointing tasks, it is unclear whether it applies when the target may be outside the initial display region. Furthermore, it is likely that the size of the display also influences the task performance. We aim at developing a model for target acquisition performance in this peephole pointing situation, which incorporates the display size *S* in addition to *A* and *W*.

5.2 Problem Setup

Our present work is one step towards providing a theoretical foundation and empirical data that can guide the design of interfaces utilizing the flashlight metaphor. We sought to obtain an understanding of the factors that affect peephole pointing performance, and develop a quantitative model that characterizes such behaviors.

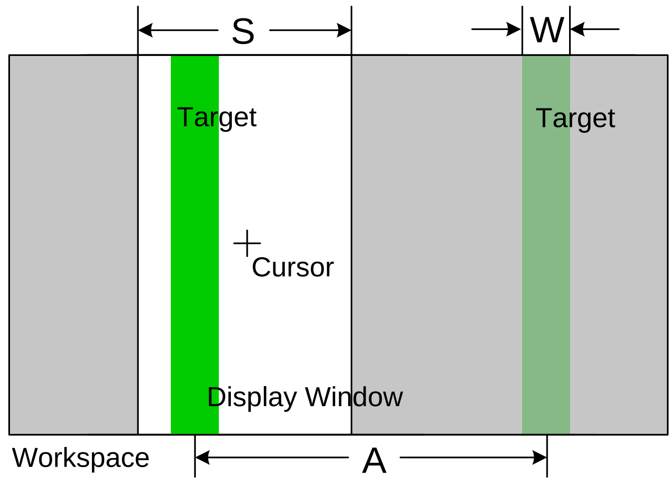


Figure 5.2: Problem setup of peephole pointing.

For the sake of simplicity, we choose to study a one-dimensional pointing task, with the constraint imposed by the target only in the direction collinear to the pointing motion. This is an idealized representation of real world pointing tasks, and is also the behavior that Fitts' law originally models. In practice, the targets are displayed as ribbons with infinite height (Figure 5.2).

The following terms are used throughout this chapter:

Workspace: the entire space in which the targets can reside. Only a portion of the workspace is visible and accessible at any given time.

Display Window: a window that can be moved within the workspace. The workspace region within the display window is visible and can be accessed by the user.

In practice, the shape of the display window may vary depending on the device and the purpose, and could be a rectangle [41, 130], circle [90] or arbitrary quadrangle as in our handheld projector system. However, in our study, because the targets are 1D, we choose to make the display window 1D as well. The display window has a finite size along the direction collinear to the pointing motion, but conceptually extends infinitely along the direction perpendicular to the pointing motion, just as the targets do. By doing so, we maintain the 1D nature of the task. Using a 2D display window may impose unintended 2D constraints on the movement, since only the region within the display window is accessible.

A typical pointing task in this scenario consists of first moving the display window to reveal the target (if it is not inside the initial display window already), and then move the cursor to actually reach it. Three parameters characterize such a task:

A – distance between the initial cursor position and the center of the target.

W – width of the target.

S – size of the display window.

where A and W are variables included in the Fitts' law equation, while S is unique to peephole pointing, and thus needs to be incorporated into our model. In the case of a 1D display window, S is defined as the width of the display window, collinear to the pointing motion.

We also consider a few other factors that may affect the user's performance:

Prior knowledge of target location

When the target is not initially visible, whether the user has knowledge about the target location could affect the amount of effort required to search for the target, and in turn influence the task performance. The prior knowledge may come from previous visits to the target, or from techniques that visualize off-screen targets such as Halo [11].

Cursor control mechanism

Researchers have explored two mechanisms to control cursor position in these situations:

1. Co*upled cursor*: the cursor position is fixed with respect to the display window (typically in the center). Moving the display window also moves the cursor in the workspace. Thus the display also serves as a pointing device, and requires only one input to operate. This mechanism is used by ours (in the projector-only interaction case) as well

as some previous handheld projector systems [90], where a dedicated pointing device is not available or preferable.

2. Decoupled cursor: the cursor position is controlled by a dedicated pointing device that is different from that controlling the display window's position. One possible scenario is where the user moves the display window using the non-dominant hand, while controls the cursor using the dominant hand. The cursor pointing device can move independently of the display window, but is sensed only when within it. This mechanism has been utilized on PDAs and handheld screens, using a stylus for pointing [116, 130].

5.3 Proposed Model

Building upon existing research, we perform a theoretical analysis of peephole pointing, and develop a tentative model for movement time that incorporates *A*, *W* and *S*.

In general, a peephole pointing task consists of two stages: first *M*oving the *D*isplay window to reveal the target (noted as *MD*), and then *M*oving the *C*ursor to hit the target (noted as *MC*). Note that this distinction is made more in terms of the cognitive process rather than necessarily the actual dynamics of the input device movements. In the coupled cursor case, there is no physical distinction between moving the display window and moving the cursor. Even in the decoupled cursor case (e.g. using a stylus on a PDA), the user may choose to rest the stylus (and the hand holding it) on the screen when the screen is moving. In this case the cursor is roughly synchronized with the display window until the user sees the target and stops the display window. The user will then adjust the cursor within the display window. Cognitively, the stylus moves in the reference frame defined by the display window, rather than by the world. On the other hand, *MD* and *MC* may in practice also overlap in time, and thus may not have a clear boundary. Nevertheless, we feel this conceptual distinction is useful in framing our analysis.

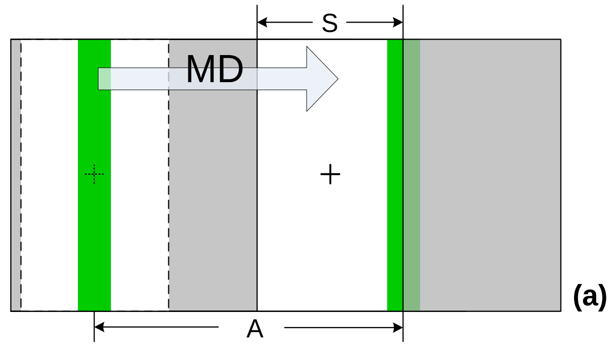
Depending on whether the user has prior knowledge of the target location, the performance of *MD* and *MC* can differ:

5.3.1 Without Prior Knowledge (Case 1)

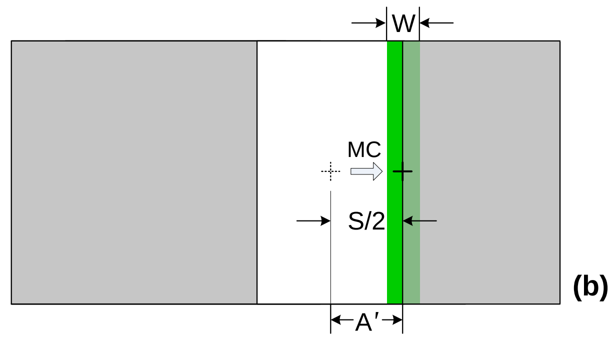
When the user does not know the target location, *MD* requires searching for the hidden target with the display window (Figure 5.3a). (Here we are assuming that the user knows which direction to search for, and leave the more complex case that involves uncertainty of the target direction for future investigations). Unfortunately, to the best of our knowledge, there is no existing model for this type of searching. (Prior research on visual search [128] mainly focused on identifying a target among several distracters, with both target and distracters visible all the time, which is fundamentally different from our task). However, intuition and observations during our pilot studies suggest that the time required for *MD* (T_{MD1}) should increase as A increases, since a longer distance requires more time to search through. T_{MD1} should also decrease as S increases because: (1) a larger display window reaches the target earlier, resulting in a shorter search distance; (2) a larger display window makes it easier to visually identify the target; (3) the user tends to move the display window faster when it is larger, with less concern of missing the target (observed in both pilot study and formal experiment). With these considerations, we posit that T_{MD1} plausibly follows a Fitts' law relationship with A and S:

$$T_{MD1} = a_{D1} + b_{D1} \log_2\left(\frac{A}{S} + c\right) \tag{2}$$

where a_{D1} and b_{D1} are constants that depend on the device and task setting, and c is 0, 0.5 or 1 (according to [70]). We neglect W in this equation as W does not impose a significant movement constraint in the *MD* stage, especially when W is considerably smaller than S.



(a) Moving the display window (*MD*).



(b) Moving the cursor (MC).

Figure 5.3: Peephole pointing without prior knowledge of the target location.

Once the target is revealed, the user moves the cursor to reach the target (*MC*) (Figure 5.3b). *MC* is a standard pointing task that can be modeled by Fitts' law, where the target width is W, and the effective target distance A' is the distance between the target and the current cursor position. In the ideal situation, *MC* starts once the target is revealed in the display window, and the cursor is residing at the center of the display window at the time (which is always true in the coupled cursor case), then A' is between $(S - W) / 2$ (i.e., target fully revealed) and $(S + W) / 2$ (i.e., target barely revealed), with an average of $S/2$.

However, in practice, due to limits of reaction time, the user will often overshoot the display window before *MC* starts. Since in the *MD* stage the user tends to move faster with a larger display window, this overshoot is likely to be larger with a larger S. Further, in the decoupled cursor case the cursor may have a positive or negative offset (with a mean of zero) from the center of the display window, depending on the previous action, the way the pointing device is held etc. Considering the somewhat unpredictable nature of these factors, we assume the mean value of the effective target distance as $\overline{A}' = k_1 S$, where k_1 is some constant coefficient (because of overshooting, k_1 is likely smaller than 1/2). Therefore the average time needed for *MC* is

$$T_{MC1} = a_{C1} + b_{C1} \log_2 \left(\frac{k_1 S}{W} + c \right) \tag{3}$$

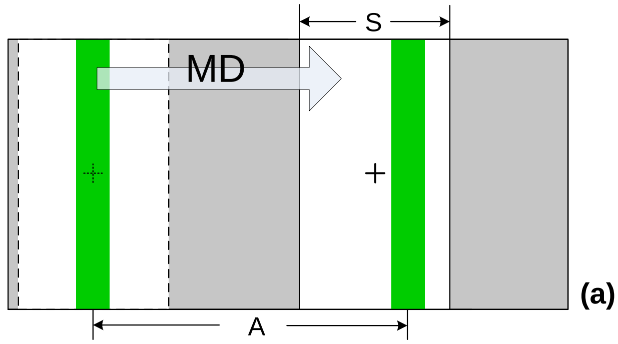
where a_{C1} and b_{C1} are the Fitts' law parameters for this setting, and c is 0, 0.5 or 1.

For the ease of mathematical derivation, we choose c to be 0 for both Eq. (2) and (3) (which is the equivalence of the original form of Fitts' law [39]), then total movement time is:

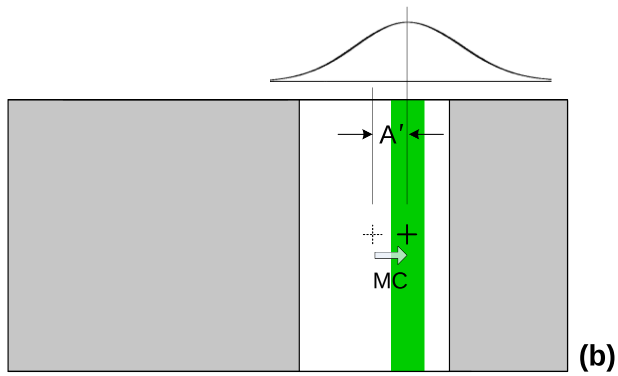
$$\begin{aligned} T_1 &= T_{MD1} + T_{MC1} \\ &= a_{D1} + b_{D1} \log_2 \left(\frac{A}{S} \right) + a_{C1} + b_{C1} \log_2 \left(\frac{k_1 S}{W} \right) \\ &= a_{D1} + \log_2 \left(\frac{A}{S} \right)^{b_{D1}} + a_{C1} + \log_2 \left(\frac{S}{W} \right)^{b_{C1}} + b_{C1} \log_2 k_1 \\ &= (a_{D1} + a_{C1} + b_{C1} \log_2 k_1) + \log_2 \left(A^{b_{D1}} S^{-(b_{D1} - b_{C1})} W^{-b_{C1}} \right) \end{aligned} \tag{4}$$

Because T_{MD1} describes searching without knowing the target location, which is more difficult than pointing to a visible target (described by T_{MC1}), we expect that $b_{D1} > b_{C1}$, resulting in a negative relationship between S and T_1.

5.3.2 With Prior Knowledge (Case 2)



(a) Moving the display window (MD).



(b) Moving the cursor (MC).

Figure 5.4: Peephole pointing with prior knowledge of the target location.

When the user already knows about the target location (such as from spatial memory of previous visits), *MD* is not truly a searching action, but a planned motion towards the target to cover it with the display window (Figure 5.4a). This is essentially the behavior of an area cursor, also modeled by Fitts' law as reported by Kabbash and Buxton [64], where *W* is set to be the width of the area cursor. Although they used a point target in their experiment, the relationship should still hold when the target width is considerably smaller than the area cursor. The only difference is that here the user does not see the target until it is revealed by the display window, and prior knowledge is used to guide the initial movement. Thus we have:

$$T_{MD2} = a_{D2} + b_{D2} \log_2\left(\frac{A}{S} + c\right) \tag{5}$$

Similar to case 1, *MC* is a typical pointing action with target width *W* and effective target distance A' being the distance between the target and the current cursor position (Figure 5.4b). However, here A' is caused by the spread (distribution) of display window positions resulting from *MD*. *MD* is a rapid approximate pointing task towards the target without initial visual cues, therefore has an open-loop (or "ballistic") characteristic to some extent. Several researchers [51, 78, 123] have investigated the spread of endpoints from such an action. They showed that (in the 1D case) the position of the endpoints X can be approximated by a normal distribution $N(X_0, \sigma)$, where the mean value X_0 is the center of the target, and the standard deviation $\sigma = fA$, f being an empirically determined constant. Therefore $A' = | X - X_0 |$, and from the property of normal distribution we know the mean value $\overline{A}' = \sigma\sqrt{2/\pi} = fA\sqrt{2/\pi} = k_2 A$, where $k_2 = f\sqrt{2/\pi}$ is a constant coefficient. (Again, in the decoupled cursor case, the offset between the cursor and the center of the display window is averaged out in the mean). Thus,

$$T_{MC2} = a_{C2} + b_{C2} \log_2\left(\frac{k_2 A}{W} + c\right) \tag{6}$$

Similarly, we choose $c = 0$ for both Eq. (5) and (6), and the total movement time

$$\begin{aligned}
T_2 &= T_{MD2} + T_{MC2} \\
&= a_{D2} + b_{D2} \log_2\left(\frac{A}{S}\right) + a_{C2} + b_{C2} \log_2\left(\frac{k_2 A}{W}\right) \\
&= a_{D2} + \log_2\left(\frac{A}{S}\right)^{b_{D2}} + a_{C2} + \log_2\left(\frac{A}{W}\right)^{b_{C2}} + b_{C2} \log_2 k_2 \\
&= (a_{D2} + a_{C2} + b_{C2} \log_2 k_2) + \log_2\left(A^{b_{D2}+b_{C2}} S^{-b_{D2}} W^{-b_{C2}}\right)
\end{aligned} \tag{7}$$

Summarizing case 1 and 2 (Eq. (4) and (7)), we have

$$T = a + \log_2\left(A^{b_A} S^{-b_S} W^{-b_W}\right) \tag{8}$$

where a, b_A, b_S, b_W are constants that depend on device and task property, and $b_A = b_S + b_W$ (b_A, b_S, $b_W > 0$). An alternate expression of Eq. (8) is

$$
\begin{aligned}
T &= a + \log_2\left(AS^{-b_S/b_A}W^{-b_W/b_A}\right)^{b_A} \\
&= a + \log_2\left(\left(\frac{A}{S}\right)^{b_S/b_A}\left(\frac{A}{W}\right)^{b_W/b_A}\right)^{b_A} \\
&= a + \log_2\left(\left(\frac{A}{S}\right)^{n}\left(\frac{A}{W}\right)^{1-n}\right)^{b} \qquad (let \quad n = b_S/b_A\,, \quad b = b_A) \\
&= a + b\left(n\log_2\left(\frac{A}{S}\right) + (1-n)\log_2\left(\frac{A}{W}\right)\right) \qquad (9)
\end{aligned}
$$

where a, b and n ($0 < n < 1$) are empirically determined constants that vary depending on the device and task property. Specifically, n describes the relative importance of S in terms of impact on movement time, compared to W.

Given the similarity between Eq. (9) and Fitts' law, we can define the Index of Difficulty for peephole pointing with a display window of size S as:

$$ID_S = n\log_2\left(\frac{A}{S}\right) + (1-n)\log_2\left(\frac{A}{W}\right) \qquad (10)$$

Inspired by the widely preferred variation of Fitts' law: $ID = \log_2(A/W+1)$, we can propose a variation of Eq. (10):

$$ID_S = n\log_2\left(\frac{A}{S}+1\right) + (1-n)\log_2\left(\frac{A}{W}+1\right) \qquad (11)$$

so as to guarantee ID_S is positive, especially when $S > A$, which is valid in practice. Eq. (11) also suggests that

$$\lim_{S\to\infty} ID_S = (1-n)\log_2\left(\frac{A}{W}+1\right) \qquad (12)$$

since the task becomes a standard pointing task when the display window is infinitely large, modeled by Fitts' law. Eq. (10) does not have this desirable property as it approaches $-\infty$ when $S\to\infty$.

Note that Eq. (8) may also be equivalently expressed as

$$T = a + b\left(n\log_2\left(\frac{A}{S}\right) + (1-n)\log_2\left(\frac{S}{W}\right)\right)$$
$$or \quad T = a + b\left(n\log_2\left(\frac{A}{W}\right) + (1-n)\log_2\left(\frac{W}{S}\right)\right) \tag{13}$$

with a slightly different definition of b and n ($0 < n < 1$), while keeping the signs of b_A, b_S, and b_W. Similar to Eq. (11), we may generate two other candidate variations:

$$ID_S = n\log_2\left(\frac{A}{S}+1\right) + (1-n)\log_2\left(\frac{S}{W}+1\right) \tag{14}$$

$$ID_S = n\log_2\left(\frac{A}{W}+1\right) + (1-n)\log_2\left(\frac{W}{S}+1\right) \tag{15}$$

Eq. (14) is reasonable given its similarity to our derivation of Eq. (4) in case 1. However, it does not regress to Fitts' law when $S\rightarrow\infty$. Eq. (15) satisfies this regression property, but does not have an intuitive interpretation. In addition, Eq. (11) yields $ID_S = 0$ when $A = 0$, which is intuitive since the required movement is completely eliminated. Neither Eq. (14) nor Eq. (15) satisfies this requirement.

Nevertheless, we will use our experiment data to verify each variation of ID_S (Eq. (10), (11), (14), (15)) and determine which model best fits the data.

5.4 Experiment Setup

5.4.1 Apparatus

We chose to simulate a flashlight/peephole display on a desktop monitor (19-inch, 1600 x 1200 pixels) (Figure 5.5). The entire screen represents the workspace, with only the region corresponding to the display window revealed, and the rest (virtually) masked in black. A Wacom™ Intuos2 tablet (12 x 18 inch) is used for input. The decision to simulate the flashlight/peephole display on a desktop computer rather than using a spatially aware handheld device was to ensure the most reliable and highest quality possible input and output devices, and to enable the abstract 1D display window in our experiment. This prevents our results from being confounded by tracking errors and other performance limitations of current spatially aware handheld displays. Nevertheless, based on this current work, testing our model under various realistic situations would be desirable to further generalize our findings.

(a) Coupled Cursor.

(b) Decoupled Cursor.

Figure 5.5: Experiment Setup.

5.4.2 Techniques

Two techniques are investigated:

Coupled cursor (Figure 5.5a): a stylus is used to control the position of both the display window and the cursor when held over the tablet. The cursor, displayed as a crosshair, always stays at the center of the display window.

Decoupled cursor (Figure 5.5b): a Wacom™ puck held in the non-dominant hand can be moved on the tablet to control the position of the display window. The dominant hand holds the stylus used to control the cursor position, independent of the display window. Although the user is free to move the cursor outside the display window and still able to see it, the system only responds to the stylus tap (for selecting the target) when the cursor is inside. To simulate the real world scenario of using a stylus on a moving PDA (as in

[130]), a thin board is glued to the bottom of the puck and moves with it. The user can tap the stylus on the board, and rest the stylus hand on it while moving the puck.

5.4.3 Task

A reciprocal 1D pointing task is used, in which both the display window and the targets are one-dimensional, vertically extending to the border of the screen. In each sequence of trials, the participant selects two equal-width targets back and forth in succession. The distance between the targets is set according to the experiment condition, while the exact positions of them are randomized for each sequence, so as to prevent the participant from guessing the target position. Within the same sequence the target positions remain unchanged across the trials. At the beginning (first trial) of each sequence, the position of the first target is displayed as a red X, in order to guide the participant to move the display window toward it and reveals the first target. However, before the display window actually moves over the red X, the display window is rendered only as a wireframe silhouette without revealing the content inside. This is to prevent the participant from seeing the second target if the display window happens to pass it. Once the display window reaches the red X, it reverts to its normal appearance, and the first target is revealed. The participants can then use the cursor to select it (by tapping the stylus), and immediately starts to search for (if not already inside the display window) the second target. The position of the first target provides a cue as to which direction to search. For example, if the first target lies in the left half of the screen, the second target will always be to its right (but not necessarily in the right half of the screen), and vice versa.

After selecting the second target, the participant moves back to select the first one, and so forth. Note that unlike a conventional reciprocal pointing task, where all trials in a sequence are essentially equivalent, in our task the first trial is fundamentally different from the others in the sequence. The first trial (i.e. the first pointing action from the first target towards the second target) represents pointing *without prior knowledge* (apart from the directional cue discussed above) and requires extensive searching, but in all following trials the participant has already gained knowledge about the target location from previous selections. This provides us a way to cover both conditions of *prior knowledge of target location* in a single sequence.

The participant must successfully select the target before s/he can proceed to the next trial. Once the target is selected, its color turns from green to gray, and a short beep is played to indicate success.

Because we are only interested in the movement along the horizontal axis, we constrained the cursor to be always vertically centered on the screen. The input area of the tablet is cropped to provide a 1:1 control-display gain. All buttons on the puck and the stylus are disabled.

We measured the completion time (including possible error correction time), and the number of errors made in each trial.

5.4.4 Design

A fully crossed within-participant factorial design was used. Independent variables were target distance *A* (128, 256, 512, 1024 pixels), target width *W* (8, 16, 32, 64 pixels), display window size *S* (32, 64, 128, 256, 512 pixels), c*ursor control mechanism* (coupled, decoupled), and *prior knowledge of target location* (yes, no). The range of *A*, *W* and *S* values covered both typical and extreme (e.g. $S < W$) scenarios.

Each participant used both cursor control mechanisms, within which, three consecutive blocks were performed. Within every block, each combination of *A*, *W* and *S* were tested for one sequence of trials. The presentation order of these combinations was randomized within each block. We chose to have only 2 trials in each sequence, in which the first trial (i.e., pointing from the first target towards the second target) represents pointing *without prior knowledge*, and the second trial (i.e., returning from the second target to the first target) represents pointing *with prior knowledge*, so that we have equal numbers of data points for both conditions. We evenly randomized the movement direction (left or right) in the first trial (the second trial has the reverse direction), in order to prevent the movement direction from confounding with the *prior knowledge* factor. An input error was counted when the participant taps outside the target. Participants were asked to keep the error rate under 4% during each block. 20 practice trials were performed before each cursor control condition started.

As a baseline, each participant also performed a standard pointing task, with the entire workspace always visible (i.e. the display window is infinitely large). The participant used the stylus to control the cursor. A similar experiment structure was used for the baseline task, keeping all factors but the display window size S. Each participant performed the peephole pointing task using two different cursor control mechanisms plus the baseline task, with the order of the three counterbalanced across participants.

5.4.5 Participants

12 right-handed volunteers, aged 18–33, participated. The experiment lasted about 1.5 hours for each participant. Participants were encouraged to take breaks between blocks. Participants did not receive monetary compensation.

5.5 Experiment Results

5.5.1 Movement Time

For the peephole pointing task, the overall average movement time T is 2118 ms. T increases monotonically as A increases, and decreases monotonically as W or S increases, as suggested by our proposed model. An analysis of variance showed that all these effects are statistically significant (p<.001). Pair-wise means comparisons also showed that every condition of A, W, or S is significantly different from the others (p<.001).

There was a significant main effect for *cursor control mechanism* ($F_{1,11}$ = 55.4, p<.001), with mean times of 1884 ms for *coupled cursor* and 2351 ms for *decoupled cursor*. Different *cursor control mechanisms* result in different behaviors, and in turn different performance. However, the trend that *coupled cursor* performs faster than *decoupled cursor* should not be regarded as a general conclusion, since it may depend on the specific device used. For example, we might expect *coupled cursor* to become slower if we used the non-dominant hand puck to control the cursor rather than the dominant hand stylus.

There was also a significant main effect for *prior knowledge of target position* ($F_{1,11}$ = 491, p<.001),with the first trial (*without prior knowledge*) of each sequence averaging 2562 ms, while the second trial (*with prior knowledge*) averaging 1674 ms. As a comparison, for the baseline task, there is no significant difference ($F_{1,11}$ = 2.14, p = .171) between the two trials. Therefore we conclude that the performance difference between the two trials in the peephole pointing task is indeed caused by the knowledge of target position, rather than a practice effect.

There were significant interactions for *cursor control mechanism* x *prior knowledge* ($F_{1,11}$ = 4.83, p = .050), S x *cursor control mechanism* ($F_{4,44}$ = 32.8, p < .001), and S x *prior knowledge* ($F_{4,44}$ = 89.5, p<.001). In particular, S has a larger effect on T when there is no prior knowledge of the target location (Figure 5.6). This can be explained that without prior knowledge, MD is a pure searching task that relies heavily on the visual feedback in the display window, while with prior knowledge MD has a more ballistic nature, for which the display window plays a less important role. However, the effect of *prior knowledge* decreases as S increases. When S is small, only a small amount of information is revealed in the display window, so any prior knowledge of the target location would add considerable information and significantly benefit the task performance. As S becomes larger, the additional information gained from prior knowledge becomes smaller, resulting in less performance gain. Eventually when *S* approaches infinity, the task becomes a standard pointing action, and the effect of prior knowledge is eliminated, as shown by the lack of significant difference between trials in the baseline task.

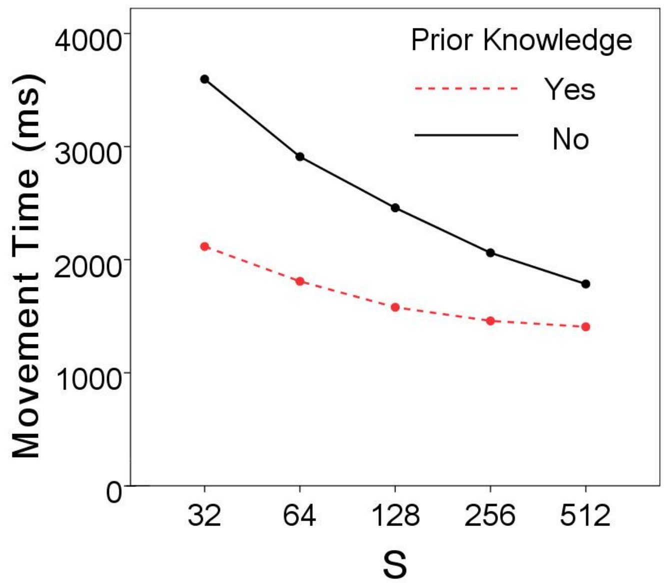


Figure 5.6: Movement time by display window size and prior knowledge.

There was a significant main effect for block number on *T* ($F_{2,22}$ = 18.7, p<.001), indicating learning between blocks.

5.5.2 Errors

An error occurred when the user tapped outside the intended target. The overall average error rate was 3.5%. The factors that significantly affected the error rate include: *cursor control mechanism* ($F_{1,11}$ = 7.50, p=.019) with 2.9% errors for *coupled cursor* and 4.1% errors for *decoupled cursor*, possibly caused by the coordination skills required for bimanual input; *W* ($F_{3,33}$ = 27.3, p<.001), where the error rate increases as *W* decreases, because of the increased task difficulty; and *S* ($F_{4,44}$ = 2.65, p=.046). The error rate is lowest (2.4%) when *S* = 32, largest (4.4%) when *S* = 512, and roughly 3.5% for all other conditions between. This is an interesting and somewhat counterintuitive phenomenon, which may be because users became over-careful with a tiny display window, and over-relaxed with a huge display window. No significant interaction exists between any factors for error rate.

ID_S	Cursor Control Mechanism	Prior Knowledge	a (ms)		b (ms/bit)		n		R^2
			Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	
$n\log_2\left(\frac{A}{S}\right)+(1-n)\log_2\left(\frac{A}{W}\right)$ (Eq. 10)	Coupled	Yes	444	46.9	316	12.1	0.304	0.0302	0.900
		No	561	115	652	29.5	0.533	0.0363	0.868
	Decoupled	Yes	615	65.8	437	16.9	0.441	0.0306	0.897
		No	782	139	796	35.6	0.575	0.0363	0.873
$n\log_2\left(\frac{A}{S}+1\right)+(1-n)\log_2\left(\frac{A}{W}+1\right)$ (Eq. 11)	Coupled	Yes	171	46.2	382	13.2	0.376	0.0281	0.923
		No	-168	108	843	31.0	0.606	0.0287	0.906
	Decoupled	Yes	175	63.4	548	18.1	0.515	0.0260	0.924
		No	-143	128	1040	36.7	0.646	0.0275	0.913
$n\log_2\left(\frac{A}{S}+1\right)+(1-n)\log_2\left(\frac{S}{W}+1\right)$ (Eq. 14)	Coupled	Yes	-244	67.3	707	26.3	0.625	0.00974	0.913
		No	-697	160	1276	62.6	0.720	0.0150	0.891
	Decoupled	Yes	-259	96.9	899	37.8	0.679	0.0120	0.905
		No	-737	187.3	1524	73.1	0.738	0.0152	0.902
$n\log_2\left(\frac{A}{W}+1\right)+(1-n)\log_2\left(\frac{W}{S}+1\right)$ (Eq. 15)	Coupled	Yes	-84	65.8	743	54.2	0.454	0.0268	0.897
		No	-942	205	1948	169	0.341	0.0241	0.780
	Decoupled	Yes	-276	112	1189	92.1	0.380	0.0238	0.845
		No	-1176	255	2514	210	0.321	0.0221	0.775

Figure 5.7: Summary of model fitting results.

5.5.3 Model Fitting

We fit the movement time T to the four candidate formulations (Eq. (10), (11), (14), (15)) of ID_S using least-square regression. Since both *cursor control mechanism* and *prior knowledge of target location* affect T, we break down the data into four categories according to these two factors, and fit them separately. Figure 5.7 shows the parameter estimates and standard errors for those estimates. The last column provides the R^2 values for the regression.

Eq. (11) has the best fit with experiment data in all 4 categories, with R^2 always greater than 0.9 (Figure 5.8). In addition, Eq. (11) has several desirable properties as discussed earlier. Therefore our model for peephole pointing is best represented as:

$$T = a + b\left(n\log_2\left(\frac{A}{S}+1\right)+(1-n)\log_2\left(\frac{A}{W}+1\right)\right) \qquad (16)$$

where *a*, *b* and *n* ($0 < n < 1$) are empirically determined constants that may vary depending on the device and task.

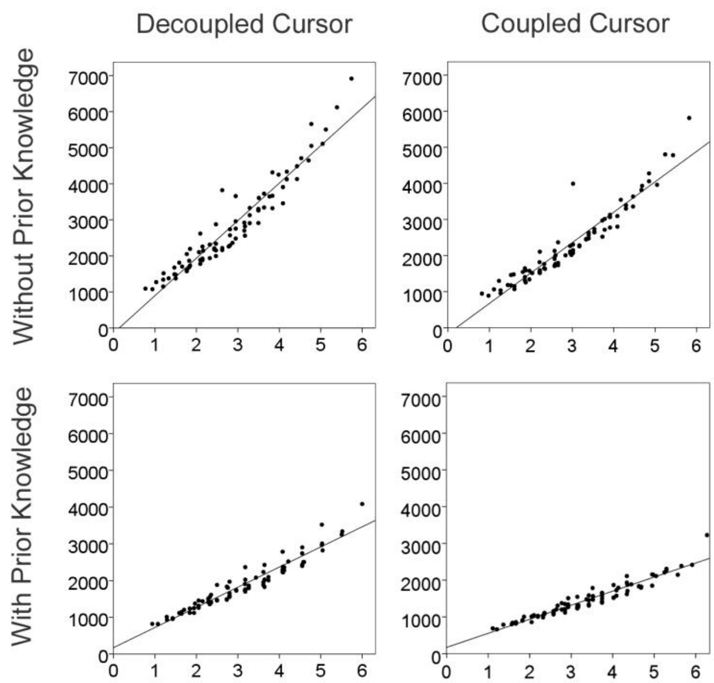


Figure 5.8: Model Fitting of Eq. (11).

(horizontal axis: ID_S ; vertical axis: movement time (ms))

As a baseline, we also fit the data to the three variations of standard Fitts' law: $ID = \log_2(A/W+1)$, $ID = \log_2(A/S+1)$, and $ID = \log_2(S/W+1)$, which all yield average R^2 under 0.66. This further confirmed the advantage of our proposed model which yielded significantly better fits.

In particular, *n* describes the relative importance of display window size *S* in terms of impact on the movement time. Compared to *a* and *b*, which reflect mainly the performance of the device or user group, *n* captures more of the essence of the task itself. As an example, Figure 5.9 compares the *n* values between the four categories we fit to.

n	Coupled Cursor	Decoupled Cursor
With Prior Knowledge	0.376	0.515
Without Prior Knowledge	0.606	0.646

Figure 5.9: *n* values under different conditions.

For both *cursor control mechanisms*, n is smaller when there is prior knowledge about the target location, indicating a lesser impact of S. This is validated by the data trend we analyzed in the previous section. However, n is larger in the *decoupled cursor* condition than in the *coupled cursor* condition. The fact that the stylus needs to be inside the display window to be responsive imposes an additional motor constraint related to S, while this constraint does not exist in the coupled cursor case. This may explain the larger impact of S in the *decoupled cursor* condition. Further, the difference of n caused by *cursor control mechanism* is smaller *without prior knowledge*, since the searching behavior that accounts for the majority of the task effort is mainly a cognitive process, relatively independent of the control mechanism. When prior knowledge becomes available, the cognitive aspect of the task is reduced and the motor control aspect dominates, which is more vulnerable to influences of the control mechanism.

In addition, Eq. (16) shows that when S becomes sufficiently large (e.g. when $S > 2A$, the target is inside the initial display window), the benefit of further enlarging it becomes marginal. For example, enlarging S from $2A$ to $4A$ only results in ID_S decreasing by $0.26n$, versus $0.74n$ when S changes from $A/4$ to $A/2$. This can also be observed from the trend in Figure 5.6: the relationship curve becomes flatter as S increases, and will eventually converge as $S \to \infty$, at which point the action becomes a standard pointing task. Our model suggests that the movement time follows Fitts' law as $S \to \infty$, validated by the unsurprising good fit of the baseline task performance to Fitts' law ($R^2 = 0.976$).

5.6 Discussion and Implications

One interesting observation about our model is its mathematical similarity to the very first model for 2D pointing tasks, suggested by Crossman [32]: $T = a + b\log_2\left(\frac{A}{W}+1\right) + c\log_2\left(\frac{A}{H}+1\right)$ (W and H are target width and height respectively), which can be reformulated to resemble Eq. (16). This model suggests that W and H affect T independently. However, better models for 2D pointing tasks have been developed later on [2, 54, 78], which more properly address the strong interaction found between W and H. In our task, the somewhat separable stages of MD and MC make it more appropriate to tackle S and W separately in the model, as supported by its good match with experimental data.

As we discussed, the parameter n captures in large the nature of a given peephole pointing task, and is less influenced by the device used. Thus to analyze behavior under a particular setting, we may first acquire the n value (by theoretical analysis or empirical study), and make design decisions according to it. In general, a larger n means we need to focus on increasing the display size (or virtual size, such as using a Fisheye [47] visualization); while a smaller n suggests we should endeavor to enlarge the interface components (or enlarge their effective width, such as by dynamically expanding the targets as the cursor approaches them [77]).

In our experiment, we dealt with cases that the user has either no prior knowledge of the target location, or near-perfect knowledge of it (all participants expressed they could remember very well the target location during the second trials). However in practice there are many cases that lie between these two extremes. For example, the user may obtain partial (or "soft") knowledge through visualizations of off-screen targets. We expect the resulting performance also lies somewhere between the two cases we studied. How exactly the performance varies with the amount (and format) of the information available is worth further investigation.

During the experiment we observed different strategies employed by participants to search for the targets. Some people moved the display window at a more or less constant speed within each trial, until they saw the target and stopped the display window. Some others first moved the display window in a quick movement, hoping to get a glance of the target, and then backtrack to it. This strategy might provide some benefit with large display windows, but was not always successful when the display window was small. In the latter case, the participant might have to revert to the first strategy after an initial failure to find the target, resulting in performance loss. Analysis of the movement logs indicates 32.1% of the trials involved some amount of backtracking of the display window, including cases of either the intentional quick movement strategy, or an unintentional overshooting. The number of backtrackings decreases as S becomes larger. Regardless of the variety of the strategies, our model is general enough to account for the data. However, it is worthwhile to consider how one might accommodate these strategies when designing flashlight/peephole interfaces, for example making the interface components more “glanceable”.

It is worthwhile to note that the *decoupled cursor* mechanism is a typical example of asymmetric bimanual action [52], where the non-dominant hand performs coarse movements and sets the reference frame, within which the dominant hand performs fine movements and operations. Analysis of the movement logs revealed that the cursor resides inside the display window for 88.0% of the movement time on average. The instant movement velocity of the cursor (calculated every 0.1 sec, relative to the display window) has a mean of 65.7 pixels/sec versus 223.9 pixels/sec for the display window (absolute). These both conform to the asymmetric work division pattern between the non-dominant and dominant hands.

The peephole pointing model is a first step towards a comprehensive understanding of the Flashlight metaphor’s effect on user performance of fundamental interaction tasks. Many obvious further investigations could be conducted in the same direction, such as for crossing, steering, or 2D pointing tasks. Validation of the models on various devices is also important. However, guided by the overall goal of the thesis, that is to explore and set the groundwork of a board spectrum of research aspects of handheld projector interaction, we chose to leave further pursuit in this direction for future explorations (discussed in Chapter 8). On the other hand, we have so far investigated interactions with a single Flashlight (i.e. a single handheld projector). Given the envisioned prevalence of handheld projectors in the future, as well as the sharable viewing experience created by them, it is natural to consider how multiple handheld projectors can be used simultaneously and interact with each other. Supported by this, multiple co-located users could enjoy more powerful viewing experience, easily exchange information with each other, and engage in collaborative and social activities. Thus, in the next chapter we explore interaction techniques and usage scenarios using multiple handheld projectors.

Chapter 6

Interaction using Multiple Handheld Projectors

The large-sized displays generated by handheld projectors naturally afford multi-person viewing and thus have the potential to support co-located collaboration. In particular, when each user has a handheld projector, the interactivity between projectors can result in a rich design space for multi-user interaction. Building upon our explorations on single projector interactions in Chapter 4, we sought to explore how multiple handheld projectors are to be used together to support interaction between users.

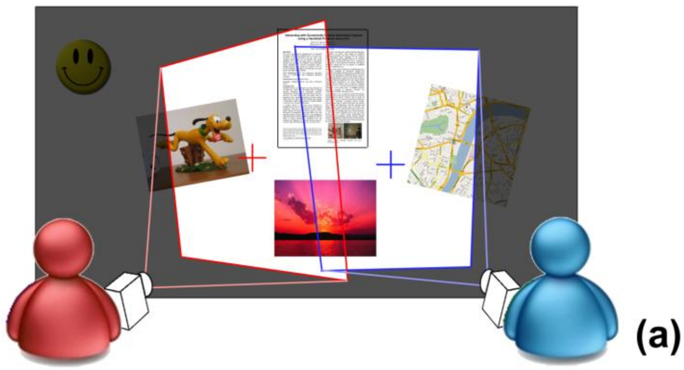
Although many current handheld or portable devices have the ability to exchange data with other devices via wireless connections, the interaction required to facilitate such exchange often requires cumbersome and explicit authentication procedures. Although such procedures are generally unavoidable when devices (and their users) are not physically co-located, they may be unnecessary if we can design interaction that exploits the user and device co-locality to facilitate connectivity and collaboration. Researchers have explored co-located collaboration between people using shared displays on tabletops [104, 126, 129] and walls [22, 54, 62, 106]. Because the workspace is shared between all users, information exchange and multi-user operations can be easily realized. However, these shared displays are not portable for ubiquitous use, and every user shares the same view of the workspace. Private and personalized information are not easily accommodated, and global conflicts [80] may occur, in which one user's action affects the entire shared display and disrupts other users.

In contrast, the use of multiple handheld projectors may open up a novel interaction paradigm for co-located users, in which they can share the same physical display and interaction space, while at the same time individually creating and controlling parts of the overall virtual display with their own projectors. In this chapter, we explore the design space of multi-user interaction using multiple handheld projectors. Expanding from the single-projector interaction techniques we described in Chapter 4, we develop an additional set of interaction concepts and techniques to specifically suit multiple users working in a shared physical space, each using their own projector. These designs could enable a variety of multi-user usage scenarios as illustrated in our explorations.

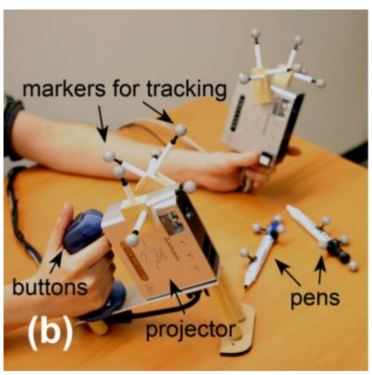
6.1 Multi-Projector Prototype Overview

Extending the single handheld projector prototype described in Chapter 3, our multi-projector prototype uses two Mitsubishi™ PK10 Pocket Projectors (Figure 6.1b). Similarly, each projector is augmented with two buttons for input (primary button for selection and operations, and secondary button for triggering menus), and can be easily handled and moved using one hand. Two passive pens are included for writing on surfaces. Both the projectors and the pens are tracked by the same Vicon motion tracking system, providing 6-dof (position + pose) information at millimeter precision. Both projectors are connected to a same 2.4GHz P4 PC, which produces the images and handles the interaction.

Again, the flashlight metaphor is used as the basis for the interactions (Figure 6.1a). The image projected by each projector reveals a portion of a large workspace stationary on the projection surface. When the projector is moved, the projected image content changes accordingly, as if the projector is used as a flashlight to explore in darkness. Multiple workspaces can be associated to different physical surfaces in an environment, such as walls, tables, and bulletin boards. The workspaces are shared among the projectors. Different projectors may reveal different (or overlapping) regions of a workspace simultaneously. In addition, each projector's view of the workspace may also be personalized depending on which user is using it.



(a) System concept.



(b) Handheld projectors. (c) System in use.

Figure 6.1: Multi-projector prototype.

Tracking and calibration inaccuracies may result in slightly imperfect image alignment in overlapping projection areas. To avoid unpleasant double images, we provide the option to blank out one projector in the overlapping area, and let the other projector handle the display for both. We expect that advances in computer vision and calibration techniques will solve this problem in the long run.

Two users, each having their own projectors and pens, can interact with the system simultaneously. The system architecture is also scalable to support three or more projectors/users. Each user is identified by a unique color, which is reflected by the cursor displayed by the projector, and the marks drawn by the pen. With two buttons, each user can independently manipulate virtual objects using the cursor, and trigger commands using crossing-based widgets, as described in Chapter 4. In this chapter, we explore techniques to support interactions that involve multiple users. Similar to the single projector techniques described in Chapter 4, these multi-user interactions are designed as a generic vocabulary to demonstrate the capabilities of multiple handheld projectors. How these interactions will be combined, switched between, and semantically interpreted will depend on the particular applications that utilize them, which are beyond the focus of this chapter.

6.2 Interaction Concepts and Techniques

6.2.1 Ownership & Access Control

Each object in the workspaces may either have no ownership (accessible by all users), or be owned by a particular user. In the latter case, the owner of the object has full control over it. How other users can interact with it is determined by its access level, which is one of the following:

Public: The object is visible to all users (i.e. all projectors will display it), and all users can operate on it. Any object without an owner is implicitly public.

Semi-Public: The object is visible to all users, but only operable by its owner.

Private: The object is visible and operable only by its owner. It is not displayed in other users' projectors.

The ownership and access level of an object is indicated by the flags on the top right and top left corners of the object, respectively (Figure 6.2). The color of the ownership flag matches that of the owner, and the color of the access flag indicates the access level: green for public, yellow for semi-public, and red for private. We choose to use colors to indicate both the ownership and the access level for the sake of simplicity, which works well for a small number of users. Should we need to scale the system to include a larger number of people, we may employ other ways to differentiate the owners such as textual IDs to avoid confusion. The owner can cycle through the access flag levels by crossing it from outside the object to inside while holding the primary button down. Note that the term “visible” used within this access level context is not to be interpreted in the strictest sense, since other users can still peek at the object when it is being viewed by its owner. To completely hide the object’s content, the owner can toggle the visibility flag by crossing it. The object will then be shown as a blank frame in its owner’s view, and invisible to other projectors. Hiding an object’s content using the visibility flag in this manner implies setting its access level to private.

Figure 6.2: Object with flags.

6.2.2 Information Exchange

Exchanging information between users is a common task in multi-user interaction. Compared with current handheld devices, which rely on indirect procedures, in our system all information exchanges can be achieved by direct manipulation using several techniques each suited to different situations:

Passing Ownership

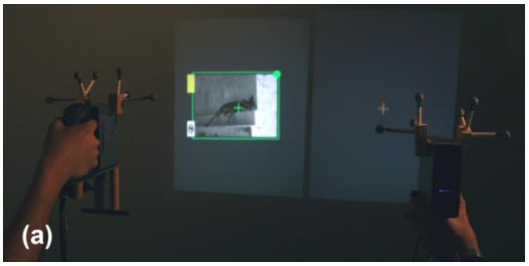
Users can pass the ownership of their objects to others, much like handing over a physical object in the real world. The receiving user can then operate on the object freely, or drag it into their personal folder (a container that stores all personal files) for later use. The ownership passing action is completed by the following steps:

Step 1: User A (owner) captures the object by clicking on it and holding the primary button down. (Figure 6.3a).

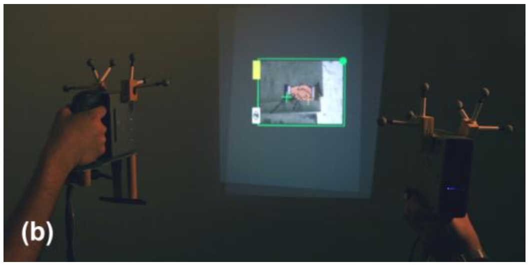
Step 2: User B clicks on the object and holds down the primary button.

Step 3: Both users dwell their cursors. (Figure 6.3b).

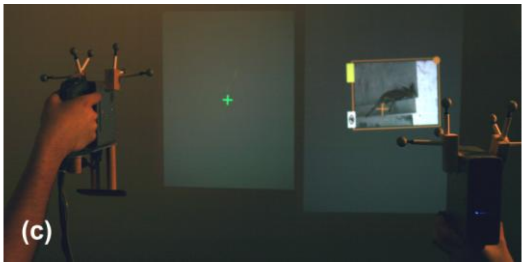
Step 4: After a brief period (≈ 1sec), User B "captures" the object and becomes its owner (Figure 6.3c).



(a) Before.



(b) During.



(c) After.

Figure 6.3: Passing object ownership.

While dwelling is often undesirable in interactive systems, it serves an important purpose here in that the dwelling required of both users ensures quasi-explicit consent from both parties to perform the action. A handshake icon starts to fade in at Step 3, and reaches its full opacity by Step 4, giving users an indication of the upcoming passing action. During this stage, User A can either move away or release the object to prevent undesirable or unintentional passing.

Dropping into Personal Folder/Portal

For more efficient information exchange, objects can be directly dragged and dropped into one's personal folder. The dropping action is regulated by the access levels of both the object and the folder. Users can drop anything that they have operation access to into their own folder. In order to drop objects into other users' folders, either the folder needs to be set as public, or the folder's owner needs to hold down the primary button over the folder while the object is dropped into it, analogous to the real world action of holding a bag open to allow others to put things in it.

To provide access to folders far away, or to protect privacy of a folder's content, users can create a portal to their folder (Figure 6.4). Ownership is indicated by the portal's color. Anything dropped into the portal will be transported into the associated folder. The portal follows the same access control policies as the personal folder.

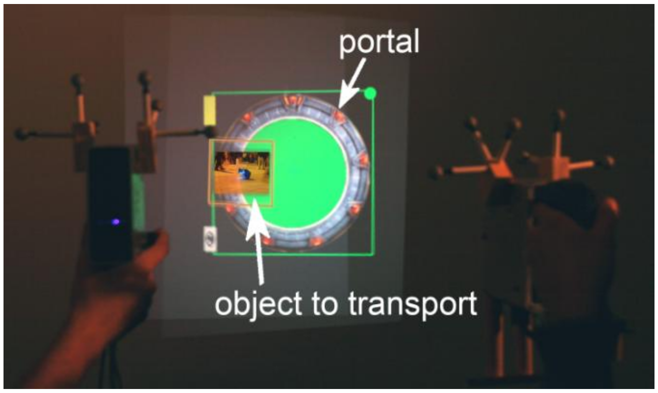


Figure 6.4: Dropping objects into a portal.

6.2.3 Compositing Projections

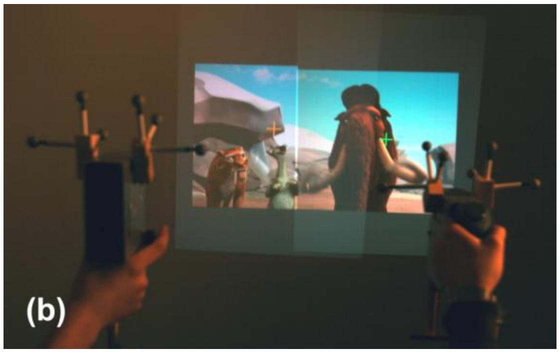
When multiple handheld projectors are available, their displays can be composited to improve the viewing experience, beyond what is possible with a single projector.

Expanding the Display Area

As Raskar et al. [91] suggested, multiple projections can be aligned side by side to create a larger display area than a single projection without sacrificing image resolution. This is clearly helpful for viewing a large document or map.

In addition, our system can also intelligently adapt the view of the object to exploit the enlarged display area provided by multiple projectors. For example, when watching a movie, a cropped version of the movie is displayed when viewed with one projector, but seamlessly switches to a widescreen version when two projection displays are aligned horizontally to accommodate it. (Figure 6.5)

(a) Cropped view.



(b) Widescreen view.

Figure 6.5: Viewing a movie.

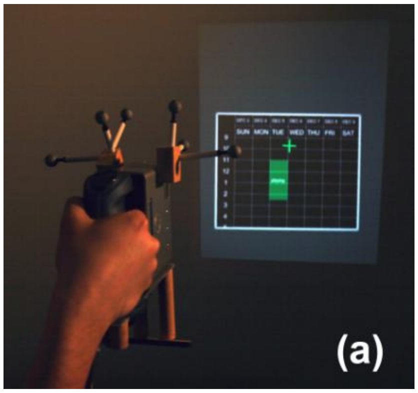
Different projectors can also point at different regions of the workspaces, thus creating multiple viewing/operating areas, especially on different projection surfaces. Similar to Hinckley et al.'s [55] example using multiple tablet displays, one user can click on thumbnails of photos in an album projected on a surface convenient for operation (e.g. a table), while another user projects the full view of the selected photo on a larger surface (e.g. a wall). (Figure 6.6) Other applications include exploring related parts of a large graph, or transporting information between places.

Figure 6.6: Photo album browsing using two displays.

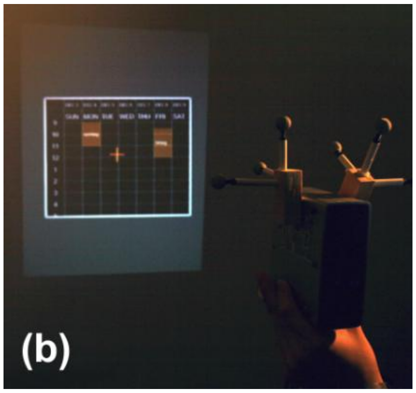
Accommodating and Combining Multiple Views

One unique characteristic of handheld projectors is that each user is creating their own display while sharing the same physical and information space. This fluidly enables different views of the same object when displayed by different projectors, and allow users to see personalized information relevant to themselves. For example, a calendar shows appointments of the user who is projecting it as colored blocks (Figure 6.7a, b), and a photo frame shows a photo of the user who is projecting it (Figure 6.9a, b).

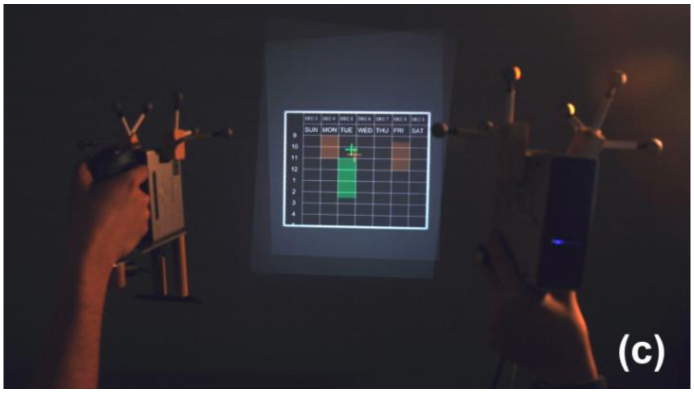
When multiple projections overlap on the same object, the different views (if applicable) are seamlessly blended by the optical overlaying of projection images. The result is a combined view that is relevant to all projecting users. For example, a calendar displayed by two overlaying projectors shows events for both users, and the empty spaces in it are timeslots available for both people to have a meeting (Figure 6.7c). This provides an intuitive and efficient way for scheduling meetings. In order to maintain privacy, text labels describing the events are hidden when the calendar is viewed by multiple users.



(a) Viewed by User A.



(b) Viewed by User B.



(c) Viewed by both users.

Figure 6.7: Multi-view calendar.

As described in Chapter 4, handheld projectors can support image resolution gradation and multiple information granularities depending on the distance between the projector and the projection surface. As the user comes closer to the surface, the projection area shrinks and a higher pixel density is achieved in the area, resulting in higher local resolution that can be used to display more detailed information. Utilizing this feature, multiple projectors can be combined to create a viewing experience similar to that of a focus plus context display [10]. One projector can be held afar to create the low-resolution coarse-granularity surrounding context in a larger area, and another projector is used close to the surface creating a focus region to explore high-resolution fine-granularity details within that context. Because the projection image also becomes brighter when the projector is nearer, the projection of the focus region automatically overlays the context information beneath it. Figure 6.8 shows this with a multi-granularity city map. The context region shows main streets only, while the focus region shows all small streets. Compared with previous focus plus context screens [10], where the resolution and position of both focus and context displays are fixed, our solution is more flexible in that users can dynamically move, resize and change the resolution of both projections. We may also achieve nested focus regions by overlaying three or more projectors.

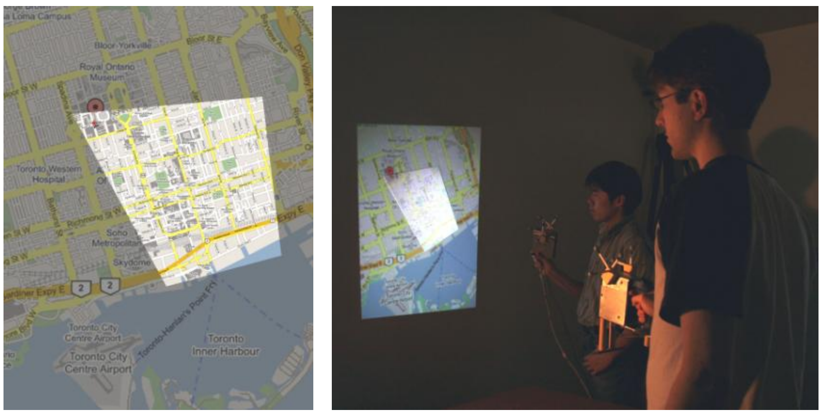
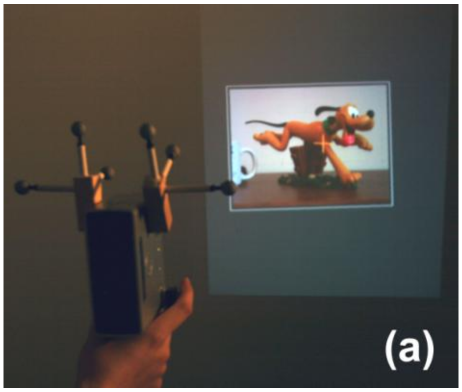


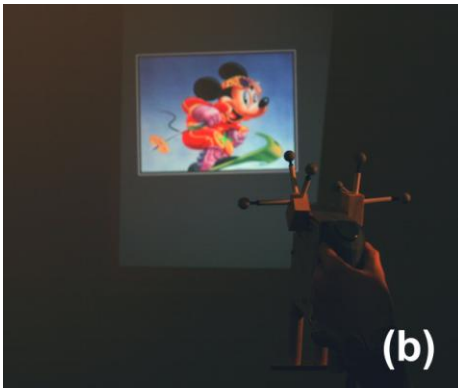
Figure 6.8: Focus plus context display.

Direct blending of multiple views may not always be sufficient. The system can also render a semantic combination of different views when it detects multiple projectors overlaying on an object. For example, the photo frame shows a group photo of both users when two projectors overlap on it (Figure 6.9c). Alternatively, some critical information may only be revealed when multiple users look at them at the same time. For example, any single user can only see the cover page of a group assignment. Only when two or more users overlap their projections on it, can they read the content of the assignment. An extreme case of this is an object that is completely hidden in any single projector's view, but become visible when projections overlap.

Similar to our focus plus context usage, semantically combined views can also be useful when one projection is contained in another. A different view is rendered in the overlapping region only, as if using a magic lens [19]. Figure 6.10 shows a user exploring the inner structure of a car model using one projector like a magic lens.



(a) Viewed by User A.



(b) Viewed by User B.

(c) Viewed by both users.

Figure 6.9: Multi-view photo frame.

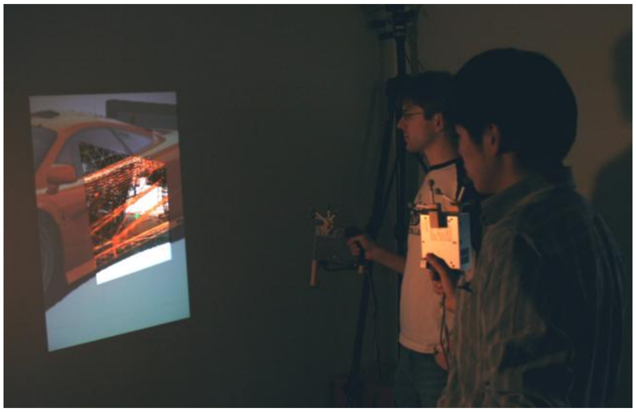
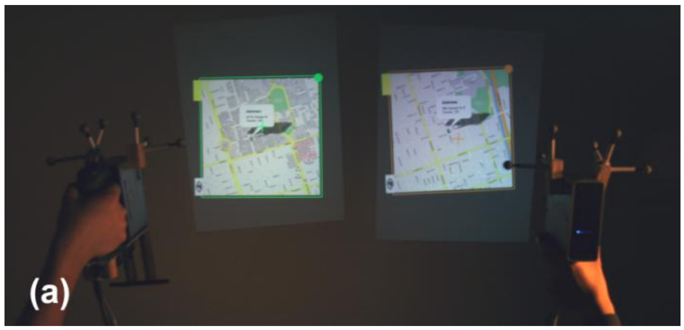


Figure 6.10: Emulating a magic lens with semantically combined multiple projections.

6.2.4 Linkage between Objects

Two users can create linkages between their objects (one from each user) for information or operations that involve both (if applicable). Depending on their needs, two types of linkages can be created.



(a) Separate.

(b) Snapped.

Figure 6.11: Snapping.

Snapping

Snapping provides a lightweight transient way to quickly view information that involves two objects. If two objects are compatible for snapping, then when they are moved close enough, they will snap to each other side by side. To unsnap them, users simply drag either or both of the objects in any direction past a small distance. When snapped together, either or both the objects will change their appearance to reflect information that relates to its partner. For example, when two users snap maps of their home addresses together, the maps change to show the directions between the two addresses, both in drawing and in text (Figure 6.11). When a clock is snapped to a city map, it changes its time zone to reflect the local time of the city.

Docking

To perform operations that affect (possibly modify) multiple objects, two users can create a more explicit linkage between their objects by docking them together. Compared with snapping, which can be initiated by one user, docking is a more strict action that requires consent from both users to prevent operations not authorized by the object owner. In addition, only two objects of the same kind (two documents, two calendars, etc.) can be docked.

The docking action is completed by the following steps:

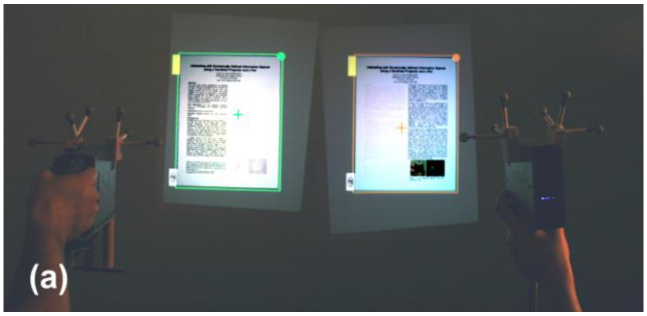
Step 1: Each user capture his/her own object by clicking on it and holding down the primary button (Figure 6.12a).

Step 2: Both users move their objects to roughly overlay them together, and dwell the cursors. (Figure 6.12b).

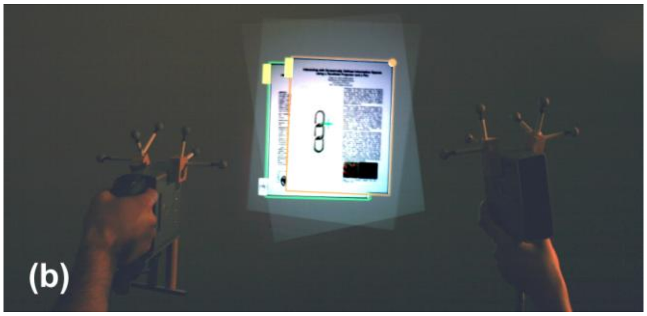
Step 3: After a brief period (≈ 1sec), both objects become docked together (Figure 6.12c).

Similar to the ownership passing action, the dwelling ensures consent from both parties. A linkage icon starts to fade in at Step 2, and reaches its full opacity by Step 3. While the icon is fading, either user can move away or release the object to prevent unwanted docking.

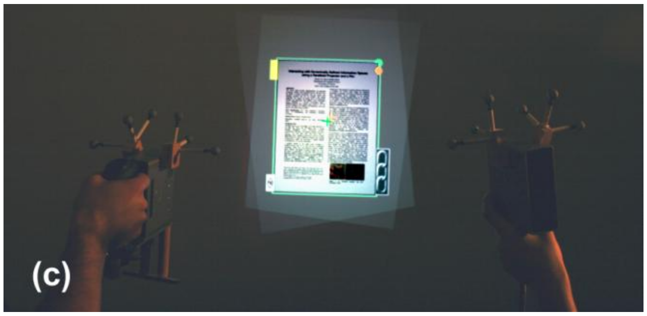
Two docked objects become precisely superposed, and will be operated as a whole. Both ownership flags are shown side by side. The objects may also change their appearance to reflect information from their partners. A linkage flag on the bottom right corner indicates that the linkage has been established. (Figure 6.12c) The pair of objects can be operated by both object owners, and any operation on the pair will affect both objects. For example, two half-completed documents (each worked on by a different user) can be docked to preview the combined document. Both users can then write annotations on it using their pens. The annotations are shown on both documents when undocked.



(a) Before.



(b) During.



(c) After.

Figure 6.12: Docking action.

To undock a pair of objects, either user can toggle the linkage flag by crossing it. Then users can move their own objects away. Alternatively, two users can "tear apart" the objects by both capturing the docked pair at the same time, and dragging them in different directions. The "tear apart" action can also be performed immediately after Step 3 before either user releases the button. This enables "transient" docking to get a glimpse of the docked view.

Other examples include docking two personal calendars. Either user can then click in an empty timeslot to create a meeting for both people. For each user, the meeting will be labeled as “Meet X”, with X being the name of the other user (the label will not be shown until the calendars are undocked). Two users may also dock their portals, resulting in a two-way portal between their personal folders. Objects dropped into the docked portal by either user will be transported to the other user’s folder. This provides an efficient way for them to quickly exchange files.

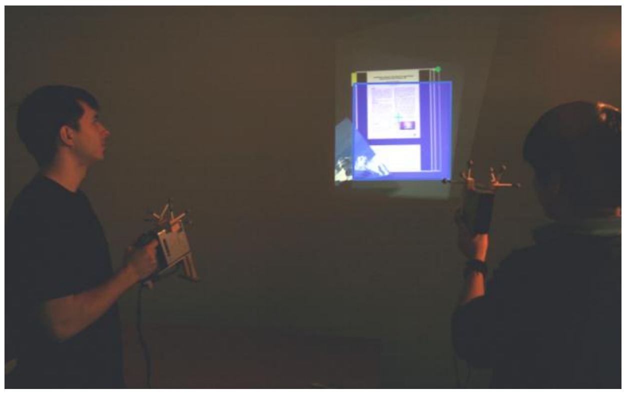


Figure 6.13: Snapshot.

6.2.5 Snapshot

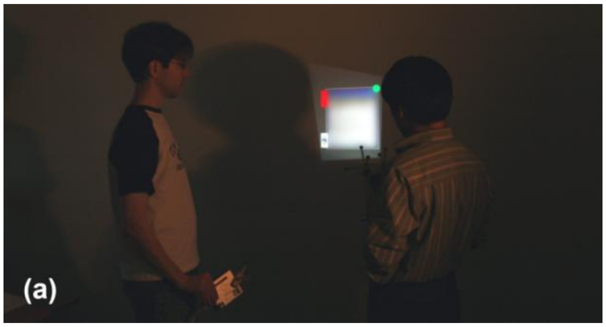
While working in a shared workspace, sometimes a user may want to record information for later reference, especially when the information is from other users or created collaboratively. Triggered by a menu command, a user can take a snapshot of his/her region of interest. A translucent square (“viewfinder”) inscribed within the projection area indicates the region to shoot at, which can be moved and resized by moving and rotating the projector respectively (Figure 6.13). Pressing the primary button takes the shot as an image copy of what the user’s projector displays inside the viewfinder, which can be then manipulated and saved as an ordinary object owned by the shooter. The snapshot can also be used to take small parts from a large object (e.g. a document or a map) to reflect the point of interest. Note that a user cannot take peep shots of private information displayed by others, since that information is not shown in the shooter's projector.

6.2.6 Spatial Relationship between Users

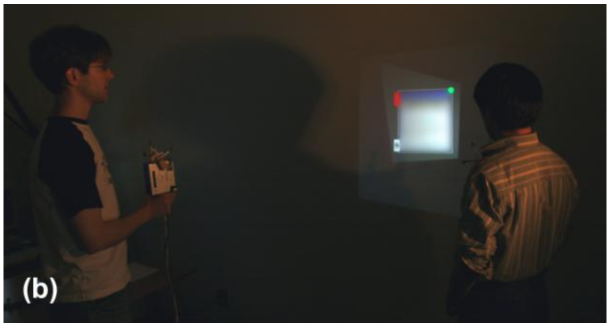
The spatial proximity between people plays an important social role in terms of privacy. People get close to each other to have private conversations. Conversely, people feel uncomfortable if somebody else comes nearby when they are viewing private information. Our system can estimate user proximity and face orientation from the spatial locality of the projectors, and use this information to facilitate subtle interpersonal interaction that exploits real-world social protocols.

When a user is viewing a private object, if another user comes nearby, the private object becomes blurred so as to prevent the second user from peeking at its content (Figure 6.14a). Similarly, private objects also get blurred when other users cast their projection onto it, which also suggests that they are looking in that direction (Figure 6.14b). On the other hand, only when two users are close to each other can they perform private communications such as passing ownership of a private object. For another example, when a private letter is addressed to both users, its content will only be revealed when the two people stand close to each other, and overlap their projection on the letter (Figure 6.15).

Another feature that exploits the spatial relationship between users is to avoid shining the projector into people's eyes. When a projector is pointing at any other users' estimated position, it temporarily turns off to avoid hurting their eyes.



(a) when another user is nearby.



(b) when within another user's projection.

Figure 6.14: Blurring private information.

(a) Content hidden when they are afar.

(b) Content revealed when they come close.

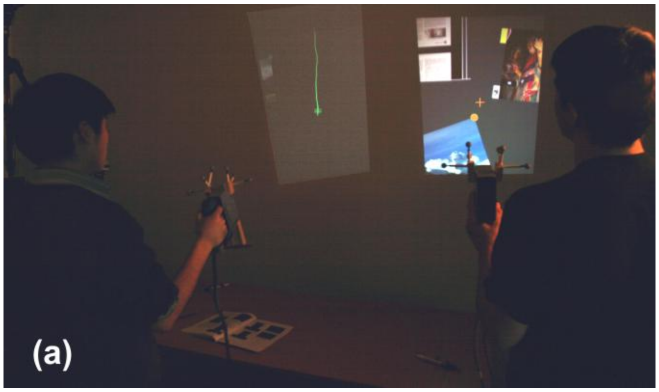
Figure 6.15: Private letter for two users.

6.2.7 Independent Work

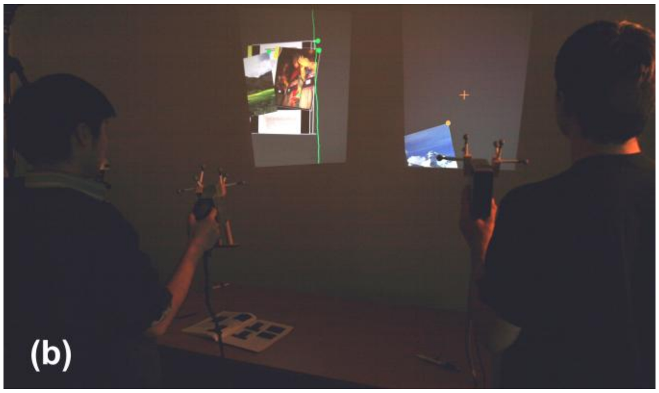
Users sharing the same physical space may not always be collaborating with each other. Compared with systems using a single shared display, the use of multiple handheld projectors allows users to see information relevant only to them. This largely reduces the likelihood of interpersonal conflict when people are working independently in a shared space. To further facilitate independent work, we provide the ability to create a “fence” around one’s work.

Triggered by a menu command, a user can sketch a line in the workspace using the cursor, which turns into a fence between his/her territory and another user’s. Once drawing the line is finished, all objects in the workspace that belong to the user will be pulled back to his/her side of the fence, making room for the other user to work (Figure 6.16). The other user’s objects will stay where they are. The fence will stay visible as an informal demarcation between two users’ territories. However, it does not actually prevent users from moving objects beyond it. It is up to the users to maintain the notion of the boundary. The reason for this design is that we provide this feature only to facilitate but not override or enforce social protocols. Users will still have the flexibility to dissolve the boundary between them, or completely ignore it and start collaboration. This is also why the fence only pulls back the owner’s objects but does not push back other users’ objects. By doing so, people accommodate and do not compete with each other. The use of the fence should only be the result of a well-negotiated common understanding between people. Both users can draw a fence multiple times, only the most recent one in the workspace will stay visible.

A user can also set the access level of all their objects at once to prevent other users from operating or viewing any of their belongings. Both this and the fence feature can be particularly helpful when a user starts by working alone, and is subsequently joined by another user who then shares the workspace.



(a) Before.



(b) After.

Figure 6.16: Drawing a fence.

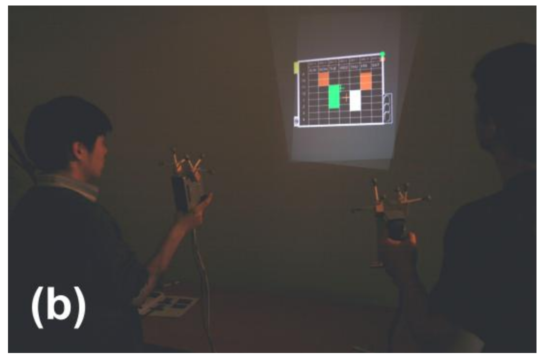
6.3 Usage Scenarios

The interaction concepts and techniques discussed above can support a variety of potential usage scenarios, including but not restricted to the following:

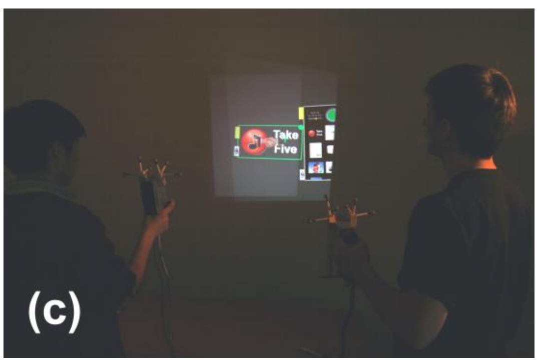
6.3.1 Casual Communication

Given the portability of handheld projectors, it is natural to use them to facilitate casual communication between people when they encounter one another. For example, exchanging contact information can be as easy as dragging avatars of people between personal folders (Figure 6.17a). Scheduling a meeting can be done by docking two personal calendars, without the hassle of separately comparing each other’s schedule (Figure 6.17b). Sharing information such as music and photos can be achieved in several ways, including passing ownership and dragging into a personal folder or portal (Figure 6.17c). A user can also use the pen to write a note and drop it into another person's folder as a reminder (Figure 6.17d). Using current mobile devices, the above activities usually involve explicit and indirect data transmission and synchronization steps that do not fully leverage the co-locality of the users.

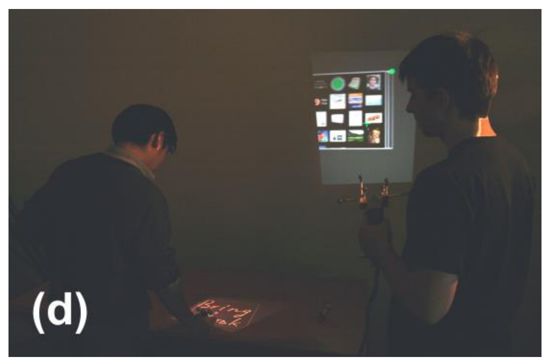
(a) Exchanging contacts.



(b) Scheduling a meeting.



(c) Music sharing.

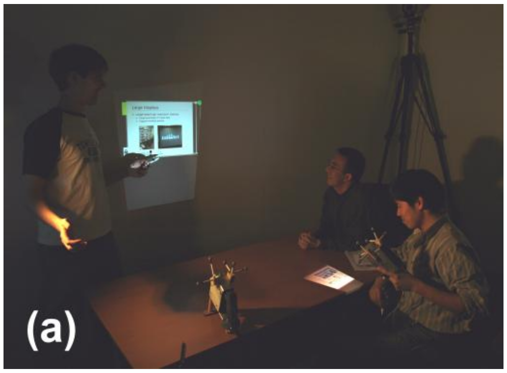


(d) Writing a reminder for the other person.

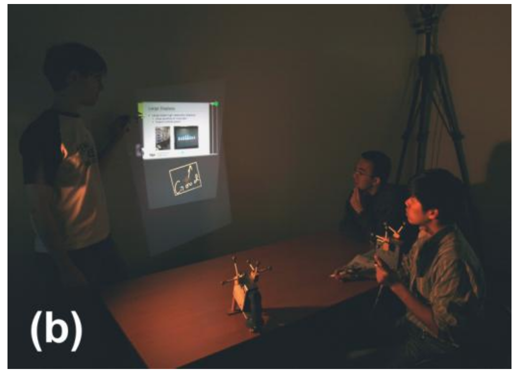
Figure 6.17: Casual communication.

6.3.2 Group Meeting

Although mostly designed for mobile usage, multiple handheld projectors may also support more organized group meetings such as presentations or brainstorming. One projector can be dedicated to display the presentation slides or project a virtual whiteboard to write on. Attendees can use their personal projectors to access additional information on the table (Figure 6.18a). They can also use the pen to write comments on the table first, and then drag it to the wall to post it when desired (Figure 6.18b). In addition, they may take snapshots of the presentation slides using the projector, and then write notes on them (Figure 6.18c). Taking snapshots of presentation slides has become a common practice using digital cameras, but annotating them is currently inconvenient to do without the slides available beforehand.



(a) Looking up additional information.



(b) Posting a comment.

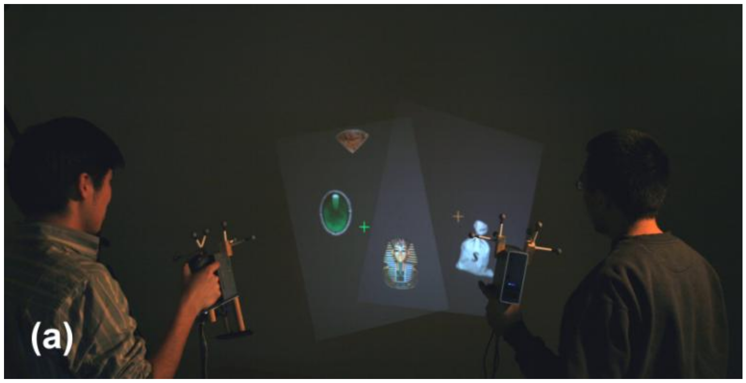
(c) Taking a snapshot of the presentation.

Figure 6.18: Group meeting.

6.3.3 Games

Even without dedicated design for gaming, the current system features can already support a few simple interesting games. For example, utilizing the access levels and multiple views of objects, we can create a collaborative treasure hunting game. Players need to complement and combine their projectors' viewing powers to discover and collect treasures scattered around the physical space. Some treasures are only discoverable by a particular user, others only when projections overlap (Figure 6.19a).

Utilizing the snapshot function, people can also play an ad hoc jigsaw puzzle game. Players take parts from a large picture using snapshots, and try to reassemble them later (Figure 6.19b). These possibilities also lead to our exploration on public games supported by handheld projectors in Chapter 7.



(a) Treasure hunting.

(b) Jigsaw puzzle.

Figure 6.19: Games using multiple projectors.

Mobile social games [9, 15, 70] in which players physically walk around a city to complete tasks such as collecting treasures according to directions given by mobile devices, have attracted attention recently. These games encourage players who are not familiar with each other to collaborate face to face, therefore promoting social interaction between people. With the assistance of handheld projectors, the experience of mobile social games may be altered and arguably improved by projecting game information into the physical environment, thus further blurring the boundary between the game and reality.

6.4 User Feedback

For preliminary user feedback on our designs, we asked four graduate students, working in pairs, to try the prototype system. Three of the four participants are regular cell phone users, and one of them owns a PDA. We demonstrated all the system features to each pair of participants, and then asked them to freely try out the techniques, especially those which involve interaction between them. Each session lasted about an hour. We observed participants' behaviors, and conducted a post-study interview.

All participants grasped the system concepts quickly, and did not show any difficulty learning the interaction techniques. As we expected, the feature that they found most appealing is the ability to easily exchange information in a shared workspace. The multi-view calendar also especially resonated with users, as it largely simplifies one of their most frequent tasks – scheduling meetings. Other highly welcomed features include the movie player that adapts to multiple projections, the snapshot function, the focus plus context map, and the fence to support independent work.

The participants' experience seemed to be affected by the imperfect alignment between projectors, as well as the somewhat jittery projection caused by unideal image update rate (≈25Hz). These could be reduced with technical advances. They also had some reservations about projecting private information in public space, although they all agreed that the system designs surrounding privacy protection alleviated their concerns to some extent. Exploiting the embedded small screen on handheld devices for highly private information and operations may be one way to address this concern.

Some participants asked for more advanced support for collaboratively authoring and annotating text. Another participant suggested having a selection box that can be moved and resized similar to the snapshot viewfinder to quickly select and operate on multiple objects. For independent work support, participants suggested expanding the fence to other shapes such as a circle.

6.5 Discussion

Handheld projectors provide interesting design challenges compared to other co-located collaborative technologies such as a shared tabletop display. For example, users can create their individual displays with their projectors, allowing for easy support of personalized views, which is seldom the case in other settings. This also enabled the three-level access control we proposed as opposed to simply public vs. private in most other systems. Users can also easily point the projector to virtually anywhere in the workspace with little physical constraints, whereas in the tabletop setting the reachability is constrained by the user's sitting position and arm length. This makes some social protocols that work in the tabletop scenario less applicable in our setting. Therefore we considered alternative ways to coordinate users, especially for supporting independent work. Further, when working with handheld devices, the somewhat stable demarcation of personal and group territories used in tabletop interaction [103] is less applicable because of the constant change of users' positions. However, changes in spatial relationship between users can be exploited to facilitate subtle interpersonal interaction.

In our prototype, both projectors are connected to the same computer, thus all data transmission is done locally. In the real world, we can expect the data exchange between handheld projectors to be backed either by peer-to-peer connections such as Bluetooth or infrared, or by centralized services such as WiFi or cell phone networks. The shared workspace created by the projectors can make the background connection mechanism transparent to the users. Identity verification is achieved simply by looking at the person's face, thus eliminating the need for passwords or other complex authentication schemes.

An interesting issue is that although our system provides various ways to support privacy, in some social contexts the very fact that the user is projecting data may be perceived as an indication that the information is public, and viewed as an invitation for other people to participate. This disparity indicates that more delicate design may be needed to convey subtle privacy cues to others, without changing current social protocols. One possible solution may be complementing the projection display with the small screen embedded on the handheld device to accommodate different scenarios. However, the implicitly public nature of projected imagery suggests that handheld projectors may be an ideal platform for mobile social games, which encourages ad hoc participation and initiates social interaction between strangers.

In this chapter, we have explored interactions between multiple users supported by handheld projectors, which naturally leads to the question how the social interaction patterns between people might evolve to reflect the usage of such a new technology. The unique affordances of handheld projectors, such as the dual display attribute supporting both public and personalized views, might result in social experiences different from those with current mobile devices. Handheld projectors may also support new social applications such as multi-player public games, creating new experiences of public spaces. To explore these possible outcomes, in the next chapter we create a public game supported by handheld projectors, deployed as a conduit to investigate influences of handheld projector usage on social and public lives.

Chapter 7

Public Game Supported by Handheld Projectors

Our exploration of multi-user interaction opened the realm of social interactions mitigated by handheld projectors. With the vision of projection devices embedded in all handheld devices, we can expect that they will have interesting influences on people's daily and social life. The somewhat public display created by handheld projectors will likely result in a different user experience and possibly different social interaction patterns and protocols than those with today's mobile devices. As illustrated in Chapter 6, handheld projectors can be involved in a rich set of social activities, which might benefit and evolve from the usage of them. In order to investigate the impacts of handheld projectors in social settings, we chose to concentrate on the detailed study of one example application, namely public games, as a starting point.

Public spaces have always been hotspots for social game activities. It is common to see children playing games in city squares, people playing Frisbee in parks, or chess-lovers "duking it out" in bars. These game experiences weave seamlessly into public space usage, and create a lively and social atmosphere for players as well as spectators around. Recently, the increasing presence of large-sized public displays starts to offer a unique opportunity to position computer games into these public spaces. If designed well, these highly visible interactive games might invite casual play and spectatorship and promote social interaction amongst the public, similar to traditional non-computerized public games. Several researchers [95, 101, 132] have explored the design of such public display entertainment experiences. Similarly, the implicitly public nature of projected imagery also suggests that handheld projectors as a social game platform may create enjoyable experiences in public spaces. In addition to the public display attribute, the personalized views created by handheld projectors can also be exploited to further improve the game designs. As a specific case of social interaction, public multi-player games may provide a valuable viewport into the influences of handheld projector usage, as well as the roles of public display games in general.

To investigate handheld-projector-supported public games, in this chapter we create a lightweight ad-hoc multi-player game prototype, *Flashlight Jigsaw*, enabled by multiple handheld projectors and played in public spaces. The game design is guided by deliberate consideration of both the affordances of handheld projectors and the attributes of public social spaces. Although designed for handheld projectors, the game experience can be easily simulated on static public displays as well. Hence, for our research to be more widely applicable, we sought to use this game as a tool to investigate phenomena not just specific to handheld projectors, but related to public display games in general. To collect real-world usage patterns, we deployed the game using a simulated setup in both a shared lab space and a public atrium. We investigated the experiences and behaviors of game players and spectators, as well as the interrelationship between the game and the spaces it is deployed in. The research resulted in findings regarding game play, communication, social interaction, spectatorship, and space and location around such public display games. We also use our findings to develop design implications for future games suited for handheld projectors and public displays in general.

7.1 Game Prototype

7.1.1 Design Goals

To guide our game design, we considered several goals that a public display game should achieve:

Casual and lightweight: Most games are designed for dedicated intense play. Games that blend seamlessly into public space experiences must be casual and lightweight enough to be concurrent or multiplexed with other activities, especially social interactions between people.

Simple to understand and operate: Unlike dedicated spaces where people come to play (e.g. casinos, arcades), people typically enter public spaces without forethought of game play, and any participation in games will be ad-hoc. Thus, the game must be simple and quick to understand for both players and spectators without intensive instructions.

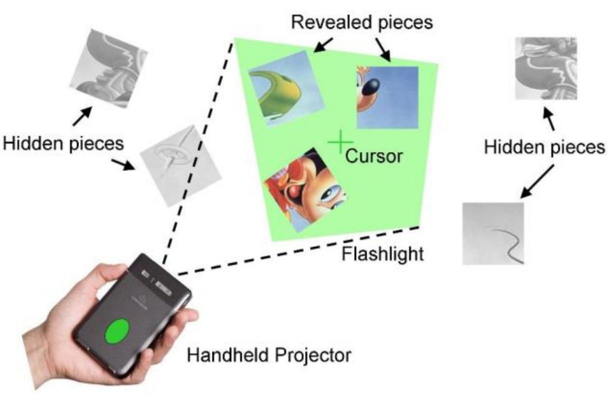
Suitable for various populations: Given the usually diverse and dynamic population in public spaces, the game should broadly appeal to people of different ages, educational and ethnic backgrounds, technology experiences, and so on.

Ad-hoc joining and leaving: Similar to some traditional games, the game should allow people to initiate, join or leave at any time in an ad-hoc fashion ("drop in, drop out") without interrupting the game experience.

Encouraging group play and communication: As a part of the social experience in public spaces, the game should not only accommodate but also encourage people playing and communicating in both preformed and spontaneous groups.

7.1.2 Game Design

We create a game prototype called *Flashlight Jigsaw*, which is a multi-player jigsaw puzzle game played using handheld projectors (Figure 7.1). The design can also be directly migrated to a situated large public display and wireless controllers tracked in 3D, which simulate the effect of handheld projectors. We chose jigsaw puzzles because they are familiar to most people and not very intense, but still challenging enough to be interesting. Unlike traditional jigsaws where all pieces are simultaneously visible, pieces in *Flashlight Jigsaw* are only revealed within the "flashlights" casted by the handheld projectors on a large projection surface (Figure 7.1a). The object manipulation technique we described in Chapter 4 is applied: each player can use the primary button with the crosshair cursor in the center of their flashlight to select and move jigsaw pieces. Rotating the projector about its Z axis (Figure 3.2) rotates the selected piece. When a piece has been correctly placed alongside others, they will connect to each other and be moved as a whole. Each player is assigned a different color as marked on the handheld projector, and which also virtually identifies the relevant flashlight and cursor.



(a) Flashlight experience.

(b) Types of jigsaw pieces.

Figure 7.1: *Flashlight Jigsaw* concept.

The current prototype supports up to 3 concurrent projectors/players. To elicit interaction between players, each projector has a different viewing and operating power, being able to reveal (with flashlight) or move (with cursor) different jigsaw pieces. Figure 7.1b illustrates. The type (and associated color if applicable) of each piece is randomly assigned when the puzzle starts. Therefore, players need to work together to discover and fit pieces. When there are fewer than 3 players, a player can switch to another available projector by pressing the "switch" buttons on both projectors, in order to access other pieces. In the case that only one player is present, after the projector switching, all the "star" pieces are converted to pieces of other types, so that a single player can access them.

Players may join or leave the game at any time regardless of the status of the puzzle and other players. Therefore the game seamlessly transitions between different numbers and sets of players. Once a puzzle is completed, another randomly selected puzzle is generated. During play, hint messages appear periodically or when triggered by context, such as reminding players of pieces they cannot access, or suggesting that players find other people to play with. If a situated projector is available, when there is no one playing, a simulated flashlight randomly scans throughout the display area to allow passers-by to see the puzzle content, enticing them to play.

A player scores points each time s/he places a jigsaw piece correctly. When the entire puzzle is completed, all players present get bonus points, ranging between 50~75% of the puzzle's total score. The player who places the last piece of the puzzle gets the highest bonus. To encourage group play, all scores are multiplied by 1.5 when 2 players are present, and by 2 when 3 players are present. There are no time requirements or incentives. Players can check their records (score, number of puzzles completed, total playing time) by pressing the secondary button on the projector.

7.2 Deployment Study

Because of the technical limitations of our current prototype handheld projectors (weight, brightness, connecting cables), they are not yet suitable for public deployments. Instead, we use a static public display and wireless controllers to simulate the same game experience. On the other hand, this also enables us to investigate the game in an environment more similar to ordinary public display situations. We use a short-throw projector to create a public display measuring roughly 3m x 2m. The game controllers are built from wireless remote presentation controllers with several built-in buttons, and tracked using the same Vicon tracking system as in our handheld projector prototype (Figure 7.2). Given the position and pose of the controllers, and the pre-calibrated geometry of the projection surface, the position, size, and shape of the flashlights are calculated using a mathematical projection model. The game content is then rendered within the flashlight areas accordingly. This results in a realistic-feeling flashlight experience.

Figure 7.2: Game controllers used in simulated setup.

We deployed *Flashlight Jigsaw* in two locations in an urban university campus, one school week (Mon-Fri) per location:

In Week 1, the game was deployed in a lab of around 30 computer science graduate students, researchers, and software engineers, who all knew each other previously (Figure 7.3a). Inside the room there is a wall-sized projection screen, in front of which is an open space measuring roughly 5m x 3m. Two people work in the room regularly, while others work in adjacent offices. The room also hosts the lab printer, and serves as the passageway between the lab pantry/meeting room and two large shared offices. Thus people frequently pass through the space. The game was available daily between 1-7pm.

In Week 2, the game was deployed in the ground-floor public atrium of a large academic building (Figure 7.3b). The lower floors consist of classrooms and common areas, resulting in constant traffic throughout the day, especially when classes start or end. Due to the open spaces, public seating, and nearby cafe, people also tend to linger. The probability of people in the atrium knowing each other is much smaller than in the shared lab in Week 1. The display was projected on a wall in an open space of roughly 8m x 5m, which is easily visible throughout the atrium, but not directly in the way of traffic. The game was available daily between 11am-5pm.

(a) Shared lab space.

(b) Public atrium.

Figure 7.3: Game deployment spaces.

We chose these contrasting locations to investigate how people reacted to the presence of the game in both a shared space with a relatively fixed set of occupants, and a public space with a more diverse and transient population. In both locations, a poster illustrated the concepts of the game. Flyers with the same content were available for pick up. When appropriate, the onsite game facilitator would encourage people who stopped at the game installation to play. An online blog kept updated player rankings and allowed people to leave game-related comments.

Fifty different puzzles were used in the deployments, all generated from pictures of Disney cartoons, chosen for their popularity and familiarity in North American culture. Each puzzle was segmented into 10-20 pieces.

To offer continuity of experience and personal identity to the players, as well as for the ease of organizing and analyzing the study data, each player is assigned a unique ID by the game facilitator when the player plays for the first time. The game facilitator is responsible for logging in/out the player to a specific controller. The player's record including scores is maintained across multiple play instances. A nearby desktop monitor serves as a scoreboard which lists the ranking of all player records for people to check. Note that for a non-study situation, we could choose not to require a log-in process to allow for anonymous play.

At all times, one or two researchers were in the deployment space, recording observations and conducting impromptu onsite interviews with players and spectators. The players' interaction with the game was logged by the system and video recorded. After each week of study, semi-structured follow-up interviews were conducted with several players selected from among those who opted to leave their contact information when they first played. Selected interviewees included players with the highest scores at the end of the week, as well as representative players covering different behaviors and backgrounds based on researcher observation during the week. As compensation, the interviewees were awarded prizes in cash or electronic devices. To some extent this also served to encourage people to start or return to play. Interviews were conducted individually or in groups of 2-3 players who played together. Interviews were coded using open coding. Two researchers first jointly coded 2 randomly selected transcripts from each week in order to establish baseline agreement. One researcher then coded all

21 interviews transcripts, resulting in 571 unique codes, organized into 5 major themes as reported in the findings section. Limited by time, the second researcher randomly selected and coded 3 interviews to establish inter-rater reliability (IRR), resulting in observed agreement of 97.6%, and Cohen's Kappa of 0.662 averaging across all codes, indicating good agreement between the coders.

7.3 Findings

A total of 239 people played *Flashlight Jigsaw*: 28 in Week 1 and 211 in Week 2. We interviewed 11 players from the shared laboratory space in Week 1 (noted as P1-1 ~ P1-11). All 11 were male computer science researchers or software engineers aged 23-31. We interviewed 16 players from the public atrium space in Week 2 (noted as P2-1 ~ P2-16). Of these, 6 were women and 10 were men, aged 17-25 and were undergraduate and graduate students from several departments including engineering, computer science, psychology, nutrition, and biology. We use the themes derived from analysis of the follow-up interview transcripts as a primary framework for describing our findings. We situate these themes in conjunction with other data sources including observations, onsite interviews, and system logs, to provide a holistic understanding.

7.3.1 Game Play

Players unanimously liked the game experience, which was described as "*fun*", "*enjoyable*", "*cool*", and "*rewarding*". They particularly liked that the game was simple to understand and operate (20/27: 20 out of 27 interviewees explicitly mentioned, similarly denoted hereafter), and that they could play and communicate with other players (17/27). The game was also perceived as casual and lightweight (6/27), thus enabling other activities such as conversation with spectators.

The large size of the display appeared to be an important factor in the game experience (8/27). It was mentioned that in other co-located multi-player game settings such as arcades, the screen is usually occluded by the player’s body, making it hard for spectators to be involved. The large size also contributed to a more immersive playing experience (“*Feels like you are a part of the puzzle*”). Players even suggested deploying the game on even larger scale displays such as a movie theatre screen, which could enable a massive number of players to participate.

The main complaint related to imperfections of the tracking system, which resulted in the occasionally noisy or lagging controls (20/27). The freehand pointing and standing posture were tiring after long plays (6/27). Suggestions for improvements include adding sound feedback, additional visual cues such as jigsaw edges on the pieces, and that the puzzle difficulty should adapt to players’ experience levels.

Motivation to Play

Several factors contributed to a person’s decision to play:

Novelty and curiosity (17/27): The majority of players were attracted by the novelty of the game and the technology when they first passed by, a common phenomenon to new activities and installations in public. However, 46 players played more than once, indicating that the game had an attraction beyond initial curiosity.

Entertainment and sense of accomplishment (17/27): Like any other game, the entertainment value itself is a major reason that people played. In addition to having fun, the sense of accomplishment gained by leading or defeating other players also contributed to people’s desire to play.

Filling gaps in life (14/27): The game served as a relaxation tool when people took breaks from work or study. At other times, the game helped to occupy people between daily activities, such as when waiting for classes to begin.

Influenced by other players (16/27): Influence from other players played an important role in encouraging people to play. This included being invited by other players; joining when noticing other players; joining to help other players; or using the game to socialize with friends or strangers.

Prize incentive (7/27): Unsurprisingly, the prize incentive also played a role in encouraging people to try the game and seek higher scores. However, no interviewees considered it to be their primary motivation.

Collaboration and Competition

The game required all 3 controllers to complete the puzzle. While a single player could switch between multiple controllers, the most effective way to complete the puzzle and achieve a high score is to collaborate with other players. However, each player acquires scores individually, which potentially instills a sense of competition between players. This design decision enables both collaborative and competitive behaviors to emerge.

The game yielded a mixture of play styles. Many players reported they primarily played in a collaborative way (14/27), while others reported they played competitively (4/27). Of particular interest are the players who played both competitively and collaboratively at different points in the game (9/27). Competitive play occurred primarily between people who already knew each other, whereas collaborative styles were frequent between both acquaintances and strangers. Players also adjusted their play style depending on who they were playing with.

Players orchestrated their collaboration mostly verbally. They gave directions and suggestions ("*That piece goes to the bottom-left corner*"), and shared information with other players ("*There's a block of your color here*"). When needed, players also coordinated their actions, such as moving their flashlights synchronously to search for the star pieces. Some players would also assist others to finish certain tasks, such as moving irrelevant pieces out of the way, or shining the flashlight to let others grab a star piece. Players who played collaboratively thought it was "*fun*", "*enjoyable*" and "*relaxing.*" Friendly collaborative game play appeared to fit the casual setting of the game: "*…when I rest, I don't want to be so competitive. I just want to relax, so this collaborative game is appropriate for this situation*" (P1-1). People also valued the sense of working as a team: "*I had a team that started to work together better … it was definitely rewarding*" (P2-16). The game scoring system also encouraged collaboration: "*We all get points after we're done. Then, we can get more puzzles then*" (P2-2).

When playing competitively, players employed various strategies to maximize their own scores and/or minimize others'. This was not necessarily for pragmatic reasons such as winning a prize, but rather for the enjoyment of competition itself: "*It's always fun to be a little better than somebody else*" (P2-9). Some of these competitive strategies were: playing as fast as possible to overrun others; completing the public pieces first and saving pieces of one's own color for later; creating obstacles and inconveniences for others; or even deceiving others. The most frequent behavior was that in the end of a puzzle, each player would capture one piece, and all players wanted to be the last to release the piece in order to gain the higher completion bonus. This caused a temporary stalemate until someone conceded. Some reported that the desire to increase their scores and defeat others motivated them to play repeatedly (7/27).

7.3.2 Communication

The rich inter-player interaction in the game resulted in constant communication between players during most game sessions. The communication was mainly verbal, but also included body language such as pointing (with finger or flashlight) and gestures, which was made possible and necessary by the large size of the public display.

Most in-game communication was directly related to the game play and involved giving directions and suggestions, asking for help or input ("*Somebody give me a little* *light*”), or coordinating actions (“*Let’s search some other area together*”). While more direct and imperative statements were also observed in some cases, (“*Stop moving my piece*”, “*Drop it!*”), these were delivered in a friendly manner. Both one-way directive sentences and back-and-forth conversations were frequent. Communication tended to increase towards the end of each puzzle, when the players needed to concentrate in the same area or pieces, whereas in the beginning players acted more independently and required less communication.

More experienced players often voluntarily tried to explain and teach the game to new players. This happened regardless of the player being collaborative or competitive. This spontaneous propagation of game knowledge helped to lower the entry barrier for the game, and created a welcoming atmosphere. The teaching behavior itself was also considered to be a rewarding experience by the players.

The game also provided an opportunity for players to make small side talk and to socialize briefly, which often happened when somebody had just joined, or to fill the gaps during the game, e.g. when a player did not have an available piece to move. The puzzle pictures also provided a frequent chatting topic (“*Look it’s Aladdin*”, “*I grew up with it [Disney]*”). These non-game-related communications tended to be short and opportunistic. But interestingly enough, a few players mentioned that during the few occasions of a system outage, they did have a chance to have extended chats while waiting for the system to recover, which also helped to ease the wait.

The in-game communication was found smooth and enjoyable by many players (21/27). One mentioned that the good communication experience during the game made him more confident talking to people in general: “*Usually I wouldn’t have the initial to talk to someone at a public event… Playing the game definitely made me more confident*” (P2-1). For players who did not know each other previously, the amount of communication tended to increase over the period they played. The game served as an icebreaker between people: "*We sort of got to know each other and got comfortable with talking with each other so it was a lot more fun by the end*" (P2-13).

In addition to communication during the game, players also communicated about game-related topics later. While most after-game conversations happened between players who already knew each other before playing, there were also cases where strangers continued talking after they left the game, potentially providing a basis for further socialization. One player also tried to advertise the game to other friends through word-of-mouth and e-mails.

7.3.3 Social Interaction

Alone vs. Together

Most players preferred playing with others (25/27), of which 10 said they would not play the game if they were alone. This was partly because all 3 controllers were needed to complete the puzzles, requiring a single player to switch between controllers (although several players became skilled in this). However, social interaction was considered the main reason for preferring multiple players (18/25): "*Playing with other people, like I said, you feed off the fun that they're having*" (P1-8), "*[if played alone] then it's just like any other puzzle at home*" (P2-3).

Players indicated that puzzles were easier to complete and that they played longer when with other people, claims we validated by system log data. In both weeks, as the number of concurrent players increased, average completion time decreased, and average length of play session (time between when a player joined and left) increased (Figure 7.4).

Number of players	1	2	3
Avg. Puzzle completion time (minute)	7.1	6.5	3.7
Avg. Play session length (minute)	6.5	10.5	12.2

Figure 7.4: Influence by number of players.

Group Behavior

Players who played together can be considered to be a spontaneously formed group. Players implicitly assumed different group roles such as leader or follower without explicit assignment. Players also adjusted their group roles according to others' behaviors, e.g. choosing to follow or lead depending on whether there was already a leader or not.

In Week 2 (public atrium), often a group of 2 or 3 friends arrived together and wanted to play. In the case that there were not enough available spots for them to join together, they would often choose to wait until some player left, rather than split the group. Comparably, in Week 1 (shared lab space), players often tried to recruit a group before starting to play. Most of the time, the group would continue playing until all players left together: "*It's a team game. If you leave, probably they will give up*" (P1-1).

Sometimes, we observed in-group isolation, i.e. two players who talked to each other while isolating the third. Reasons for this included: two players knew each other (better) but not the third player; one player fell behind because of lower skill or technical difficulty; one player closely interacted with a spectator, which isolated herself from other players; or one player's competitive strategy resulted in isolation.

Joining and Inviting

In addition to the individual motivations we mentioned previously, many players decided to join because other people were playing or going to play. New players frequently watched others play before joining. For returning players, the presence of others lowered the barrier to further play ("*Saves me some time to recruit people*"). This happened particularly frequently in the shared lab of Week 1, where there were not constantly people playing; it was not always easy to recruit partners because of the smaller population and work-oriented nature of the lab. However, the lab members who worked near the game installation would join when they saw or heard other players ("*If I* *hear a game I would move in for the kill*”). Those who walked through the room to pick up printouts or coffee got sidetracked to play if others were playing.

Players frequently invited others to join, either before they started to play, or in the middle of their games. Many players invited others for pragmatic reasons, such as to find hidden pieces or increase their scores, especially when they encountered a bottleneck in the game (9/12 of the interviewees who invited others); players also invited others to improve their game experience, given more players generally resulted in more fun (4/12); they also wanted to invite other people, either friends or strangers, to share the game experience that they enjoyed (6/12). The invitation was ad hoc, especially when in the middle of a game. Spectators were frequently invited to play, and general invitations were sometimes yelled out (“*Anybody wants to join me?*”, “*We need one more brain!*”). In addition, in Week 1 there were several cases that a person wanted to have a break from work and invited his officemates to go to play together. In Week 2, one player went to the student common room to invite her friends to try the game together, who then played and remained as a group. Compared to people invited by the game facilitator, invitations from players were more successful.

Socializing

Brignull and Rogers [23] explored using public displays to elicit socializing behavior between people. *Flashlight Jigsaw* had a similar effect to a certain extent. People tried to initiate socializing conversation when new players joined or during gaps in the game: “*It was an excuse to socialize briefly with lab members that I don't talk to so much*”(P1-5).

Interestingly, the game provided a tool through which players learned about other people’s personalities, both for friends and strangers (11/27): “*I guess that <P2-3> is a little more aggressive than we thought.*” (P2-5). These impressions were often strong, such that in Week 2, 6/16 players recalled particular strangers they played with.

Although the game served as an icebreaker to initiate socializing, players' concentration on the game play impeded further socializing behaviors within the game itself. Players also said that the deployment spaces we chose were not particularly appropriate for follow-up socializing. They suggested that spaces where people would linger longer such as parks or cafes would better accommodate this.

7.3.4 Spectating

The nature of public display games ensures that spectating will always be part of the experience. We investigated the experience of both spectating and being spectated, through the follow-up interviews of players who also spectated, onsite impromptu interviews of spectators, and observations.

Reasons to Spectate

The reasons that attracted people to spectate were as diverse as those for playing. To list the most important ones:

Novelty (8/19 of interviewees who also spectated): The novelty of the game and technology stopped passersby wanting to figure out what it was. Some people also wanted to learn the game through watching before they played.

Puzzle and picture (9/19): Interestingly enough, many spectators simply enjoyed watching how jigsaw pieces went together, which was described by one person as "*mesmerizing*". The curiosity about what the final picture would be also caused people to stay and watch. The choice of Disney pictures seemed to echo well with many spectators.

Attracted by players (10/19): Players' performance and behavior served as the biggest attractor. Spectators enjoyed watching players both excel and struggle, as well as the interaction between the players. The excitement of the players also appealed to passersby. In particular, people were more attracted to spectate when they saw players they knew.

Attracted by other spectators (2/19): Like Brignull and Rogers [4], we observed the "honey-pot" effect: hurried passersby slowed their pace to observe the game, and as a larger mass of people accumulated around the display, more and more people joined them ("gravitation" effect).

Waiting to play (10/19): People who waited for spots to play usually watched other players while they waited.

Spectator Behaviors

There were relatively few spectators during Week 1 given the small and static population. In Week 2 there were always spectators whenever players were present. The number of spectators depended heavily on the time pattern in the building, ranging from 1 or 2 in light times, up to 30-40 when classes began or ended.

Spectators chose different standing positions to reflect their willingness to interact with the players. People who did not want to be involved watched quietly from afar, while those who stood near the players usually tended to communicate with them, especially when they knew some of the players.

Spectators' communication with players included directions and suggestions, commenting on the players' performance, and asking players about the game. In particular, some players consulted with their spectating friends on what moves to make, often resulting in a "co-playing" situation.

There was also frequent discussion amongst spectators. Much of this was commenting on the players and discussing puzzle solutions. Former players or more experienced spectators also often tried to explain the game to other spectators. This increased the public awareness of the game and potentially prepared more people to play.

We also observed "quasi-spectator" behaviors, where people stayed in the deployment area and watched the game, while engaging in other activities such as having their own conversations or eating lunch. The game acted almost like a “water cooler” to create a social hub for people to linger in.

Being Spectated

Players had different attitudes about being watched. Some did not care (12/18 amongst the interviewees who noticed being spectated), while some particularly enjoyed spectators’ presence (7/27), partly out of the desire to show off (“*Hey, check out my skills!*”), and more importantly because the spectators contributed to the excitement in the atmosphere: “*It really reminds me of Dance-Dance Revolution because it’s not just the two people who are dancing who are part of the game, but it’s pretty much everybody around who’s cheering them on or looking at them*” (P1-11).

The presence of spectators also influenced players’ behavior and attitudes while playing. Some players felt distracted, nervous, or intimidated if there were many spectators (10/27), especially when they were performing badly or being discussed. Conversely, players tried to perform better or more cautiously when being watched (4/27) (the social facilitation effect [133]). No interviewee explicitly opposed to being watched given the public setting of the game, as one player noted: “*If I’m playing in a public space, I’m prepared for other people watching me*” (P2-1).

Spectating and Playing

Many people experienced the game as both spectators and players at different times. In particular, many people started as spectators and then joined the game. The transition from a spectator to a player was smooth (9/14 of the interviewees who both played and spectated), partly because the spectator learned about the game as well as the players’ strategies through watching. Spectating also encouraged people to start playing themselves, and performance of other players encouraged them to do better when they played (6/14). This ease of movement between spectatorship and playing appeared to be critical to the success of *Flashlight Jigsaw*.

7.3.5 Space and Location

Playing electronic games in public spaces with ad-hoc partners was a novel experience to most players, although many of them did so with real-world games or sports. Some players had played games on personal devices in public, alone or with a few friends. Compared to this, the public display game created a shared experience involving all people in the space. Similar to the findings by Brignull and Rogers [23], the few interviewees who had experience with public displays stated social embarrassment as the key factor that prevented them from using ("*I didn't want to seem dumb*"). In *Flashlight Jigsaw*, the presence of other players helped alleviate the concern and made the public display more socially acceptable: "*I'm not too worried about doing it because other people are doing it like me*" (P1-9). The fact that players were co-located in the space (compared to online games with remote players) also added to the game experience because of the rich communication and interaction involved (13/27): "*It's a little dangerous to remove yourself from all your contacts and sit in a room by yourself, so having people there with you and really engaging with them and communicating is more fun… it would develop their communication a lot better*" (P2-16).

Players generally thought the deployment spaces were appropriate for the game given the regular traffic in them (18/27). Their pre-existing knowledge that the building housed computer science and engineering departments may have prepared people for seeing these kinds of installations. However, this also means the population in the spaces was relatively uniform, mainly consisting of university students with reasonable technology experience. More diverse populations would be present in more general public spaces such as parks or plazas. Nevertheless, among the players were representatives of various age groups, including 3 elderly people. One elderly woman initially shied from the technology ("*Could work for teenagers, for me it's much easier to play on a table*"), but as she watched, she began to give the players suggestions, and finally grabbed the controller from a player to start playing herself. The spectators included various employees who worked in the building, such as janitors, cashiers, and electricians.

Players considered the game suitable for playing for short sessions in a casual manner (11/27), which nicely fit in their life patterns in public spaces. The presence of the game also changed people's experiences of the spaces themselves (9/12 respondents to a follow-up email). The spaces became more "*social*", "*vibrant*", "*relaxing*", and "*approachable*", as opposed to "*boring*", "*dead*", and "*empty*" before the game deployment. For some people, the technology caused them to vary a pre-established walking path: "*After I knew that the game was where it was set up, every time I had a class in the building, I would want to pass by.*" (P2-7) The spaces were converted from a passageway that people only passed by into a social hub where people would stop, meet or come purposely. As a result, the spaces themselves also received more attention during or even after the deployment: "*(I noticed) there exists a payphone underneath the stairs, but after walking around the building for almost every school day for the last 4 years, I've barely noticed it*"(P2-1), "*Whenever I walk past the large display, it reminds me of the different puzzles that I played with my colleagues*" (P1-8). A study of public plazas [125] found that the successful public spaces were those that "stimulate people into new habits – al fresco lunches – and provide new paths to and from work, new places to pause". Similarly, public display games encouraged people to form new life patterns around a space that was otherwise banal or irrelevant to them.

Conversely, in Week 1 the proximity of the game to work spaces was occasionally disturbing: "*When I was deep in thought or in 'work mode' I would sometimes be annoyed at the pandemonium happening in front of the game*" (P1-9).

Given the different characteristics of the two deployment spaces, players also showed different behaviors and playing patterns. In addition to those discussed earlier, Figure 7.5 summarizes the player records throughout each week of game deployment. In the shared lab space of Week 1, people played more frequently (also reflected in more total playing time and completing more puzzles) but in shorter sessions. The vicinity of the game to players' work areas resulted in the relatively "frequent and short" playing pattern, as opposed to the public atrium in Week 2, which had a more dynamic and transient population. The existing social relationship between the people in the lab also resulted in more plays influenced by other players. We expect other different playing patterns and behaviors would emerge if the game were deployed in other public spaces such as cafés or parks.

Week		1	2
Space		Shared	Public
Number of players		28	211
Number of play sessions per player	Mean (SD)	2.86 (2.69)	1.26 (0.99)
	Median	1.5	1
Total playing time per player (minute)	Mean (SD)	26.6 (29.0)	23.9 (29.1)
	Median	16.0	10.0
Completed puzzles per player	Mean (SD)	5.64 (6.99)	2.91 (5.72)
	Median	3	1
Play session length (minute)	Mean (SD)	9.6 (7.9)	13.0 (11.3)
	Median	7.1	9.5

Figure 7.5: Playing pattern statistics.

7.4 Design Implications

Our study indicates that *Flashlight Jigsaw* was a successful multi-player game, satisfying our initial design goals. We expect many of these findings could not only apply to handheld projectors, but also generalize to other public display games in general.

Based on them, we draw some design implications for improving the experience of *Flashlight Jigsaw* as well as public display games in general.

Encouraging Initiators

We found that existing players acted as a strong attraction for other people to play public display games. However, during our study when there were no other players present, first-time players often needed invitation or encouragement from the game facilitator to take the initiative, although some were already interested due to the game display or the poster. The encouragement from a real person was much more effective than lifeless messages on the display. For long-term real-world installations of public display games, which would ideally not require a human facilitator, we could consider using on-screen conversation agents to invite players. For example, when nobody is playing, the display can show an animated or recorded character. The character starts to talk and invite people to play when optical or acoustic sensors have detected passersby. The life-size of the personated character on the public display could potentially produce the similar effect of a real inviting person. Similarly, Vogel and Balakrishnan [118] explored showing an onscreen video of a person on public displays, but used it for explaining system operations only.

Promoting Socialization

Flashlight Jigsaw acted as an icebreaker to initiate socializing, but did not further accommodate it very well. The concentration on the game itself distracted players from holding conversation about other topics. However, we did observe players socializing during the gaps of the game such as an occasional system outage. Inspired by this, we could consider a “gapful” design, i.e. intentionally introducing gaps into the game. Real-world games often have this gapful nature, often reflected in turn-taking such as in billiards, which was mentioned as a good socializing tool by players. Similarly, turn-taking or periodic breaks in the public display game could provide opportunities for players to socialize. However when doing so we must be cautious not to compromise the game flow and dynamics. On the other hand, the game itself could be designed to include themes for people to chat about, just like the Disney pictures in *Flashlight Jigsaw*. By doing these, we could make these games a more effective platform for socializing.

Improving Single-Player Experience

The single-player experience in *Flashlight Jigsaw* was considered inferior to the multi-player experience. The controller switching procedure was annoying enough for some players to give up. In our design it was difficult to balance between improving single-player experience and encouraging group playing, and at the same time supporting seamless transition between the two. We expect this to remain a challenge that has to be addressed for most public display game designs, although the specific solution will vary from game to game. One general possibility is to have automated "ghost" players controlled by the system to group with single players, which will be replaced by real players as they join. The ghost players could be combined with on-screen inviting agents mentioned previously.

Facilitating Group Forming

In addition to spontaneously joining a game, players also enjoyed forming a group before starting. Without sacrificing the ad-hoc nature, we could provide tools to deliberately facilitate this group forming behavior. For example, a matching service could help individual players seek out game partners by sending text messages to interested parties in the vicinity. The game could also maintain group profiles for more persistent groups over time. The game should also support a more volatile number of players so a group can always join together without having to wait for spots.

Designing for Spectatorship

Spectator experience around the game is an inseparable part of public display games. Based on the taxonomy proposed by Reeves et al. [94], the spectator experience of *Flashlight Jigsaw* was "expressive", where both the operations and the effects of the player actions were revealed to the spectators. This enabled the spectators to enjoy and comment on the players' behaviors, as well as prepared them for playing. To further improve the experience of spectating as well as being watched, we could consider introducing participation from the spectators in the game, for example letting spectators vote on players' performance using their personal devices. On the other hand, the game could include a "spectator" mode that plays by itself, allowing people to spectate even when nobody is actively playing. In this mode the game (or the ghost player) may occasionally ask for input from spectators, resulting in intermediate levels of involvement, and potentially elicit spectators to start playing. We could also capitalize on the collaborative relationships between players and spectators. For example, the game could require additional attention, so having a spectating advisor ("spotter") would be the best way to win.

Situated Design

As we showed, people's game experience and behavior were largely influenced by the spaces where the game was deployed. For public display games to successfully blend into public spaces, the nature of the space itself needs to be accounted for in the design: How big is the space? What is it used for now? Is it indoor or outdoor? Is it noisy or quiet? How many people are there? Who are they? Do they know each other? Are they standing or sitting? Are they familiar with the space? Answering these questions would help to situate the game design in the space from the beginning, and guide the decisions we make throughout. In addition, we need to consider what roles the game will play in the space, be it for people to kill time while waiting, for people to relax from work, or for attracting people to gather in the space, and so on. The same game design could result in various experiences in various spaces.

Our exploratory investigation on public games resulted in qualitative knowledge of higher-level social interactions resulting from handheld projector usage. Complemented by the quantitative user performance model for the generic peephole pointing task reported in Chapter 5, we provided a more holistic understanding about the user experience with interactive handheld projectors.

Chapter 8

Conclusions

As an exploration on interactive handheld projectors, I explored several research issues relating to handheld projector interaction, including technical implementation, interaction design, and the user performance and behaviors regarding the usage of them. This work exemplifies a human-computer interaction research methodology that combines theoretical analysis, creative design, technical development, and empirical studies. For this particular research problem, this involves: identifying the interaction situations of handheld projectors by both understanding current user practice and envisioning emerging possibilities; developing a prototype system that supports interactive handheld projectors using a flashlight metaphor; designing interaction concepts and techniques to suit both a single user and multiple users using handheld projectors, and empirically evaluate them; theoretically analyzing and modeling user performance on pointing tasks under the flashlight metaphor, and experimentally validating the model; and empirically investigating social behaviors that emerge from handheld projector usage in a public game situation. In this chapter I summarize the contributions of this research, discuss limitations of the work, and propose directions for future exploration.

8.1 Contributions

The first contribution of my work is the design and implementation of a handheld projector interaction prototype system, which served as the experiment platform for all research in this thesis. The prototype supports multiple handheld projectors sharing the same physical environment and virtual workspaces. The ability for the user to interactively define workspaces in the physical environments enables the system to be potentially deployed ubiquitously. The prototype platform and its architecture and algorithms as described in Chapter 3 could be utilized by the research community for further exploration on handheld projector interaction.

On the interaction design front, I developed and evaluated a rich set of interaction techniques for use with handheld projectors. These techniques cover general operations (e.g. object manipulation and parameter adjustment) useful for various applications, and take into account the special characteristics and affordances of handheld projectors (such as the dynamic image resolution determined by projection distance). The result is a generic interaction vocabulary that could be widely applied by future researchers and designers. I then expanded the interaction designs to support multiple users working in a shared physical space, each using their own handheld projector. The designs deliberately facilitate both collaborative and independent work between users, and emphasize new possibilities enabled by ad hoc composition of multiple projections, such as the blending of multiple personalized views. This work unleashes the power of handheld projectors to support co-located interpersonal interaction, in contrast to traditional handheld devices which are almost exclusively single-user devices. I also explored usage scenarios enabled by these designs, covering various activities in both single- and multi-user situations. My work is an exploration of the interaction design space of handheld projectors. In addition, many of the designs may also be adapted to other interaction situations, such as large display and tabletop interaction, co-located groupware, and mobile augmented reality.

At the fundamental user performance level, I investigated users' abilities to perform generic interaction tasks in such a setting, especially when using the flashlight metaphor that underlies all the interactions. Based on previous research findings on target pointing, I rigorously analyzed target pointing behavior when the targets are dynamically revealed by a moving display window ("peephole pointing"). I proposed a model for the movement time under such circumstances, taking into account the size of the display window (*S*):

$$T = a + b\left(n\log_2\left(\frac{A}{S}+1\right) + (1-n)\log_2\left(\frac{A}{W}+1\right)\right)$$

This model was validated by the experimental data under various settings including different cursor control mechanisms, and both with and without prior knowledge of the target location. This model can help us more deeply understand the impact of the flashlight metaphor, and guide future interaction designs for handheld projectors as well as other devices that utilize such a metaphor.

On the social application and behavior side, I created *Flashlight Jigsaw*, a multi-player public game played using handheld projectors, as a research conduit to investigate higher level social behaviors that emerge from handheld projector usage in a public space. I studied the game experiences and behaviors from both players and spectators through the game deployment in both a public atrium and a shared lab. The *Flashlight Jigsaw* game proved to be a successful example in designing games that leverage the special affordances of handheld projector, and blend into the experience of public spaces. Through the study findings, I have gained a detailed understanding of social behaviors around such public space games, which not only relates to handheld projectors, but may also generalize to games played on other public displays. I draw several design implications that could guide future development of handheld projector games and potentially other social applications.

In summary, my current work is a multi-faceted investigation of handheld projector interaction, and will provide the groundwork for future research and development in this area.

8.2 Limitations

Some limitations also exist regarding my current research, leaving space for further improvements and investigations:

The current prototype relies on high-precision tracking of the projector using a high-end motion tracking system, therefore restraining its ubiquitous usage with current technologies. With the rapid improvement of lightweight location sensing technologies such as indoor GPS [117] or self-contained portable tracking solutions such as TrackSense [68], this limitation is likely to be alleviated so that handheld projectors can be used in a wider variety of locations. The rationale of our current approach is to enable us to experiment our designs with the most mature technology, so that when tracking technologies become more cheaply and ubiquitously available, our research results can be readily applied. On the other hand, due to the possible imprecision and lag involved in these alternative tracking technologies, as well as the physical jitter caused by the hand holding the projector, the resulting projected image may appear jittery or unsynchronized. To address this potential problem, we will need to incorporate smoothing techniques to stabilize the input and output digitally (*e.g.*, by using temporal filtering techniques such as the Kalman filter [65]) and/or physically (*e.g.*, by using optical or mechanical image stabilizers [114]).

Secondly, when exploring the interaction techniques and usage scenarios, I focused my research on demonstrating the interaction concepts, and did not address how they are to be technically supported in a realistic mobile computing system. For these designs to be deployable for real usage, one key challenge is to seamlessly store, communicate, and broadcast data in and of the environment, both between the projector and the environment, and between different projectors. When multiple projectors are interacting with each other, it is important to consider how to establish the shared workspace between devices both efficiently and securely. These issues themselves deserve a thorough investigation from the systems perspective.

Regarding the modeling of user performance, the peephole pointing experiment used simulated apparatus with high quality tracking and display in order to ensure reliable data and exclude confounding factors. However, when interacting with a real handheld projector, many other elements come into play. The most notable is that the size and shape of the flashlight are constantly changing as the projector moves. My current model assumes a display window with a constant size, therefore needs to be revised to account for the more complex situation with handheld projectors. In addition, the hand movement to point a handheld projector to different places may be a combination of rotation and translation, different from the pure translation movement involved in the experiment. How different styles of physical movement would affect the user performance is worth investigating. In order to fully answer these questions, and to confidently generalize this model to various situations, experimental investigation using the real handheld projector prototype as well as other peephole display devices is desirable. Nonetheless, given the rigorous theoretical analysis that the model is based on, we have reasons to expect that it would be general enough to account for situations beyond those investigated in the current experiment. In addition, for the sake of simplicity we investigated a one-dimensional task. In realistic scenarios where targets can reside anywhere in a 2D space, the behavior of searching for the target is likely to be different from the linear search behavior that we have discussed, another important factor to be addressed in future exploration on 2D tasks.

The user evaluations of the interaction designs for both single and multiple users were relatively short and preliminary. In addition to the qualitative feedback I collected, quantitative controlled experiments might help us to better understand users' abilities to use each technique. We could also request users to perform higher-level tasks, so as to investigate how the individual techniques can be effectively used together. On the other hand, the 2-week deployment of *Flashlight Jigsaw* resulted in many valuable findings about the social behaviors regarding handheld projector usage, but might not be long enough to reveal phenomena that evolve over time, *e.g.*, the building of long-term social relationships between players. An extended deployment of the similar nature, ideally with real handheld projectors when they are technically ready, would help answer these questions.

8.3 Future Directions

I have explored a wide range of research issues related to handheld projector interaction. In addition to addressing the limitations I discussed, in each of the aspects many interesting questions remain open for further and deeper research in the future:

System Platform

Several extensions and improvements could be made to the current handheld projector prototype platform. One possibility is to attach a camera to the handheld projector, similar to that experimented by Beardsley et al. [13] and Raskar et al. [93]. Instead of relying on the camera to achieve image correction, the camera could be used as an additional input device that captures and stores information from the physical environment. The user can take a snapshot of a physical region or object of interest, which can be later projected into other physical environments, or processed by image recognition algorithms. By doing so, the boundary between the physical and virtual worlds is further blurred, as information can seamlessly transform back and forth between its digital and physical forms.

With simple computer vision techniques, the camera may also be used to analyze the attributes of the projection surface and environment, such as color, material, illumination, size or even simple geometry. By taking these parameters into account, the system could adjust accordingly to create the optimal projection quality and experience. Higher-level applications may also take them as auxiliary input to support a variety of interactions. For example, in a game application the game character may present different behaviors depending on the particular "terrain" it is projected on.

Facilitated by the attached camera and the visual patterns created by the handheld projector itself, we may apply more advanced computer vision technologies to assist tracking of the projector (similar to the TrackSense system [68]), therefore alleviating the requirement of an external tracking system. The projector-camera pair may also help recover the 3D geometry of the physical environment, by applying principles similar to those of 3D scanning [46]. By doing so, we may automate the workspace definition techniques, further lowering the overhead for using handheld projectors in new environments.

Our current system design assumes the workspaces and their corresponding surfaces remain static once they have been defined. Although in practice the user could easily redefine these surfaces if they have been rearranged, in a realistic scenario it may also be desirable for the system design to explicitly take into account the change of the environment over time. As the interactions expand into various environments and situations, it is also important to support workspaces embedded in moving surfaces (such as a handheld notepad) and non-planar surfaces (such as human bodies), which would require dynamic recovery and/or tracking of these surfaces. Finally, as technologies become mature, we will need to experiment embedding the projectors into current handheld devices such as mobile phones or PDAs, and address the relevant technical issues including power, data transmission, and alternative tracking technologies.

Interaction Design

Our interaction design explorations focused on the situation that the handheld projector is used as the primary channel of interaction, with the assistance of a passive pen in some cases. However, given the various environments that handheld projector are likely to be used in, it would be natural to consider how they can be complemented by other interaction manners when available.

The most obvious of such additional channels is the small screens currently embedded in most handheld devices, the very motivation for developing handheld projectors. Nonetheless, these embedded small displays, possibly touch-sensitive, are still valuable for activities not suited to large projected displays, especially those concerning personal or private information. They are also important for tasks that require precise input such as entering text, which are difficult by using freehand pointing. Other scenarios, such as collaborative activities or multi-player games, may benefit from using the projector and the embedded display simultaneously by distributing information and operations into different channels that suit them best. As a result, a careful design exploration is needed to integrate these two complementary interaction channels that will likely coexist on future handhelds.

Another possibility lies in the increasing prevalence of large-sized displays. Where such situated large displays are available, handheld projectors may be used to overlay additional content on top of the information already displayed on them, thus providing access to multiple layers of information. The combination of handheld projectors and static large displays may leverage the unique affordances of both technologies where applicable.

The majority of my explorations have been concentrated on lower-level interaction techniques designed to suit a large variety of potential applications, with conceptual demonstrations of possible usage scenarios supported by these techniques. Upon completion of this step, it is worthwhile to move on to identify and explore higher-level activities supported by handheld projectors, and take an iterative user-centered-design approach to create appropriate applications using these building blocks. Some of these potential applications may include personal information management, collaborative authoring, or educational activities. It is also interesting to consider how to adapt existing legacy applications to leverage the affordances of handheld projectors. Finally, for the multi-user interaction techniques I developed, it is important to consider how they may scale to a large number of concurrent users.

User Performance

User performance with handheld projectors is a complex research topic with many open questions to investigate. Our work on modeling pointing tasks under the Flashlight metaphor ("peephole pointing") is but the first step towards this direction. Within the same research framework, we can continue to experimentally study and model other generic tasks under the Flashlight metaphor, such as navigation, crossing [2], steering [1] or pointing at 2D targets with a 2D display window. In addition to task completion time, it is also important to look at the user errors involved, and investigate the speed-accuracy tradeoff in such tasks, as researchers [102, 136] have explored for tasks under traditional situations. Furthermore, guided by these research findings we could design and evaluate interaction and visualization techniques to improve user performance under the Flashlight metaphor, especially when the user works with multiple workspaces embedded in the physical environment.

In addition to the Flashlight metaphor, other characteristics of handheld projectors may affect the user performance as well. For example, to make informed design of interaction techniques, it is important to measure the precision and fatigue relating to freehand pointing in 3D space using the handheld projector, inspired by similar research on laser pointer interaction [81, 84]. On the other hand, the handheld projector creates an experience that the user interacts with multiple workspaces (display surfaces) positioned in the same physical environment, although not simultaneously. It is worthwhile to study users' abilities to navigate through these workspaces, or perform tasks that span multiple workspaces. The results would also make an interesting comparison with existing research on multi-display-surface environments [127].

Games and Social Interactions

Our exploration with the *Flashlight Jigsaw* game revealed valuable findings and implications regarding game experiences supported by handheld projectors, as well as the social interaction patterns that emerged from them. Handheld projectors seemed a successful platform for supporting games that facilitate interactions between people. Inspired by this, we may explore a larger variety of design alternatives of such games. *Flashlight Jigsaw* mainly exploits the dual display nature (public vs. personal) of handheld projectors. Future games could be designed to fully leverage the other affordances as well, especially by expanding the game experience to cover the entire physical environment, and more effectively exploiting the mobility aspect of handheld projectors. Examples may include a treasure hunting game, where players collaboratively discover treasures hidden in the physical world using their handheld projectors; or a Tamagotchi game where players raise virtual pets within their projections on different physical "terrains", and let them socialize or fight with each other when projections overlap. These games would be designed to fit into people's daily life patterns, and to be played in various physical environments. In order to further understand the relationship between these game experiences and the physical and social spaces, it is important to deploy them in spaces and locations of various natures, including outdoor spaces as the technologies become feasible. The questions I raised about "situated design" of such games may provide a good starting list of factors to consider and investigate for deployment in various spaces.

Games are only one example of social activities that could be supported by handheld projectors. To obtain a holistic understanding of social behaviors and protocols emerging from handheld projector usage, we could take a similar approach to research into other social applications. One example of these potential applications is education. The handheld projector may serve as a tool to help teachers and students to exchange knowledge either face-to-face or through the physical environment as a conduit, ideal for scenarios such as field trips. Compared to games, education provides a slightly more "serious", yet still flexible enough scenario for exploration. The interaction and communication patterns among the teachers and students would provide valuable insights for future designs of such applications.

Finally, instead of studying handheld projectors deployed for a specific application, it would certainly be interesting to distribute them to people for daily usage, as the technology becomes feasible. This would help us understand the real usage patterns and resulting social behaviors when the handheld projector eventually becomes available as an everyday personal device.

Bibliography

1. Accot, J. and Zhai, S. (1997). Beyond Fitts' Law: Models for trajectory-based HCI tasks. ACM CHI Conference on Human Factors in Computing Systems. p. 295-302.
2. Accot, J. and Zhai, S. (2002). More than dotting the i's - foundations for crossing-based interfaces. ACM CHI Conference on Human Factors in Computing Systems. p. 73-80.
3. Accot, J. and Zhai, S. (2003). Refining Fitts' law models for bivariate pointing. ACM CHI Conference on Human Factors in Computing Systems. p. 193-200.
4. Apitz, G. and Guimbretière, F. (2004). CrossY: a crossing-based drawing application. ACM UIST Symposium on User Interface Software and Technology. p. 3-12.
5. Azuma, R.T. (1997). A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments, 6(4). p. 355-385.
6. Baar, J.v., Willwacher, T., Rao, S., and Raskar, R. (2003). Seamless multi-projector display on curved screens. Eurographics Workshop on Virtual Environments (EGVE). p. 281-286.
7. Bajura, M., Fuchs, H., and Ohbuchi, R. (1992). Merging virtual objects with the real world: seeing ultrasound imagery within the patient. Proceedings of the 19th annual conference on Computer graphics and interactive techniques. p. 203-210.
8. Barfield, W. and Caudell, T., eds. (2001). Fundamentals of wearable computers and augmented reality. Lawrence Erlbaum Associates: Mahwah, New Jersey.

9. Barkhuus, L., Chalmers, M., Tennent, P., Hall, M., Bell, M., Sherwood, S., and Brown, B. (2005). Picking Pockets on the Lawn: The Development of Tactics and Strategies in a Mobile Game. UbiComp. p. 358-374.

10. Baudisch, P., Good, N., and Stewart, P. (2001). Focus plus context screens: combining display technology with visualization techniques. ACM UIST Symposium on User Interface Software and Technology. p. 431-440.

11. Baudisch, P. and Rosenholtz, R. (2003). Halo: a technique for visualizing off-screen objects. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 481-488.

12. Baudisch, P., Xie, X., Wang, C., and Ma, W.-Y. (2004). Collapse-to-zoom: viewing web pages on small screen devices by interactively removing irrelevant content. Proceedings of the 17th annual ACM symposium on User interface software and technology. p. 91-94.

13. Beardsley, P., Baar, J.V., Raskar, R., and Forlines, C. (2005). Interaction using a handheld projector. IEEE Computer Graphics and Applications, 25(1). p. 39-43.

14. Bell, B., Feiner, S., and Höllerer, T. (2001). View management for virtual and augmented reality. ACM UIST Symposium on User Interface Software and Technology. p. 101-110.

15. Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., Brown, B., Rowland, D., and Benford, S. (2006). Interweaving mobile games with everyday life. In Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM Press: Montreal, Quebec, Canada. p. 417-426.

16. Benford, S., Anastasi, R., Flintham, M., Drozd, A., Crabtree, A., Greenhalgh, C., Tandavanitj, N., Adams, M., and Row-Farr, J. (2003). Coping with Uncertainty in a Location-Based Game. IEEE Pervasive Computing, 2(3). p. 34-41.

17. Benford, S., Crabtree, A., Reeves, S., Sheridan, J., Dix, A., Flintham, M., and Drozd, A. (2006). Designing for the opportunities and risks of staging digital experiences in public settings. Proceedings of the SIGCHI conference on Human Factors in computing systems. p. 427-436.

18. Benford, S., Seager, W., Flintham, M., Anastasi, R., Rowland, D., Humble, J., Stanton, D., Bowers, J., Tandavanitj, N., Adams, M., Row-Farr, J., Oldroyd, A., and Sutton, J. (2004). The error of our ways: The experience of self-reported position in a location-based game. Ubicomp. p. 70-87.

19. Bier, E., Stone, M., Pier, K., Buxton, W., and DeRose, T. (1993). Toolglass and Magic Lenses: The see-through interface. ACM SIGGRAPH Conference on Computer Graphics and Interactive Techniques. p. 73-80.

20. Blaskó, G., Coriand, F., and Feiner, S. (2005). Exploring Interaction with a Simulated Wrist-Worn Projection Display. ISWC'05 IEEE International Symposium on Wearable Computers. p. 2-9.

21. Blaskó, G. and Feiner, S. (2004). An Interaction System for Watch Computers Using Tactile Guidance and Bidirectional Segmented Strokes. ISWC. p. 120-123.

22. Brignull, H., Izadi, S., Fitzpatrick, G., Rogers, Y., and Rodden, T. (2004). The introduction of a shared interactive surface into a communal space. In Proceedings of the 2004 ACM conference on Computer supported cooperative work. ACM Press: Chicago, Illinois, USA. p. 49-58.

23. Brignull, H. and Rogers, Y. (2003). Enticing people to interact with large public displays in public spaces. INTERACT Conference. p. 17-24.

24. Brown, B., Chalmers, M., Bell, M., Hall, M., MacColl, I., and Rudman, P. (2005). Sharing the square: collaborative leisure in the city streets. European Conf. Computer Supported Collaborative Work. p. 427-447.

25. Buxton, W. (1997). Living in augmented reality: Ubiquitous Media and Reactive Environments. In Video Mediated Communication, K. Finn, Sellen, A., Wilber, S., Editor. Erlbaum: Hillsdale, NJ. p. 363-384.

26. Cable, A.J., Buckley, E., Mash, P., Lawrence, N.A., Wilkinson, T.D., and Crossland, W.A. (2004). Real-time Binary Hologram Generation for High-quality Video Projection Applications. SID Symposium Digest of Technical Papers, 35(1). p. 1431-1433.

27. Cadiz, J.J., Gupta, A., and Grudin, J. (2000). Using Web annotations for asynchronous collaboration around documents. Proceedings of the 2000 ACM conference on Computer supported cooperative work. p. 309-318.

28. Cham, T.-J., Rehg, J.M., Sukthankar, R., and Sukthankar, G. (2003). Shadow Elimination and Occluder Light Suppression for Multi-Projector Displays. IEEE Conference on Computer Vision and Pattern Recognition,. p. 513-520.

29. Chang, M. and Goodman, E. (2004). FIASCO: game interface for location-based play. Proceedings of the 2004 conference on Designing interactive systems: processes, practices, methods, and techniques. p. 329-332.

30. Cooper, N., Keatley, A., Dahlquist, M., Mann, S., Slay, H., Zucco, J., Smith, R., and Thomas, B.H. (2004). Augmented Reality Chinese Checkers. Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology. p. 117-126.

31. Crabtree, A., Benford, S., Rodden, T., Greenhalgh, C., Flintham, M., Anastasi, R., Drozd, A., Adams, M., Row-Farr, J., Tandavanitj, N., and Steed, A. (2004). Orchestrating a mixed reality game 'on the ground'. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 391-398.

32. Crossman, E. (1956). The measurement of perceptual load in manual operations. University of Birmingham.

33. Dietz, P., Raskar, R., Booth, S., Baar, J.v., Wittenburg, K., and Knep, B. (2004). Multi-projectors and implicit interaction in persuasive public displays. In Proceedings of the working conference on Advanced visual interfaces. ACM Press: Gallipoli, Italy. p. 209-217.

34. Drascic, D. (1993). Stereoscopic Vision and Augmented Reality. Scientific Computing & Automation, 9(7). p. 31-34.

35. Ducheneaut, N., Yee, N., Nickell, E., and Moore, R.J. (2006). "Alone together?": exploring the social dynamics of massively multiplayer online games. Proceedings of the SIGCHI conference on Human Factors in computing systems. p. 407-416.

36. Edwards, W.K., Mynatt, E.D., Petersen, K., Spreitzer, M.J., Terry, D.B., and Theimer, M.M. (1997). Designing and implementing asynchronous collaborative applications with Bayou. Proceedings of the 10th annual ACM symposium on User interface software and technology. p. 119-128.

37. Feiner, S., MacIntyre, B., Haupt, M., and Solomon, E. (1993). Windows on the world: 2D windows for 3D augmented reality. Proceedings of the 6th annual ACM symposium on User interface software and technology. p. 145-155.

38. Feiner, S., MacIntyre, B., and Seligmann, D. (1993). Knowledge-based Augmented Reality. Communications of the ACM, 36(7). p. 52-62.

39. Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47. p. 381-391.

40. Fitzmaurice, G., Khan, A., Pieke, R., Buxton, B., and Kurtenbach, G. (2003). Tracking menus. ACM UIST Symposium on User Interface Software and Technology. p. 71-79.

41. Fitzmaurice, G.W. (1993). Situated information spaces and spatially aware palmtop computers. Communications of the ACM, 36(7). p. 38-49.

42. Flagg, M. and Rehg, J.M. (2006). Projector-guided painting. Proceedings of the 19th annual ACM symposium on User interface software and technology. p. 235-244.

43. Flintham, M., Benford, S., Anastasi, R., Hemmings, T., Crabtree, A., Greenhalgh, C., Tandavanitj, N., Adams, M., and Row-Farr, J. (2003). Where on-line meets on the streets: experiences with mobile mixed reality games. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 569-576.

44. Forlines, C., Balakrishnan, R., Beardsley, P., Baar, J.v., and Raskar, R. (2005). Zoom-and-pick: facilitating visual zooming and precision pointing with interactive handheld projectors. In Proceedings of the 18th annual ACM symposium on User interface software and technology. ACM Press: Seattle, WA, USA. p. 73-82.

45. Forsyth, D. and Ponce, J. (2003). Computer Vision -- A modern approach: Prentice Hall.

46. Franca, J.G.D.M., Gazziro, M.A., Ide, A.N., and Saito, J.H. (2005). A 3D scanning system based on laser triangulation and variable field of view. IEEE International Conference on Image Processing. p. 425-428.

47. Furnas, G. (1986). Generalized fisheye views. ACM CHI Conference on Human Factors in Computing Systems. p. 16-23.

48. Ganoe, C.H., Somervell, J.P., Neale, D.C., Isenhour, P.L., Carroll, J.M., Rosson, M.B., and McCrickard, D.S. (2003). Classroom BRIDGE: using collaborative public and desktop timelines to support activity awareness. Proceedings of the 16th annual ACM symposium on User interface software and technology. p. 21-30.

49. Greenberg, S. and Rounding, M. (2001). The notification collage: posting information to public and personal displays. ACM CHI Conference on Human Factors in Computing Systems. p. 514-521.

50. Grossman, T. and Balakrishnan, R. (2004). Pointing at trivariate targets in 3D environments. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 447-454.

51. Grossman, T. and Balakrishnan, R. (2005). A probabilistic approach to modeling two-dimensional pointing. ACM Transactions on Computer Human Interaction, 12(3). p. 435-459.

52. Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. Journal of Motor Behavior, 19(4). p. 486-517.

53. Hernandez, W.S. (2003). Micro Laser Personal Projector. Master of Science Thesis, Massachusetts Institute of Technology.

54. Hilliges, O. and Terrenghi, L. (2006). Overcoming mode-changes on multi-user large displays with bi-manual interaction. MU3I Workshop on Multi-User and Ubiquitous User Interfaces.

55. Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P., and Smith, M. (2004). Stitching: Pen gestures that span multiple displays. ACM AVI Conference on Advanced Visual Interfaces. p. 23-31.

56. Holcomb, M.O., Mueller-Mach, R., Mueller, G.O., Collins, D., Fletcher, R.M., Steigerwald, D.A., Eberle, S., Lim, Y.K., Martin, P.S., and Krames, M. (2003). The LED lightbulb: are we there yet? progress and challenges for solid state illumination. Conference on Lasers and Electro-Optics. p. 4-7.

57. Holmgren, D.E. (1992). Design and Construction of a 30-Degree See-Through Head-Mounted Display, UNC Chapel Hill Department of Computer Science technical report TR 92-030.

58. Huang, E.M. and Mynatt, E.D. (2003). Semi-public displays for small, co-located groups. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 49-56.

59. Inkpen, K.M., Ho-Ching, W.-l., Kuederle, O., Scott, S.D., and Shoemaker, G.B.D. (1999). This is fun! we're all best friends and we're all playing: supporting children's synchronous collaboration. Proceedings of the 1999 conference on Computer support for collaborative learning. p. 31.

60. Intille, S.S. (2002). Change Blind Information Display for Ubiquitous Computing Environments. the Fourth International Conference on Ubiquitous Computing. p. 91?06.

61. Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. ACM CHI Conference on Human Factors in Computing Systems. p. 234-241.

62. Izadi, S., Brignull, H., Rodden, T., Rogers, Y., and Underwood, M. (2003). Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media. ACM UIST Symposium on User Interface Software and Technology. p. 159-168.

63. Joffe, B. (2005). Mogi: Location and Presence in a Pervasive Community Game. Ubicomp Workshop on Ubiquitous Gaming and Entertainment.

64. Kabbash, P. and Buxton, W. (1995). The "Prince" technique: Fitts' law and selection using area cursors. ACM CHI Conference on Human Factors in Computing Systems. p. 273-279.

65. Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82(1). p. 35–45.

66. Kanayama, H., Maeda, M., Miwa, T., Ikeda, T., Murata, H., and Chihara, K. (2006). Ultra small projector with high efficiency illumination system. International Conference on Consumer Electronics. p. 127-128.

67. Kim, W.S. (1996). Virtual Reality Calibration and Preview / Predictive Displays for Telerobotics. Presence: Teleoperators and Virtual Environments, 5(2). p. 173-190.

68. Köhler, M., Patel, S.N., Summet, J.W., Stuntebeck, E.P., and Abowd, G.D. (2007). TrackSense: Infrastructure Free Precise Indoor Positioning Using Projected Patterns. Pervasive. p. 334-350.

69. Lee, J.C., Hudson, S.E., Summet, J.W., and Dietz, P.H. (2005). Moveable interactive projected displays using projector based tracking. Proceedings of the 18th annual ACM symposium on User interface software and technology. p. 63-72.

70. MacKenzie, S. (1992). Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction, 7. p. 91-139.

71. MacKenzie, S. and Buxton, W. (1992). Extending Fitts' law to two-dimensional tasks. ACM CHI Conference on Human Factors in Computing Systems. p. 219-226.

72. Maes, P. (1993). Artificial Life Meets Entertainment: Lifelike Autonomous Agents. Communications of the ACM, 38(11). p. 108-114.

73. Magerkurth, C., Cheok, A.D., Mandryk, R.L., and Nilsen, T. (2005). Pervasive games: bringing computer entertainment back to the real world. Comput. Entertain., 3(3). p. 4-4.

74. Mandryk, R.L. and Inkpen, K.M. (2004). Physiological indicators for the evaluation of co-located collaborative play. Proceedings of the 2004 ACM conference on Computer supported cooperative work. p. 102-111.

75. Matveyev, S., Göbel, M., and Frolov, P. (2003). Laser pointer interaction with hand tremor elimination. HCI International. p. 376-740.

76. Matysczok, C., Radkowski, R., and Berssenbruegge, J. (2004). AR-bowling: immersive and realistic game play in real environments using augmented reality. Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology. p. 269-276.

77. McGuffin, M. and Balakrishnan, R. (2002). Acquisition of expanding targets. ACM CHI Conference on Human Factors in Computing Systems. p. 57-64.

78. Meyer, D.E., Smith, J.E., Kornblum, S., Abrams, R.A., and Wright, C.E. (1988). Optimality in human motor performance: Ideal control of rapid aimed movements. Psychological Review, 95. p. 340-370.

79. Morris, M.R., Paepcke, A., Winograd, T., and Stamberger, J. (2006). TeamTag: exploring centralized versus replicated controls for co-located tabletop groupware. Proceedings of the SIGCHI conference on Human Factors in computing systems. p. 1273-1282.

80. Morris, M.R., Ryall, K., Shen, C., Forlines, C., and Vernier, F. (2004). Beyond "social protocols": multi-user coordination policies for co-located groupware. In Proceedings of the 2004 ACM conference on Computer supported cooperative work. ACM Press: Chicago, Illinois, USA. p. 262-265.

81. Myers, B., Bhatnagar, R., Nichols, J., Peck, C.H., Kong, D., Miller, R., and Long, C. (2002). Interacting at a distance: measuring the performance of laser pointers and other devices. ACM CHI Conference on Human Factors in Computing Systems. p. 33-40.

82. Olsen, D.R. and Nielsen, T. (2001). Laser pointer interaction. ACM CHI Conference on Human Factors in Computing Systems. p. 17-22.

83. Patel, S.N., Rekimoto, J., and Abowd, G.D. (2006). iCam: Precise at-a-distance Interaction in the Physical Environment. Pervasive 2006. p. 272 – 287.

84. Peck, C. (2001). Useful parameters for the design of laser pointer interaction techniques. Extended Abstracts of the ACM CHI Conference on Human Factors in Computing Systems. p. 461-462.

85. Pingali, G., Pinhanez, C., Levas, A., Kjeldsen, R., Podlaseck, M., Chen, H., and Sukaviriya, N. (2003). Steerable Interfaces for Pervasive Computing Spaces. PerCom'03 IEEE International Conference on Pervasive Computing and Communications. p. 315--322.

86. Pinhanez, C. (2001). The Everywhere Displays Projector: A device to create ubiquitous graphical interfaces. Ubicomp, Conference on Ubiquitous Computing. p. 315--331.

87. Podlaseck, M., Pinhanez, C., Alvarado, N., Chan, M., and Dejesus, E. (2003). On interfaces projected onto real-world objects. CHI '03 extended abstracts on Human factors in computing systems. p. 802-803.

88. Preguica, N., Martins, J.L., Domingos, H., and Duarte, S. (2000). Data management support for asynchronous groupware. Proceedings of the 2000 ACM conference on Computer supported cooperative work. p. 69-78.

89. R. Sukthankar, R.S., and M. Mullin. (2000). Automatic Keystone Correction for Camera-Assisted Presentation Interfaces. International Conference on Multimedia Interfaces. p. 607-614.

90. Rapp, S., Michelitsch, G., Osen, M., Williams, J., Barbisch, M., Bohan, R., Valsan, Z., and Emele, M. (2004). Spotlight Navigation: Interaction with a

handheld projection device. Pervasive International Conference on Pervasive Computing, Video paper.

91. Raskar, R., Baar, J.v., Beardsley, P., Willwacher, T., Rao, S., and Forlines, C. (2003). iLamps: geometrically aware and self-configuring projectors. ACM Transactions on Graphics, 22(3).

92. Raskar, R. and Beardsley, P. (2001). A Self-Correcting Projector. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition. p. 504-508.

93. Raskar, R., Beardsley, P., Baar, J.v., Wang, Y., Dietz, P., Lee, J., Leigh, D., and Willwacher, T. (2004). RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors. ACM Transactions on Graphics, 23(3). p. 406-415.

94. Reeves, S., Benford, S., O'Malley, C., and Fraser, M. (2005). Designing the spectator experience. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 741-750.

95. Reid, J., Hyams, J., Shaw, K., and Lipson, M. (2004). "Fancy a Schmink?": a novel networked game in a cafe. Computers in Entertainment., 2(3). p. 11-11.

96. Reingold, H. (1992). Virtual Reality: Touchstone Books.

97. Rekimoto, J. and Nagao, K. (1995). The world through the computer: computer augmented interaction with real world environments. Proceedings of the 8th annual ACM symposium on User interface and software technology. p. 29-36.

98. Robbins, D.C., Cutrell, E., Sarin, R., and Horvitz, E. (2004). ZoneZoom: map navigation for smartphones with recursive view segmentation. Proceedings of the working conference on Advanced visual interfaces. p. 231-234.

99. Rose, E., Breen, D., Ahlers, K., Crampton, C., Tuceryan, M., Whitaker, R., and Greer, D. (1995). Annotating Real-World Objects Using Augmented Reality. Computer Graphics International. p. 357-370.

100. Sakamoto, Y. and Kuwana, E. (1993). Toward integrated support of synchronous and asynchronous communication in cooperative work: an empirical study of real group communication. Proceedings of the conference on Organizational computing systems. p. 90-97.

101. Scheible, J. and Ojala, T. (2005). MobiLenin combining a multi-track music video, personal mobile phones and a public display into multi-user interactive entertainment. Proceedings of the 13th annual ACM international conference on Multimedia. p. 199-208.

102. Schmidt, R.A., Zalaznik, H.N., Hawkins, B., Frank, J.S., and Quinn, J.T. (1979). Motor output variability: A theory for the accuracy of rapid motor acts. Psychological Review, 86. p. 415-451.

103. Scott, S.D., Sheelagh, M., Carpendale, T., and Inkpen, K.M. (2004). Territoriality in collaborative tabletop workspaces. Proceedings of the 2004 ACM conference on Computer supported cooperative work. p. 294-303.

104. Shen, C., Vernier, F., Forlines, C., and Ringel, M. (2004). DiamondSpin: An extensible toolkit for around the table interaction. ACM CHI Conference on Human Factors in Computing Systems. p. 167-174.

105. Shoemaker, G. and Inkpen, K. (2001). Single display privacyware: Augmenting public displays with private information. ACM CHI Conference on Human Factors in Computing Systems. p. 522-529.

106. Simon, A. (2005). First-person experience and usability of co-located interaction in a projection-based virtual environment. In Proceedings of the ACM symposium on Virtual reality software and technology. ACM Press: Monterey, CA, USA. p. 23-30.

107. Sims, D. (1994). New Realities in Aircraft Design and Manufacture. IEEE Computer Graphics and Applications, 14(2). p. 91.

108. State, A., Livingston, M.A., Garrett, W.F., Hirota, G., Whitton, M.C., Pisano, E.D., and Fuchs, H. (1996). Technologies for augmented reality systems: realizing ultrasound-guided needle biopsies. Proceedings of the 23rd annual conference on Computer graphics and interactive techniques. p. 439-446.

109. Sugimoto, M., Miyahara, K., Inoue, H., and Tsunesada, Y. (2005). Hotaru: Intuitive Manipulation Techniques for Projected Displays of Mobile Devices. INTERACT 2005. p. 57-68.

110. Sukaviriya, N., Podlaseck, M., Kjeldsen, R., Levas, A., Pingali, G., and Pinhanez, C. (2003). Embedding Interactions in a Retail Store Environment: The Design and Lessons Learned. International Conference on Human-Computer Interaction (INTERACT).

111. Sukthankar, R., Stockton, R.G., and Mullin, M.D. (2001). Smarter presentations: Exploiting homography in camera-projector systems. International Conference on Computer Vision. p. 247-253.

112. Symbol_Technologies. LPD: A miniature, high-resolution projection engine. http://www.symbol.com/assets/files/WP-LPD.pdf.

113. Tan, D.S. and Pausch, R. (2002). Pre-emptive shadows: eliminating the blinding light from projectors. CHI '02 extended abstracts on Human factors in computing systems. p. 682-683.

114. Teare, S.W. and Restaino, S.R. (2006). Introduction to Image Stabilization: SPIE Press.

115. Teller, S., Chen, J., and Balakrishnan, H. (2003). Pervasive pose-aware applications and infrastructure. IEEE Computer Graphics andApplications, 23(4). p. 14-18.

116. Tsang, M., Fitzmaurice, G.W., Kurtenbach, G., Khan, A., and Buxton, B. (2002). Boom chameleon: simultaneous capture of 3D viewpoint, voice and gesture

annotations on a spatially-aware display. ACM UIST Symposium on User Interface Software and Technology. p. 111-120.

117. van Diggelen, F. (2002). Indoor GPS theory & implementation. IEEE Position Location and Navigation Symposium. p. 240- 247.

118. Vogel, D. and Balakrishnan, R. (2004). Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users. ACM UIST Symposium on User Interface Software and Technology. p. 137-146.

119. Vogel, D. and Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. Proceedings of the 18th annual ACM symposium on User interface software and technology. p. 33-42.

120. Voida, S., Podlaseck, M., Kjeldsen, R., and Pinhanez, C. (2005). A study on the manipulation of 2D objects in a projector/camera-based augmented reality environment. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 611-620.

121. Wagner, D. and Schmalstieg, D. (2003). First Steps Towards Handheld Augmented Reality. p. 127-135.

122. Wanstall, B. (1989). HUD on the Head for Combat Pilots. Interavia, 44. p. 334-338.

123. Welford, A. (1968). The fundamentals of skill. London: Methuen.

124. Weng, C. and Gennari, J.H. (2004). Asynchronous collaborative writing through annotations. Proceedings of the 2004 ACM conference on Computer supported cooperative work. p. 578-581.

125. Whyte, W.H. (1980). The Social Life of Small Urban Spaces. Washington D.C.: Project for Public Spaces.

126. Wigdor, D., Leigh, D., Forlines, C., Shipman, S., Barnwell, J., Balakrishnan, R., and Shen, C. (2006). Under the table interaction. Proceedings of the 19th annual ACM symposium on User interface software and technology. p. 259-268.

127. Wigdor, D., Shen, C., Forlines, C., and Balakrishnan, R. (2007). Perception of elementary graphical elements in tabletop and multi-surface environments. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 473-482.

128. Wolfe, J.M. (1998). Visual Search. In Attention, H. Pashler, Editor. University College London Press: London, UK.

129. Wu, M. and Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. ACM UIST 2003 Symposium on User Interface Software and Technology. p. 193-202.

130. Yee, K.-P. (2003). Peephole displays: pen interaction on spatially aware handheld computers. Proceedings of the SIGCHI conference on Human factors in computing systems. p. 1-8.

131. Yokokohji, Y., Sugawara, Y., and Yoshikawa, T. (2000). Accurate Image Overlay on Video See-Through HMDs Using Vision and Accelerometers. Proceedings of the IEEE Virtual Reality 2000 Conference. p. 247.

132. Yoon, J., Oishi, J., Nawyn, J., Kobayashi, K., and Gupta, N. (2004). FishPong: encouraging human-to-human interaction in informal social environments. Proceedings of the 2004 ACM conference on Computer supported cooperative work. p. 374-377.

133. Zajonc, R.B. (1965). Social facilitation. Science, 149. p. 269-274.

134. Zalevsky, E.S.Z., Kapelner, Y., and Kapelner., S. Virtual display with low power consuming, portable micro projector. SPIE. p. 154-163.

135. Zellweger, P.T., Mackinlay, J.D., Good, L., Stefik, M., and Baudisch, P. (2003). City lights: contextual views in minimal space. CHI '03 extended abstracts on Human factors in computing systems. p. 838-839.

136. Zhai, S., Kong, J., and Ren, X. (2004). Speed-accuracy trade-off in Fitts' law tasks - on the equivalency of actual and nominal pointing precision. International Journal of Human-Computer Studies, (Special Issue of Fitts (1954) 50th Anniversary).