Spatial Interaction with Mobile Projected Displays

Dissertation zur Erlangung des Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing.)
der Naturwissenschaftlich-Technischen Fakultät I
der Universität des Saarlandes

vorgelegt von
(Dipl-Geoinf.) Markus Löchtefeld
Saarbrücken
Juni 2014

UNIVERSITÄT
DES
SAARLANDES

Dean:
Prof. Dr. Markus Bläser

Head of Committee:
Prof. Dr. Andreas Zeller

Reviewers:
Prof. Dr. Antonio Krüger
Prof. Dr. Jonna Häkkilä

Committee Member:
Dr. Sven Gehring

Day of the defense:
14th of August 2014

Abstract

The miniaturization of projection technology recently reached a point where battery powered pico-projectors can be carried in ones pocket. These projectors not only allow for mobile stand-alone devices but also the integration into other devices. These new class of devices allow to create ad-hoc mobile projected displays everywhere. Until today this development is mainly technology driven. While projector phones – smartphones with integrated projector – are available, the fact that they are mainly hand-held devices, present a variety of challenges for application and interaction techniques, for the mobile projected display they provide.

In this thesis we enable spatial interaction with mobile projected displays to overcome the current technological driven approaches. We contribute by investigating the following three different directions: Applications exploiting spatial features; Analysis of spatial alignment of mobile projected displays; and Interaction techniques exploiting spatial memory techniques. Through the development and evaluation of applications and interaction techniques that rely on spatial features we exploit the users spatial memory and allow mobile projected displays to become a valuable addition of our current mobile devices.

Kurzfassung

Die Miniaturisierung der Projektor Technologie hat kürzlich ein Level erreicht, das es ermöglicht batteriebetriebene Pico-Projektoren in der Hosentasche zu transportieren. Diese Projektoren können nicht nur als eigenständiges Gerät existieren, sondern auch in andere Geräte integriert werden. Diese neue Geräteklasse erlaubt es überall ad-hoc mobile Displays zu projizieren. Bis zum jetzigen Zeitpunkt ist diese Entwicklung allerdings hauptsächlich technologiegetrieben. Auch wenn Projector Phones – Smartphones mit integriertem Projektor – verfügbar sind, die Tatsache das es sich um in der Hand gehaltene Geräte handelt , resultiert in einer Reihe von Problemen für Applikationen und Interaktionstechniken für die mobilen projizierten Displays die diese erzeugen.

In dieser Arbeit ermöglichen wir Räumliche Interaktion mit mobilen projizierten Displays um die Probleme der aktuell technologiegetriebenen Ansätze zu überwinden. Dazu tragen wir durch die Untersuchung der folgenden drei Richtungen bei: Applikationen die räumliche Merkmale verwenden; Analyse der räumlichen Ausrichtung von mobilen projizierten Displays; und Interaktionstechniken die das räumliche Gedächtnis benutzen. Durch die Entwicklung und Evaluierung von Applikationen und Interaktionstechniken die auf räumlichen Merkmalen basieren, kann das räumliche Gedächtnis des Benutzers ausgenutzt werden, so dass mobile projizierte Displays eine wertvolle Ergänzung zu den aktuellen Mobilgeräten werden können.

"I have never listened to anyone who criticized my taste in space travel, sideshows or gorillas. When this occurs, I pack up my dinosaurs and leave the room." Ray Bradbury [31]

Für Stine und Jakob.

List of Publications

Parts of the work presented in this thesis, including figures, tables and text fragments have been published in the following publications. The chapters of this dissertation are partly based on these publications.

Chapter 4

- [132] Markus Löchtefeld, Johannes Schöning, Michael Rohs and Antonio Krüger: *Marauders Light: Replacing the Wand with a Mobile Camera Projector Unit.* MUM 2009: The 8th International Conference on Mobile and Ubiquitous Multimedia, (2009)

- [131] Markus Löchtefeld, Johannes Schöning, Michael Rohs and Antonio Krüger: *LittleProjectedPlanet: An Augmented Reality Game for Camera Projector Phones.* Mobile HCI 2009: Workshop on Mobile Interaction with the Real World MIRW, (2009)

- [188] Johannes Schöning, Michael Rohs, Sven Kratz, Markus Löchtefeld, and Antonio Krüger: *Map Torchlight: A Mobile Augmented Reality Camera Projector Unit.* CHI 2009: Adjunct Proceedings of the 27th International Conference on Human Factors in Computing Systems, (2009)

- [128] Markus Löchtefeld, Sven Gehring, Johannes Schöning and Antonio Krüger: *ShelfTorchlight: Augmenting a Shelf using a Camera Projector Unit.* UBIPROJECTION 2010. In Proceedings of the 1st workshop on personal projection, (2010)

- [133] Markus Löchtefeld, Johannes Schöning, Michael Rohs and Antonio Krüger: *Projector Phones: A new Class of Interfaces for Augmented Reality.* IJMHCI:

International Journal of Mobile Human Computer Interaction (IJMHCI) Issue 2 (3). (2010)

- [122] Markus Löchtefeld, Sven Gehring, Ralf Jung and Antonio Krüger: *guitAR - Supporting Guitar Learning through mobile Projection.* CHI 2011: Adjunct Proceedings of the 29th International Conference on Human Factors in Computing Systems, (2011)

- [123] Markus Löchtefeld, Sven Gehring, Ralf Jung and Antonio Krüger: *Using Mobile Projection to support Guitar Learning.* Smart Graphics 2011: Proceedings of the 11th International Symposium on Smart Graphics, (2011)

- [130] Markus Löchtefeld, Michael Rohs and Antonio Krüger: *Mobile Projection Interfaces for Augmented Reality Applications.* Book Chapter in Handbook of Augmented Reality (Springer, Editor: Borko Furth) (2012)

Chapter 5

- [43] Jessica Cauchard, Markus Löchtefeld, Pourang Irani, Johannes Schöning, Mike Fraser, Antonio Krüger and Sriram Subramanian: *Visual Separation in Mobile Multi-Display Environments.* UIST 2011: Proceedings of the 24th ACM Symposium on User Interface Software and Technology, (2011)

- [220] Christian Winkler, Markus Löchtefeld, David Dobbelstein, Antonio Krüger and Enrico Rukzio: *SurfacePhone: A Mobile Projection Device for Single- and Multiuser Everywhere Tabletop Interaction.* CHI 2014: Proceedings of SIGCHI Conference on Human Factors in Computing Systems (2014)

Chapter 6

- [41] Jessica Cauchard, Markus Löchtefeld, Antonio Krüger, Mike Fraser and Sriram Subramanian: *m+pSpaces: Virtual Workspaces in the Spatially-Aware Mobile Environment.* MobileHCI 2012: Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services, (2012)

- [124] Markus Löchtefeld, Sven Gehring and Antonio Krüger: *Distributed Dynamic Peephole Interfaces for Mobile Map Navigation on Projector Phones.* CHI 2011: DUI 2011 - Workshop on Distributed User Interfaces, (2011)

- [125] Markus Löchtefeld, Sven Gehring and Antonio Krüger: *Distributed User Interfaces for Projector Phones.* Book Chapter in Distributed User Interfaces: Designing Interfaces for the Distributed Ecosystem (Springer, Editors: Jose A Gallud, Ricardo Tesoriero, Victor M.R. Penichet) (2012)

Chapter 7

- [126] Markus Löchtefeld, Sven Gehring and Antonio Krüger: *Proxemic Interaction for Projector Phones.* CHI 2011: MP2 - Workshop on Mobile and Personal Projection, (2011)

- [121] Markus Löchtefeld and Michael Braun: *Towards a mobile Everywhere Display based on a depth-camera projector unit.* MobileHCI 2012 - MobiVis - Mobile Vision and HCI Workshop (2012)

- [55] Raimund Dachselt, Jonna Häkkilä, Matt Jones, Markus Löchtefeld, Michael Rohs and Enrico Rukzio: *Pico Projectors – Firefly or Bright Future?* ACM Interactions March/April 2012, (2012)

- [129] Markus Löchtefeld and Antonio Krüger: *Evaluation of an In-Call Interaction Technique for Projector Phones.* ECIS 2013: Proceedings of the European Conference on Information Systems (2013)

Contents

1

Introduction

1.1 Motivation

In recent years mobile phones transformed into multi-purpose devices. These so called smartphones allow the user to install new applications and are able to solve complicated computational problems while being on the go. In terms of memory and computational power they are already quite close to their desktop counterparts. Besides their computational power they also incorporate advanced sensors such as accelerometers, gyroscopes, GPS receivers, and high-resolution cameras that are not prominent in today's desktop computers. Especially in terms of display resolution and pixel density current high-end smartphones even outperform desktop computers and notebooks (e.g. the 2013 Apple MacBook Pro Retina 13" has with 227ppi [94] a lot smaller pixel density compared to the Samsung Galaxy S4 with 441 ppi [77]). Even though these displays have the potential to display a lot of content at the same time, the information displayed needs to retain a certain size to stay perceivable. This also marks the biggest drawback of current smartphones. To maintain a small and mobile form-factor the maximum display-size is limited and with that the amount of information that can be displayed as well. The current high-end smartphones reach display sizes of up to five inches. This reflects the demand of the users for bigger displays that allow to explore more content at the same time. The display size can not increase much more because otherwise the device would loose its mobility and would be considered a tablet. These are, due to their size, mainly used at home [153]. Smartphones are devices that can access nearly all of mankind's information but are not able to display them properly.

This leaves the user with the need to scroll, pan and zoom to holistically explore the information on the display.

To overcome these problems researchers have developed several techniques to ease the process of panning and zooming [68, 120, 135]. Besides these interaction-based approaches other technical solutions might even be more promising. One approach would be to increase the accuracy of the touch screen using user-specific touch models [211]. This allows the user to interact with the smallest perceivable object on the screen accurately. But still it would not solve the problem of the minimum size that is needed to allow the content to be perceivable. Increasing the screen real estate using multiple displays that are connected, e.g. with hinges, is one possibility. The Kyocera Echo [51] is such a multi-display device. It connects two 3.5 inch screens with hinges in such a way that it is small and pocketable when folded together but can be transformed to a large screen device as well. Flexible Displays that are bendable, rollable, or even foldable would allow a mobile device to be small while having a large screen [110, 181, 209]. On the one side this flexibility can be used as a mean of input [110, 115, 199, 209] as well as output [73, 181]. On the other side this flexibility also comes with the drawback of not being rigid, which can be straining especially for larger display sizes. The flexibility might require more than one hand to hold the device which leaves the user with no hand free for interacting with it. Furthermore flexible display technology is still experimental and while flexible-OLED displays still produce to much heat to be usable as phone displays, flexible E-Ink displays have very slow refresh rate of up to 2hz [73] which is not suitable for interactive media such as movies or games.

Not only the field of mobile phones has made several advantages in the last years. Projection technology also made significant progress. Besides becoming cheaper and more powerful, today's LED based projectors became so energy efficient that they can be operated using a battery. In combination with the advances in miniaturization, projectors have become small enough so that they can be carried in one's pocket. These so called pico-projectors not only allow for mobile stand-alone projectors but also the integration into other devices. This possibility sparked the development of new classes of devices that allow to create ad-hoc displays everywhere. The range of devices with integrated projectors include camcorders [52], tablets [59], phones [76], and even toys

Figure 1.1: Samsung Galaxy Beam - The Samsung Galaxy Beam was one of the first commercially available projector phones [76].

that create multi-touch enabled displays for gaming [66]. PMA Research predicts that the in 2015, 58 million pico projector units will be shipped [4].

The combination of a pico-projector and a smartphone - also called projector phone (compare Figure 1.1) - allows to generate a projected display with increased size, everywhere while being on the go. Every brighter surface can be utilized to project on it and with that allows the user to create a personal display in varying surrounding and for different use cases.

It has been argued that this technology is only a transitional solution to a reality with ubiquitous display wallpapers where the content can be controlled by e.g. the phone of the user [50]. An increasing amount of displays will over time become part of our everyday life including unwanted effects such as advertising [154]. It is expected that in 2015, 22 million public displays will be deployed [195] but it is questionable whether we ever reach a future in which we will have display wallpapers everywhere. Firstly we are limited by already existing artefacts and buildings that we don't want to, or are not allowed to change (e.g. historical buildings). Secondly natural resources might be too limited to allow for such a holistic deployment. There will always be situations where we either don't want to install displays permanently or where we have technical limitations that simply prevent such installation.

With all the above mentioned limitations of foldable or rollable displays and natural resources being to sparse for a holistic deployment of display wallpapers, mobile projection is a very valid alternative to create interactive displays while being on the go. One extra ordinary example that highlights the capabilities and the flexibility of projector phones is the case of the 33 Chilean miners that on August 5th 2010 got trapped in a copper mine for over two months. Due to the very small connection tube that was used to exchange food and water with the stuck miners it was not possible send down large electronic gadgets. Therefore Samsung send down an early Galaxy Beam [76] Prototype that contained a recording of the recent World Cup game of Chile against Ukraine [65]. Of course the dark environment of the mine reduced some of the limitations of current mobile projectors, such as low brightness, but still the small device allowed 33 miners to watch the soccer game on a large sized screen.

Mobile projection technology will have and already has had a major impact on the field of Ubiquitous Computing. The paradigm of Ubiquitous Computing was first defined by Mark Weiser and describes the next wave of post desktop computing devices [212], in which the computing power is slowly integrated into our everyday objects. As Weiser formulates it, these computers "... weave themselves into the fabric of everyday life until they are indistinguishable from it" [212]. As we already reached this paradigm, we interact with a variety of hidden computers already today, e.g. in our car. According to Weiser we will see new and different device classes emerging. He separates between three different sizes of devices. The first one, called Tab, is a hand-held tiny computer [208]. The second one, which he calls Pad, is a computer with a display in the size of a normal notepad [103]. And lastly Weiser envisioned a large scale interactive surface class called Board [63]. To explore these concepts in more depth Weiser and his team developed prototypes of these classes during the mid-90s which can be seen in Figure 1.2. As of today such device classes are now all commercially available: current Smartphones represent the Tab class; Tablet computers such as the iPad resemble the Pad class and with the increasing number of large-scale public displays becoming available the Board category is represented in our everyday life as well. Beside technical challenges in creating these devices the main problem lies in the interaction with them. While for Tab and Pad sized devices touch interfaces and stylus input have been proven to be efficient and useful, the ideal input for Board sized devices is rather unclear and a lot of research lately especially focuses on combining smartphones and large scale displays [28, 29].

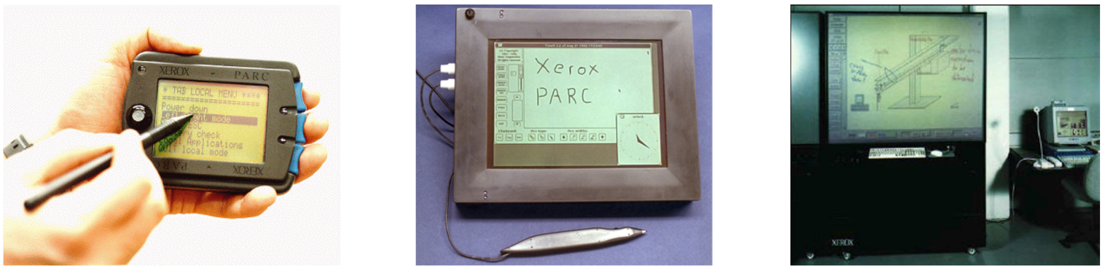


Figure 1.2: Xerox PARC Tab, Pad & Board - The devices classes developed by Xerox PARC ranged from: Tab-Sized (left) [208], over Pad-Sized (center) [103], to Board-Sized (right) [63]. Images taken from [156].

The most interesting part about Weisers device classes is that they can not be applied to projector phones. On the one side this devices resembles a hand-held device in the size of a Tab, but on the other side the integrated projector seamlessly allows to create ad-hoc displays in Pad- or Board-size if needed. Depending on distance and available projection surface a projector phone can blend between different display sizes and with that between different device classes. Naturally this comes with the need for novel interfaces. A direct-touch based user interface would for example not be possible if the projection is out of the users reach. Therefore we need to explore new interaction techniques that take the distance between the user and the projection as well as the environmental surroundings into account.

1.2 History of Interaction with Projection Devices

Mobile projection has been employed for entertainment purposes since the 17th century [83]. The *laterna magica* or magic lantern, first mentioned in 1646 [111], was a fully mobile device that was able to project images on to surfaces, therefore it is the rightful predecessor of current pico projectors. A small mirror behind an ambient light source (in the past mainly candles, later also oil lamps and electric light bulbs), allowed to focus the light through a transparent film towards a magnifying lens which then allowed to project the cone of light onto a nearby surface. The content that is on the transparent film is then magnified and projected onto the surface. Since all these elements were contained in one single device (compare Figure 1.3) that could be held by hand or worn on a belt [214], the *laterna magica* can be seen as the rightful predecessor of todays' pico projectors. If the drawbacks, such as the generated heat and its weight, are neglected, they already provided many of the same affordances of current mobile projectors as well.

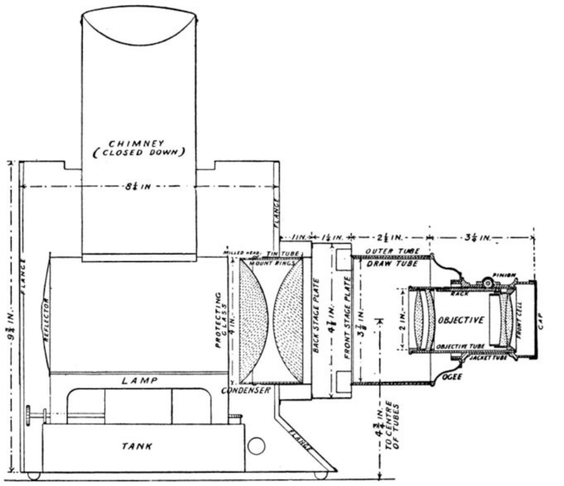


Figure 1.3: ***Laterna Magica*** - Blue print of a *laterna magica*. The light of the lamp is bundled in the mirror and creates a light cone through the transparent film and projects it contains onto a surface [83].

Besides the possibility to project on a flat surface, often also smoke or steam was used as a projection surface [83]. This allowed to create illusions in the style of holograms. Additionally to the creation of such magical illusions and entertainment (compare Figure 1.4), the *laterna magica* has also been employed for educational and scientific purposes e.g. to project astronomical diagrams [83].

The animation of the content has been realized in several ways with these devices. One possibility was to have a slideable element that contains multiple images that resemble an animation (e.g. a sprinting horse). If this element is moved fast and accurate enough it can create the same illusion as todays' movies. Another way of animating the content with the *laterna magica* was to move the device itself. As already mentioned these devices were mobile and with this capable of also creating moving imagery. Furthermore it was also possible to create the illusion of animation by occluding certain parts of the projected image before it can reach the projection surface and with that dynamically change the projected content. By hiding and showing parts of the image the viewer gets the impression of animation, very similar to a shadow play (or shadow puppetery). This concept also has been taken up as a form of interaction for pico-projectors lately [53], which is also presented in more detail in Chapter 3.

Figure 1.4: Projection of a *Laterna Magica* - These devices have been used for simple entertaining plays or educational purposes. In this drawing a person showing different slides can be seen. [72]

A detailed analysis on interaction techniques for the *laterna magica* and similar historic projection devices can be found in [214]. The development of the *Laterna Magica* emphasizes the fascination that light and moving images can have on the human mind. With the capabilities of todays' projector phones this fascination can be extended to new limits in nearly all environments as the aforementioned example of the 33 Chilean miners emphasizes as well [65].

This fascination that projected light sparks, is also present in science fiction. For example the famous R2D2 from the Star Wars movie trilogy was capable of not only capturing holographic images but also projecting them. As can be seen in Figure 1.5, R2D2 was envisioned to project a hologram without any projection surface. Even though the problem of creating holographic projections is not yet solved, approaches exist to create very similar illusions [137].

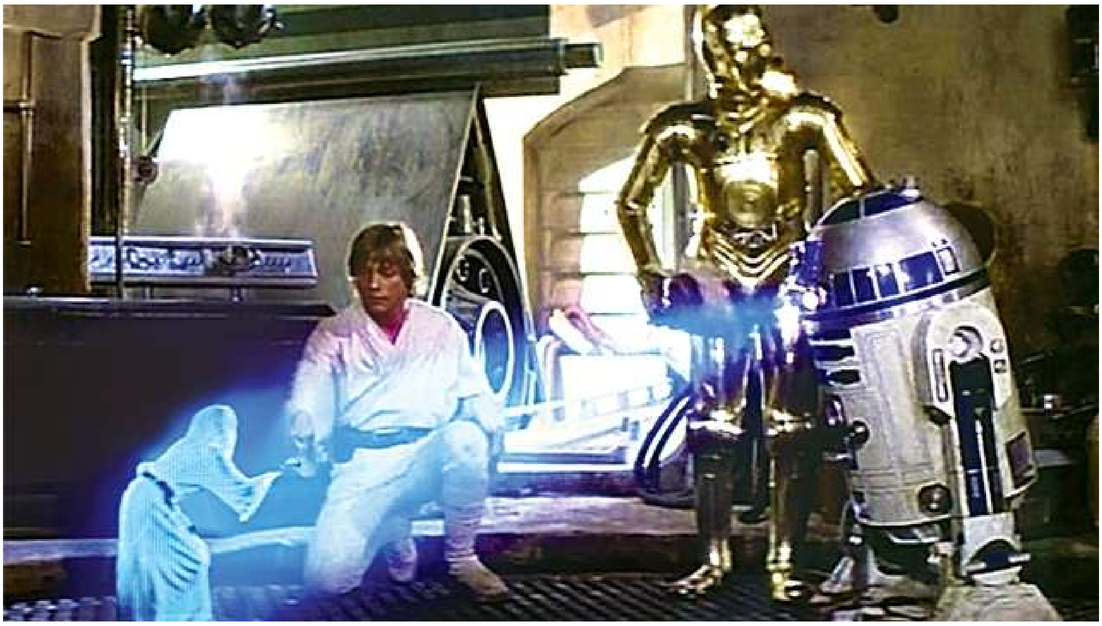


Figure 1.5: R2D2 Hologram Star Wars - The autonomous robot creating a three dimensional hologram [210].

1.3 Problem Statement

As by today, the development of mobile projectors is mainly driven from a technological side. With the advancing miniaturization of projection technology, the integration of such mobile projectors into other devices became possible. Since the availability of large personal projected displays seems a promising idea, manufacturers created products such as smartphones or video cameras that allowed to create mobile projected displays. Unfortunately they so far did not take the user's demands and needs into account. This form of technological colonialism [62] does not exploit the full capabilities that mobile projected displays may offer. One example for this is represented through the Samsung Galaxy Beam [76]. From a technical point the device featured a sufficient projection brightness, size and computational power but from an interaction and application view the device had several drawbacks. The projected display was only leveraged as a mean of output that tried to mimic static projectors. For example no API existed that would allow to create third party applications that e.g. show different content on the projected and on the physical display of the device. This limits the overall use cases drastically as third party developers only were able to mirror the image that was shown on the device's display.

In order to design enjoyable and successful interaction techniques and applications for such mobile projected displays, we need to gather an understanding of how users perceive the display in the space around them. Although one might think that such mobile projected displays can be researched as a stand-alone display, this is not the case. The environment, in which the projected display is situated in, is the main influencing factor. As the projector is mobile we have a variety of pre-conditions that decide about applicability of interaction techniques and applications compared to current desktop computer interactions. Besides privacy issues that arise in certain environments, also the physical properties of the projection surface is important. To enable mobile projected displays, some sort of suited projection surface in the environment of the user has to be available. Additionally the user is not limited in his movements. If the user is walking while holding the mobile projector and projecting, the quality and the context of the projection surface will vary. Therefore the space around the user is one of the key factors of mobile projected displays. To understand the interaction between the user and the space around him we have to investigate how he perceives his surroundings. Especially it is important to understand how the user derives his model of the current environment. This requires to understand how users create their spatial memory. To gather the needed insight we need to address the following problems:

- **How can we categorize different types of spatial memory?** First of all we need to create a basic understanding of how humans generate and organize their spatial memory. If we want to built interaction techniques and applications that successfully adapt to the surroundings of the user, this point is essential.

- **How do users perceive the mobile projected display across varying contexts?** Another crucial factor of mobile projected displays it to spawn expertise about how users perceive them. Especially in combination with a second significant smaller display coupled to the projection unit.

- **How can spatial memory aid interaction with mobile projected displays?** When moving through the environment while creating mobile projected displays it is important to incorporate the character of the environment when interacting with the displays. In which way this can be accomplished is so far not fully explored.

- **How can applications for mobile projected displays leverage the user's spatial memory?** Not only the interaction techniques need to take the environment into account. The applications should be designed in a manner that follow the characteristic of the environment. So far the design of applications for mobile projected displays where use-case driven and spatial characteristic have been neglected.

Even though current mobile projectors already reached high resolutions, for example the AAXA P300 is capable of projecting in HD 1200x800 resolution [200], they are still limited in terms of brightness. Although the power consumption has been significantly reduced in recent years, the current maximum brightness for a battery powered projector is around 160 lumens [200]. This limits possible use cases, as they can only be used in darker environments and not in bright daylight. It is expected that with further technology advancements this drawback can be reduced and the brightness can be increased to an acceptable form. In this thesis we neglect the current technological limitations of brightness and power consumption as they are expected to be solved by the manufacturers in the near future. The focus of this thesis is to create meaningful interaction techniques and applications for such future projection devices.

To lay the foundation for exploiting the full capabilities of mobile projected displays, this thesis addresses the problems mentioned above in the now following way.

1.4 Research Questions

The goal of this thesis is to enable spatial interaction with mobile projected displays to overcome the current technological-driven approaches. We will present applications and interaction techniques that exploit the possibilities of such displays. Therefore the research conducted in the course of this thesis is split into three different directions. After laying the theoretical background in terms of spatial memory development and an in-depth analysis of types of projected displays and existing interaction techniques, we will first investigate possible applications that take the characteristics of the environment into account. Next we will present an analysis of spatial alignment of projected and physical displays. Afterwards we will examine spatial interaction techniques that suit mobile projected displays. In the course of this thesis we thereby answer the following research questions:

- **How can the full capabilities of mobile projected displays be exploited through spatial features?** Besides an understanding of the users spatial memory we need to investigate suitable interaction techniques for different spatial environments. Additionally, design guidelines that allow applications to adapt to spatial features are needed.

- **What are the design requirements of future projector phones that allow for spatially aware applications?** As current projector phones do not allow seamless integration of the projected display into the environment we need to identify hardware- and system designs to make this possible.

- **How to align the projected display to support the user?** Current projector phones provide the user with two displays, on the one side the physical display of the device and on the other side the mobile projected display. How such displays need to be spatially aligned is essential to leverage the full potential of the combination of these displays.

- **What are suitable spatial interaction techniques for mobile projected displays?** Besides adaptable applications, spatial interaction techniques are needed to make mobile projected displays interactive. Therefore we need to adapt and evaluate existing techniques from mobile and desktop computing to the new setting. Furthermore novel techniques that utilize the special characteristics of mobile projected displays are required.

1.5 Methods & Approach

This thesis follows a *User-Centered Design* [148, 166, 207] approach for our theoretical, as well as for experimental contributions. Due to the technical limitations described above and the small distribution of projector phones, an *in-the-wild* [44, 176] or *Appstore-based* [88] research approach is currently not possible. Therefore this thesis draws from findings through laboratory studies. Even though this limits the generalizability of our findings, we tried to involve an as diverse as possible set of users for our studies. Since the user is, in terms of interaction and exploration of space, the main influencing factor for mobile projected displays we decided to choose this approach. Through high-fidelity prototypes [166] and expert design sessions [148] we ensure to meet the needs and demand of the users in our studies. This is further established through several feedback loops and consecutive appliance of prior findings throughout the thesis. We are convinced that this approach allows us to generate the knowledge needed to design interfaces that allow meaningful interaction with mobile projected displays which incorporate the space around the user.

For our studies we therefore rely on two kinds of factors: The first one is qualitative factor of *User Experience* (UX) in terms of acceptability and perceived effort [1, 10]. The second kind are quantitative measures such as task performance or error rate as they are by today common in the field of *Human Computer Interaction* (HCI) [166]. With this, we aim to gain a holistic insight on user interaction with mobile projected displays.

1.6 Contributions

The goal of this thesis is to enable spatial interaction with mobile projected displays to convert them into a valuable addition to our current mobile devices. The mobile projected displays enable a variety of applications to explore large scale-content. To develop such interactive mobile projected displays, this thesis contributes in the following three areas:

- **Theoretical contributions:** Up to now the usage of the term spatial memory in the field of HCI is rather diverse. We classify existing approaches and knowledge from the field of spatial cognition into micro- and macro spatial memory. This allows characterization of interaction techniques more fine-grained. Furthermore we present a classification of projected displays based on their mobility. Moreover we also classify possible interaction techniques with such as well. Additionally we developed a novel analysis technique for investigating the effects of spatial separation of mobile displays. By adopting mobile eye-tracking analysis we were capable of getting a deeper inside in the behaviour of users when interacting with mobile displays.

- **Technical contributions:** To enable interaction with mobile projected displays novel ways of sensing are needed. Since the variety of possible environments in which these displays can be created is not predictable, we can not rely on sensing techniques in the environment. Therefore we need novel sensing methods that are incorporated in the device itself. In this thesis we present two new methods for tracking direct touch interaction with a mobile projected display. Furthermore we present a sensor-fusion approach for creating a dynamic peephole. Additionally we present advanced technical solutions to create augmented reality applications for mobile projected displays. These techniques allow a seamless fusion of the projected display with the environment around the users.

- **Design contributions:** In terms of the design of interaction with projected displays we contribute on three different areas. First we investigate suited hardware designs that enable augmented reality applications. Secondly our investigations of different alignments of mobile displays contribute to a set of design guidelines for such combinations of displays. Our analysis shows that current alignments (as can be seen in Figure 1.1) do not cope with the users demands. Finally we present the design of suited interfaces for mobile projected displays that exploit the users spatial memory.

1.7 Outline of this Thesis

The remainder of this thesis is structured as follows. The second chapter presents the theoretical foundation for this thesis. Besides a definition of Augmented Reality, we present our categorization of micro- and macro spatial memory. After this theoretical background we categorize the existing interaction techniques from the related work in the field of mobile projection. Additionally we also outline a typification of projected displays based on their mobility. Afterwards we will present the design of spatial aware applications for mobile projected displays from the field of augmented reality. This section is followed by an analysis of spatial alignment of mobile projected displays and physical mobile displays. In addition to this, we also present with the SurfacePhone a specific system applying these findings. Next, we present and evaluate two different techniques for application switching and pointing interaction. As a concluding system, in Chapter 7 we then present the implementation and evaluation of a system that incorporates the prior findings. The thesis concludes by summarizing the main contributions of this thesis and identifying future work.

The majority of the work that is presented within this thesis was carried out in collaboration with researchers and students from different institutions. For this reason, the scientific plural we is used throughout the thesis. All URLs are treated as references and have been last visited on the 15th of April 2014.

2

Theoretical Background

In this chapter, the theoretical background for this thesis is presented, focusing on augmented reality (AR) as well as memory and communication techniques that are based on spatial relationships. After a definition of the term Augmented Reality (AR), we will outline the relevance of AR for Ubiquitous Computing and especially for mobile projection. Even though this thesis is rooted in the field of HCI, AR is a very promising type of application for mobile projection devices as it represents the perfect blend of a projected display and the reality. This requires applications that are highly spatially aware, which is one of the main contributions of this thesis. Other areas that will be touched can not only be found in the next chapter but they will also be presented and discussed throughout the thesis where necessary.

Besides AR, in this chapter we will discuss memory and communication techniques that have a spatial component and will be applied in this thesis. As already discussed in the introduction, the space around the user and its surfaces are of high importance for mobile projection. Therefore it is necessary to get a basic understanding of how the human cognition builds up its spatial memory and how we can appropriately exploit this to create novel mobile projection based interfaces. By combining theories from spatial cognition, cognitive psychology and physiology, we derive the concept of micro- and macro spatial memory. Additionally, we will also present the concept of Proxemics and how it can be employed for mobile projections scenarios. Proxemics is a concept of spatial relationships, which is rooted in social sciences.

Finally this chapter will present the concept of dynamic and static peephole interaction. Dynamic peephole interaction in particular provides a very natural and effective way to explore large scale content on mobile devices.

2.1 Augmented Reality

At this point we would like to define Augmented Reality and subsequently give an overview of how AR displays can be created. Furthermore, we will give a short overview of possible application scenarios. An in-depth discussion of AR applications can additionally be found in the related work chapter as well as in Chapter 4.

2.1.1 Definition

Azuma defines the term Augmented Reality as the extension of user's the field of view of a user onto the reality through digital information [5]. Through partial augmentation of the user's viewport, a fusion of the digital and real world is possible. This allows to create systems that allow the user to always have digital information connected to corresponding real world objects. If supported by the system, it would also enable the user to interact and explore these information in a more natural way [24]. According to Azuma the following requirements have to be fulfilled in order to generate an AR system [5]:

- Combines real and virtual
- Is interactive in real time
- Is registered in three dimensions

In contrast to virtual reality, which tries to let the user emerge in a newly established virtual world, AR does not create a fully new reality for the user. Milgram does describe AR as the transition phase between the users real world and a virtual reality [145].

2.1.2 Augmented Reality Displays and Application Examples

To establish the overlay described above, over the user's reality, Bimber and Raskar distinguish between five different spatial layers where a display can be created that would allow for AR [24]. The first two layers are established by attaching a device to

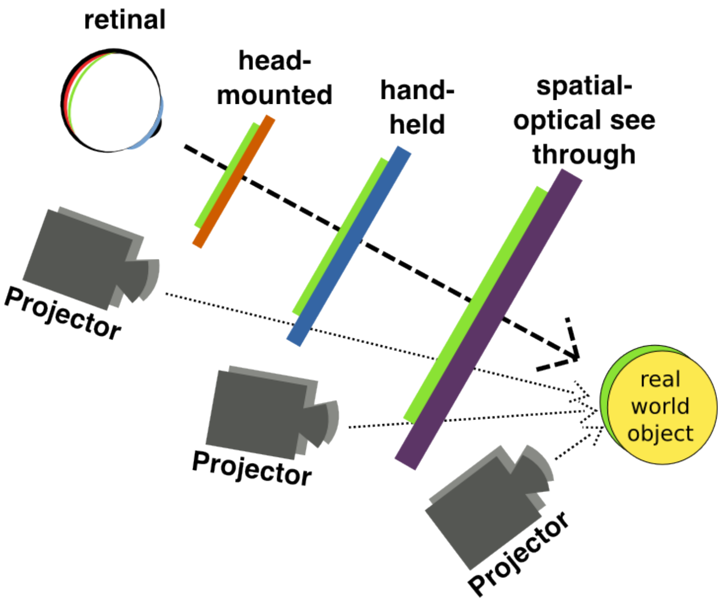


Figure 2.1: Layers for AR Display - The five different layers where an AR display can be created, starting from the retina of the user and ending on the object itself. [24]

the users head. These create a display either directly on the retina of the user or very close to the eye, e.g. to the display of Google Glass [96]. The third layer would be the display of a hand-held mobile device, such as today's smartphones. Overlays that are created on a display, which is moveable and closely mounted to the object would be the fourth layer. The last layer would be to display the information directly on the object. Bimbar and Raskar envision here to use a projector [24] that can either be attached to the user, hand-held, or placed in the environment. All these display types and their alignments can be seen in Figure 2.1.

We will now give a short overview of applications for head-attached, hand-held and spatial-optical see-through AR displays. Further examples of AR applications for mobile projection devices can be found in the next chapter as well as in Chapter 4.

Head-attached Displays For the head-attached AR displays one has to distinguish between a Retinal- and Head-Mounted Display (HMD). Retinal displays project the image using low power lasers directly onto the retina of the user [24]. This allows for high-

Figure 2.2: Human Pacman - On the left side the players equipped with an HMD can be seen, on the right side the view of players is shown. [49]

contrast images but these devices are still in a very early stage of development [213]. Until now only one commercial product exists, the Brother Air Scouter [32]. HMDs on the other side are currently gaining a momentum towards market introduction. Even though Google Glass is not yet commercially available, it is set to be released as a mass market product in 2014 [96]. Additionally, a variety of different HMDs have been announced to become available in 2014 as well. For HMDs, we have to distinguish between a video see-through and optical see-through. For the video see-through HMD, the video signal of a view-aligned camera of the real world including the overlay is shown on a non transparent display. The optical see-through HMD on the other side shows the overlay on a semi-transparent layer (e.g. a transparent LCD) so that the user can see the real world through the display as well [24]. An optical see-through HMD comes at the cost of lower brightness and a slightly blurred image of the real world.

For single user scenarios such HMDs are perfectly suited for AR. They present one of the best, easiest and most natural ways to experience AR. For multiple users, this would require an increased amount of hardware, as each user would need an HMD. One classic example for this would be the Human Pacman application by Cheok et al. [49]. Human Pacman is an AR adaption of the classic PAC-MAN game using HMDs. The players have to "eat" virtual points that are shown as an overlay on the HMD, as can be seen in Figure 2.2. To track the players, Cheok et al. used GPS [49].

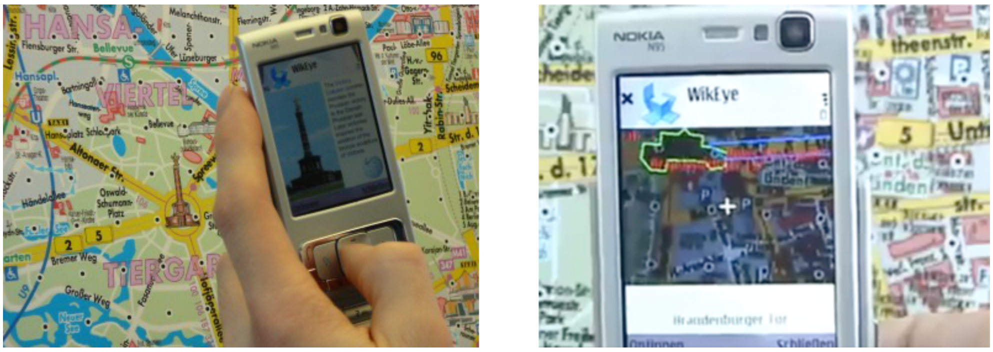


Figure 2.3: WikEye - An AR map interaction application, showing additional information to Points of Interest on a map. [86]

Hand-held Displays As today's smartphones are often equipped with a camera, they allow to track physical marker-less objects and with display sizes reaching 4.7", they are well suited to present the AR overlay. To achieve this, the physical object is tracked using the video signal and afterwards augmented with additional information before it is presented on the screen. This allows to create AR at arm-reach [24]. Even though these displays have the advantage that they can easily be observed by multiple users synchronously, they are still limited in size and come at the expense of one hand being occupied all the time. One compelling scenario is for example to augment static paper maps with digital information [86]. Even though at the time of publication of their work, completely marker-less tracking was technically not possible on a mobile device, the markers used are almost non perceivable. The static paper map in this scenario is automatically annotated with Wikipedia content, which shows how new natural user interfaces can help to explore digital content easily as it can be seen in Figure 2.3.

Spatial Displays Bimber and Raskar define spatial AR displays as a technology that is not bound to the user and rather bound to the objects itself. These are meant to be augmented and that are integrated into the environment [24]. Also, they distinguish between two different ways; in the first variant, the alignment of the camera that captures the real world and the display is irrelevant. For the second variant, the display works as an overlay over the real world showing the digital information. An example for the first variant would be the ARToolKit where the desktop computer

Figure 2.4: Fraunhofer IGD Spatial Displays - A spatial display deployed in the Allard Pierson Museum in Amsterdam that showed additional information for several exhibits [93].

display acts as viewport to the real world while the camera is free to decide on the real world aperture [105]. For the second variant the display can be moved and it shows the excerpt of the reality that is behind the display. One example for this would be the application developed by the Fraunhofer IGD [93]. They deployed several of those displays in the Allard Pierson Museum in Amsterdam that could be turned around and showed additional information for exhibits behind the display (as can be seen in Figure 2.4).

2.2 Spatial Memory

As discussed in the introduction, the space around the user of a mobile projection device is of high importance. Without a suited projection surface there is no way of creating such a projected display. Additionally, the advantage of a mobile projected display is the fact that it is not bound to one surface. Moreover, it is from high importance to understand how the users perceive their environment and derive a model from it. This knowledge can be exploited to allow for advanced interaction with the display as we will also see in the Chapters 5, 6 and 7.

Montello stated that the size of space is a perceptual problem [150]. To ease communication about space he defined four different sizes of space that vary in the amount of locomotion required to explore them.

- The first space is called the *Vista space*. It is characterized by the possibility of being apprehended completely from a single view point. In our characterization of micro and macro spatial memory, it would be reflected in the micro spatial memory. But in Montello's concept of space, it also includes complete rooms that could be apprehended without locomotion or even whole areas such as a valley that could be explored from e.g. a plateau. The *Vista space* is the most relevant space for HCI and the Spatial User Interface community as it reflect the typical interaction space that is investigated.

- The next bigger space size is called the *Environmental space*, and it characterizes a space that can not be apprehended without a significant amount of locomotion. Furthermore such spaces require time and repetitive explorations to be completely apprehended. An example for the *Environmental space* would be a city where only through consecutive exploration a complete spatial model can be store in the human mind.

- The third space is the largest and is called the *Geographical space*. It is characterized by the fact that it can not be apprehended by locomotion and must be learned from symbolic representations. Examples for such spaces would be countries or continents.

- The last type of space is the largest and smallest space at the same time. Montello defines the *Figural space* as a space that represents a larger space such as the *Geographical space* as either a 2D (Pictural space) or 3D (Object space) representations that are significantly smaller than the *Vista space*. The Pictural space consists, as the name suggests, of 2D representations such as maps while the Object spaces would cover small scale models.

One might suspect that current desktop interfaces would be covered by the Pictural space but Montello explicitly states that they are not covered in this space [150]. In later publications Montello acknowledges that acquisition of spatial memory is also possible through virtual channels and to some extent follows the *Figural space* memory learning [151]

This space classification is very crucial to understand how humans create their spatial memory. But one has to keep in mind that, when it comes to computing interfaces, often either the *Vista space* or the *Figural space* are applied to create spatial memory. The *Vista space* is of special importance as we move forwards to the age of Ubiquitous Computing [212] and the users are more frequently confronted with an interface that is distributed in their environment. The exploration of such interfaces and recalling where certain user interface elements are hidden will therefore become a task that requires spatial memory of one's surroundings, as classical interface guidelines are not applicable in such environments. Furthermore, larger areas are nowadays often explored using digital maps and other applications that are presented to the user on mobile devices. This means that also the *Figural space* is presented to the user frequently on our computing devices. These examples outline the importance of space size and classification for current computing devices.

To cope with these different space sizes' we introduce at this point the novel concept of micro- and macro spatial memory, bringing together the results of the field of spatial cognition and other working memory techniques. In current HCI literature the term spatial memory is very broadly used for nearly every memory technique that somehow is connected to space or maps. Not only does the HCI literature often neglect spatial cognition findings, the term spatial memory is arguably also misused at some points. The interdisciplinary field of spatial cognition is concerned about how the human mind acquires, manages and retrieve information about the space around the human and its environment. As these investigations normally focus on room- to city-sized environments we classify the findings as macro spatial memory. Other memory techniques that are e.g. bound to the human body are categorized as micro spatial memory.

The chosen terminology is based on the micro- and macro-navigation distinction introduced by Stahl et al. [192]. They defined macro-navigation as follows: *"Macro-navigation is concerned with tasks, in which the navigational goal is beyond the user's perception of the current environment, in the sense that the user has to move through the environment to reach their destination, which requires route and survey knowledge about the environment."* [192]. Similar to that we define macro spatial memory as spatial memory that requires locomotion to acquire this memory, while micro-spatial memory can be acquired without a change of the viewport. Here the current desktop computing metaphor presents a special case. Even though the screen can be perceived completely without a change of the users viewport, exploring all information available requires to navigate through different windows or multiple sceens and with that a special type of digital locomotion. This problem has been investigated from spatial cognition as well as from an HCI standpoint [150, 202]. Furthermore there is a whole research corpus on navigational behaviour in virtual reality. As this thesis is neither focussing on virtual reality nor on interfaces for desktop computing, we will neglect these areas for now.

At this point we now want to present several concepts of how humans develop their macro - and micro spatial memory. We by no means make any claims that this overview is complete. The selected techniques presented here are either applied in this thesis or can be applied in interface design.

2.2.1 Macro Spatial Memory

We classify macro spatial memory as the memory that the user captures of his environment through locomotion. Our classification of macro spatial memory would start with the further away *Vista space* such as bigger rooms, and everything larger in Montello's definition of space [150]. The concepts of how we develop this spatial memory are based on the findings in the field of spatial cognition. At this point we would like to refer to the overview of McNamara [143], which presents one of the most comprehensive overviews of current findings on how humans perceive and organize their spatial memory. Based on McNamara's findings we outline how humans organize and acquire their macro spatial memory and present one of the most often utilized techniques for short term spatial memory acquisition called snapshot matching.

2.2.1.1 Macro Spatial Memory Acquisition

According to McNamara humans organize their macro spatial memory in four subcategories:

1. Object-Place Knowledge - the memory about landmarks, which represent one of the most important features for human navigation [167]. Landmarks can for

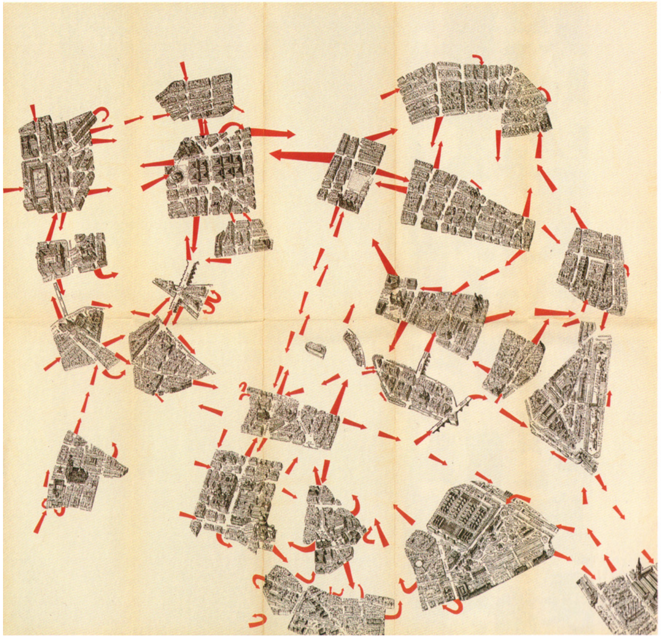


Figure 2.5: Psychogeographic Guide of Paris - A map created by the Situationists that consist of different areas that are perceived as neighbourhoods and the connections between them [58].

example be specific buildings such as the Eifel Tower or other unique objects or natural formations. They can help locate one self roughly but do not allow you to generate specific navigational aids. For example on vacation in Paris, one might be able to see the Eifel Tower from one's hotel room but this will not help in finding one's hotel when being at most other locations in the city, even though one might be able to see the Eifel Tower.

2. Route Knowledge - is the memory that allows humans to navigate a certain route along known landmarks. They, by default, are categorized as a sequence of landmarks with no metric relationships between them. Route knowledge is often used for example to give navigational instructions to strangers such as "drive down this street until you see a big blue building, then you have to turn left". In this example the big blue building would be the landmark as it separates itself from the environment enough to be helpful when navigating.

3. Environmental Shape Knowledge - not only do shapes represent important features to recognize landmarks [167], the Environmental Shape Knowledge repre-

sent the part of spatial memory that over time builds up connections and relations between known landmarks.

4. Survey Knowledge - over time humans add to the Environmental Shape Knowledge further euclidean features that allow the human to create a cognitive map of the environment. The Survey Knowledge allows to create routes that one has not travelled before and are not build on a sequence of landmarks.

As it becomes clear from these memory types, humans build up their knowledge over time, mostly through locomotion or maps. Starting from fragmented knowledge (Object-Place Knowledge) such as landmarks, humans tend to build a slightly distorted memory of their environment, in thich they try to align the known fragments in a more ideal form such as $90\,^\circ$ angles [143]. With further exploration this distortion transforms into a more accurate cognitive map. Furthermore, McNamara especially stresses that humans generate a hierarchical order of this memory where more important parts for navigation are higher arranged than less important parts. Special for this order is, that it is neutral to the personal appreciation for the specific part. One might even consider ugly buildings or the stadium of a rival football club from high importance in this order even though one has negative feelings connected to it [143]. A depiction of this evolution can be seen in Figure 2.5. Even though it was created by the Situationists [58], a group that rather focussed on psychological aspects of cities, it still is a very accurate representation of how humans first develop a memory of fragmented places and over time connect those places to a larger cognitive map.

The way humans organize and generate their spatial memory can be applied to interface design as well. One example would be the design of menus, where the user first generates an understanding which control element is associated with which menu item, before he starts memorizing the actual spatial position of the control element [8]. When designing spatial user interfaces, the construction of spatial memory is an important point and should follow the idea of landmarks as described above. This has already been successfully applied in HCI for example in [174]. Robertson et al. presented a document management sytem, which allowed users to place documents at arbitrary positions on an inclined plane in a 3D desktop environment. This way, users could align applications and locate them again by using 3D spatial memory.

2.2.1.2 Snapshot Matching

One of the oldest techniques of human navigation that has been investigated is the so called Snapshot Matching or View-based Navigation [48]. For this technique humans memorize landmarks along a route based on specific "Snapshots". If one tries to recall the way back, one simply minimizes the discrepancy between saved images and the current view. By repeating this minimization process while walking along the route, one can safely navigate back. Well suited elements in the environment for such snapshots can either be the skylines or close by static objects. The technique is also known to be applied by insects such as ants to navigate between the anthill and a point where e.g. the food is located [61].

Naturally, this behaviour can be observed in interaction with computing interfaces as well. If a user tries to recap the navigational path through for example a complex website, one tries to re-establish the positions one observed during the first exploration of the website. Of course, more distinct features that act as landmarks can help when navigating through multi media content [202]. For the future design of spatial user interfaces this particular effect should be exploited as well by providing elements that allow for snapshot matching.

2.2.2 Micro Spatial Memory

We classify micro spatial memory as the memory that the user generates about his body and the space around his body that he can reach with his extremities. In Montello's definition of space this would be covered by the *Vista space* but does not cover the complete *Vista space* [150]. This does not only include techniques where the body functions as a memory - and gesture aid but also techniques where physiological properties of the body are exploited. Compared to the macro spatial memory these techniques are independent from the users environment as they are only based on the user's body and its posture. While for the macro spatial memory a learning phase is always required this is not necessarily the case for the here defined micro spatial memory. Even though some techniques only exploit the human body for an easier learning phase, micro spatial memory also includes physiological attributes of the human body that are often underutilized in our every-day life [168]. Even though a variety of memory techniques exist [7] that could be classified as micro spatial memory, we only want to focus at this point on those that have been employed in this thesis to the field of mobile projection. Those are namely the method of loci and kinaesthetic cues.

2.2.2.1 Method of Loci

The method of loci, which is also called Memory Palace, Journey Method or Roman Room, is an ancient method that helps memorizing information using visualization. Those kinds of memory techniques that allow to retain information are in general called mnemonic devices [17]. The goal of mnemonics is to transform information into a shape that the human brain can more easily store and retrieve them as in their orignial form. According to Levin et al. the transition into a more suited form can already aid the transformation into the persons long-term memory [116]. Examples for such transformations would be the conversion from lists into an auditory form such as poems or memorable phrases. Additionally mnemonics can also be used in visual or kinaesthetic forms.

The method of loci is such a visual form and additionally one of the oldest memory techniques known to mankind. It has been proven to be used already by the ancient Roman and Greek [228] successfully. The method of loci is applied in the following way:

1. The user memorizes a layout of a room, building, street, or something comparable

2. The user associates discrete locations in these spaces with the information he wants to retain by thinking of distinct features

3. The user "walks" through the space and by thinking of the features the memory gets activated

By combining spatial locations and its relationship with other type of information, the relationships and order help structuring the novel information. It is particularly useful for unrelated information [116, 160].

Instead of memorizing a specific building or other locations, the user can also apply this technique to his own body. By touching specific locations the memory effect can be additionally extended through the haptic cues [160]. A classic example for this would be memorizing the number of days in each month of the year using one's knuckles. Each knuckle represents a month with 31 days while the gaps in between represents the months with less days.

The method of loci using the user's body as interface has been successfully applied for mobile interaction. Ängeslevä et al. presented an evaluation of possible areas of the body that are suited to act as memory aids [2]. Strachan et al. then built on top of their findings a gesture recognition system that would detect if the user is touching the corresponding body parts [194]. They created an example music player application that realized these features. Closely related to this is also the work of Chen et al. which explored on- and around body interactions with mobile devices [47]. They identified three different patterns from the related work, *Body-Centric Digital Content Storage/Retrieval*, *Body-Centric Digital Shortcuts* and *Body-Centric Controls*. All of these patterns are either based on method of loci or on kinaesthetic cues (which are discussed below). Chen et al. exploited the method of loci very similar to Strachan et al. [194] for connecting either control elements or information with body locations.

In this thesis we employ the Roman Room variant of the method of loci. As it is specific to objects in one room we employ the semantic features of real world objects to mobile applications. For example connecting a paper based calendar to a calendar application on one's phone. The approach is described in more detail in Chapter 7.

2.2.2.2 Kinaesthetic Cues

In 1887, Bastian coined the term *kinaesthesia*, which refers to the human ability to sense orientation and current movement of one's limbs around the body [12]. The kinaesthetic sense of the human body is capable of sensing these orientation without

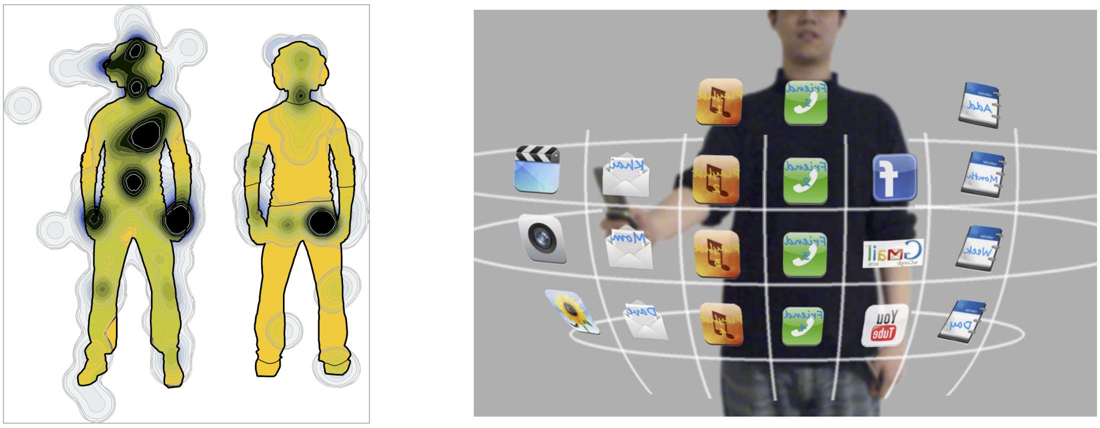


Figure 2.6: Micro Spatial Memory HCI Applications - Two HCI applications based on micro spatial memory. On the left the analysis of suited body parts to utilize the method of loci by Ängeslevä et al. [2] and on the right the Virtual Shelves application that uses the kinaesthetic senses to easily trigger controls around the user. [119]

visual feedback. Mainly responsible for this remarkable capacity are muscle spindles and skin stretch receptors in our limbs. Current research suggests that the human body has two separate senses; one to sense the limb position and another sense for current movement. A more holistic overview from a physiological and neuropsychological point has been developed by Proske and Gandevia [168].

Compared to our other senses such as vision or hearing, we as humans are not really aware of the capabilities of our kinaesthetic senses. Even though this senses exist genetically, we rarely make use of them. Therefore it cannot be anticipated that we can make full use of its possible capabilities. Compared to vision, current research suggest that we need to train this sense as well [168]. Even though the kinaesthetic sense might be degenerated over time and we would have to relearn to make use of full its capabilities, to some extend, we can still make use of it. Li et al. showed that even untrained users can easily differentiate between 7 regions on the theta and 4 regions on the phi plane around the user [119]. They further developed an application that would allow to automatically start applications or trigger certain music tracks by pointing in a distinctive direction (as can be seen in Figure 2.6). Similar results were also found by Tan et al. [198], as they showed that using kinaesthetic cues increases spatial recall. We also successfully employed this technique in a standard mobile setting as well as in a mobile projection setting. More details about this can be found in Chapter 6.

2.2.3 Discussion

This section is meant to give a rough introduction into how users acquire and organize their spatial memory. Even though by no means we make any claims that this overview is complete, it presents several techniques that we will apply in the course of this thesis. Through the distinction between macro- and micro spatial memory, it enables to combine the findings from spatial cognition as well as other working memory techniques and physiological features of the human body. We hope that this distinction can help other researchers to be more precise when applying the concept of spatial memory to the field of HCI.

2.3 Proxemics

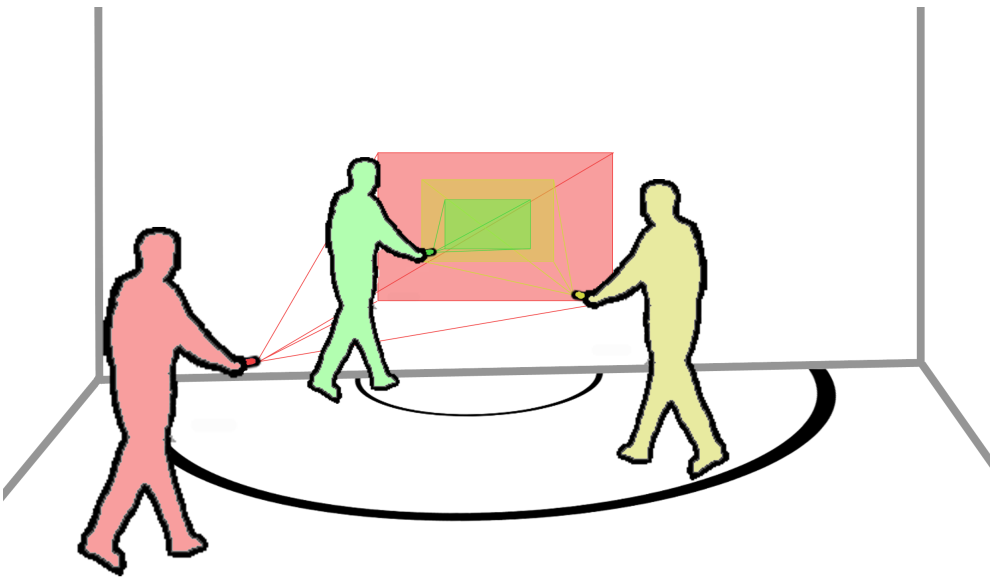


Figure 2.7: Mobile projection in the different proxemic regions. - Initmate (green), social and consultative (yellow) and public (red).

One of the main influencing factors when interacting with hand-held mobile projection devices is that size and position of the projected display is always coupled to the device and its motions. Even though there are approaches for decoupling the projection orientation from the device orientation [42], the distance from the device to the projection surface would still be a crucial aspect even in these cases. Depending on the distance, it is expected that people would be willing to show different content - e.g. if they are able to shield the projection with their body from passersby's they could project personal information [133]. Furthermore, different distances would allow and even require the usage of different interaction techniques.

Especially for AR applications the orientation and distance between the user and the object that should be augmented are always important factors. In most cases these information are only used for the tracking of the object. With mobile projection, the created AR overlay is significantly different from the magic lens approach, since it is visible for everybody. Therefore, if personal information is taken into account to create the overla,y it is possible that this information is revealed easily. Since this is normally not favoured by users, applications have to adapt to these situations. We propose to use the knowledge about the spatial relationship between the projection surface and the device as a key indicator for spatially aware applications.

The concept on how to exploit social spatial relationships - so called proxemics - as interaction techniques into ubiquitous computing scenarios was introduced by Ballendat et al. [9]. They showed how device movement between discrete proxemic regions can control and influence applications using a prototypical media player. Their concepts originated from Edward Hall‘s theory of interpersonal spatial relationships that he called proxemics [81] (compare Figure 2.7). Hall distinguished between four different regions that connect physical distance to social distance, the intimate-, the personal-, the social and consultative space and the public space. These regions can be seen in Figure 2.7 adapted to a mobile projection scenario. The intimate space is the closest "circle" around a person and the entry into this space is only acceptable for close friends and intimates. Their personal space is suited for interactions among close friends and family members. The social space is defined as the space in which people feel comfortable conducting routine and social interactions with acquaintances as well as strangers.

The public space is the area in which interactions are perceived as impersonal or anonymous. The proxemic zones would all be incorporated in Montello's Vista space [150].

It is difficult to give exact distances for theses regions, since they are perceived differently depending on the culture. Since the distance between the device and the projection surface can be sensed, it would be easy to include these proxemic regions into future spatially aware applications for mobile projected displays. In a way the projection unit itself already acts according to proxemics, since constrained by the distance the possible projection size changes. It should be kept in mind when designing applications that require a high resolution that this is only possible when the mobile projector is relatively close to the projection surface. With increasing distance to the projection surface the projection size becomes bigger and easier readable for uninvolved passersbys.

While Ballendat et al. focussed on the implications of the user being in different zones [9], we focus on the projection being moved to different zones. In projection based applications this is where the information is revealed. This idea is closely related to the approach of Vogel and Balakrishnan [206]. When moving the projected display closer to the user or further away from him, the interaction technique will be changed automatically to the most suited one. Additionally, the presented information will change. General information that are of interest for everybody or not meant to kept private can be shown in every proxemic region. Personal information on the other hand would only be shown in the intimate zone. Here the user is close enough to the projection to shield it. The following imaginable application exemplifies it more deeply: MuseoLight is an AR application for a projector phone that allows projecting information about a specific exhibit in a museum besides and on the exhibit. When the user is a couple of meters away, the projection would show information like date of creation or highlight specific parts of the exhibit that are commonly accepted to be of interest. But when the user is only a couple of centimetres away from the exhibit the projection shows a note left by the user's partner or highlights a region for a more personal reason. Since the user is closer to the projection surface, the projected information is smaller and he is able to shield the projection from uninvolved visitors and with that, he can create his small personal sphere.

2.4 Dynamic - and Static Peephole Interaction

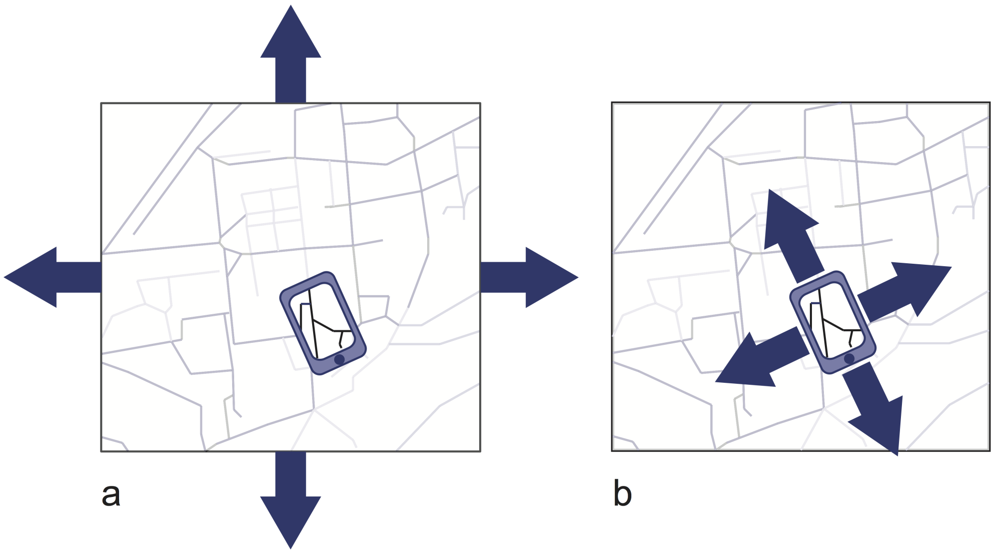


Figure 2.8: Dynamic - and Static Peephole Interaction - a) Static peephole: The viewport position is fixed while the information space is moved behind the viewport. b) The information space is fixed while the viewport is moved in front of it [34].

The exploration of large-scale content such as websites or maps is one of the most common tasks on today's smartphones. On current smartphones, this task is typically handled using a static peephole [70]. This means that the content that is presented on the display is only a zoomed clipping of a complete workspace. Thus, the display provides access to a larger virtual workspace behind a limited viewport. To be able to explore the complete workspace, it can be moved, e.g. by scrolling or panning whereas the viewport stays fixed. As can be seen in Figure 2.8 a. An analysis of this interaction method by Guiard et al. [79] shows that the navigation time increases when the view size is reduced. While modern mobile phones such as Apple's iPhone rely on direct manipulation user interfaces [191] - possible through the touch screen - these techniques are not feasible in all cases for interaction with mobile projected displays. Due to the distance that the user has to bring between him and the projection surface to create a projection from suitable size, he might not able to reach projection with one of his extremities.

To overcome this problem, Fitzmaurice [70] introduced the concept of the so-called dynamic peephole. In this technique, the viewport is physically moved above the virtual workspace that is static with respect to an external frame of reference (see 2.8 b). By the help of a spatially-aware display, it is detected which part of the virtual workspace has to be shown. The sensors embedded in current devices such as gyroscopes and accelerometers would allow such devices to track their relative position and with that let them become more spatially-aware.

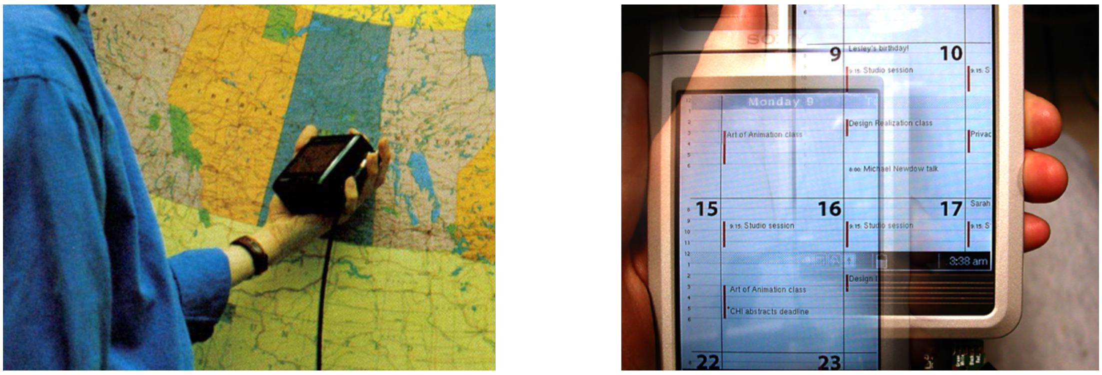


Figure 2.9: Early Peephole Prototypes - Two early prototypes of spatially aware peephole interactions. Fitzmaurice's prototype (left) was capable combining real world objects with digital information [70], while Yee investigated 2D interaction technique [230].

The system developed by Fitzmaurice, called Chameleon, allowed the user to explore different manipulation techniques for situated 3D information spaces [70]. His spatially aware prototype allowed the user to explore varying levels of information based on gestures and movements of the device. Furthermore, it allowed to connect real world objects and digital information as well, e.g. a paper map as can be seen in Figure 2.9.

While Fitzmaurice's prototype focussed on 3D interaction techniques, Yee presented a 2D spatially aware prototype that allowed for pen input (compare Figure 2.9 right). Additionally their in-depth analysis showed that a dynamic peephole interaction technique can provide significant advantage compared to current manipulation techniques. Lately Rädle et al. conducted a comparative study, investigating the effect of 2D peephole size on task performance [169]. They found that increasing the size above the approximate size a projector phone would create, would have a significant effect in terms of task completion time. Furthermore, to create a mobile projected display a sufficiently large surface is required. This forbids the display itself to be moved in three dimensions and limits it to two dimensions as well.

2.5 Summary

In this chapter we presented the theoretical background for the thesis. As one of the main goals of this thesis is to create applications that allow for seamless integration of the projected displays with the environment, we discussed the foundations of AR. As AR presents a merge of the digital - and the real world, it presents the best possible solution to merge a projected display with the real world. Furthermore mobile projection can provide the most natural and efficient way to create an AR display as the overlay is directly created on the object. Concepts and system designs for mobile projection based AR will be investigated in-depth in Chapter 4.

Another main goal of this thesis is to exploit spatial features of the environment for applications and interactions with mobile projected displays. Therefore we developed the concept of micro- and macro spatial memory that combines findings from the field of spatial cognition with physiological and working memory theories. This enables us to judge spatial interaction techniques in a better way. This knowledge is applied to interaction techniques in Chapter 6 and 7. Additionally, we present the concept of proxemics from the field of social science and provide evidence and use cases where they could be taken into account. In Chapter 7 we present a system that is relying on these social zones to change interaction techniques as well as projected content.

Moreover, we introduced the concept of dynamic - and static peephole interaction. Dynamic peephole interaction provides a promising technique for mobile projected displays and will be applied in Chapter 6 and 7. The recent study of Rädle et al. [169] provides evidence that the size of a projector phone is well suited for dynamic peephole interaction.

3

Related Work

This chapter provides an understanding, of the research and state-of-the-art on which this thesis is based. As discussed in the introduction, the work of this thesis is focussing on interfaces that support spatial interaction with mobile projected displays. We therefore will firstly present different types of mobile projected displays and categorize them based on their mobility. Afterwards we will classify and describe interaction techniques in detail, that have been applied and partially evaluated for such displays.

3.1 Types of Projected Displays

When it comes to different projected displays we can distinguish between four types that differ in the amount of mobility they provide: *Environmental Projectors*, which as the name suggests are located in the environment, e.g. projectors installed in meeting rooms; *Moveable Projectors* which are not fixed in the environment but need to be set up before operation; *Hand-Held and Body-Worn Projectors* are fully mobile projectors that are either held in the hand, e.g. a projector phone, or worn on the body; *Self-actuated Projectors* presents the most mobile setting, in which the projector can move through the environment on its own. This categorization is visualised in Figure 3.1.

3.1.1 Environmental Projectors

Environmental Projectors are characterized by the projection unit being immobile and therefore fixed in the environment. The most common example for this class would be projectors that are installed in e.g. meeting rooms. In this case the output – the

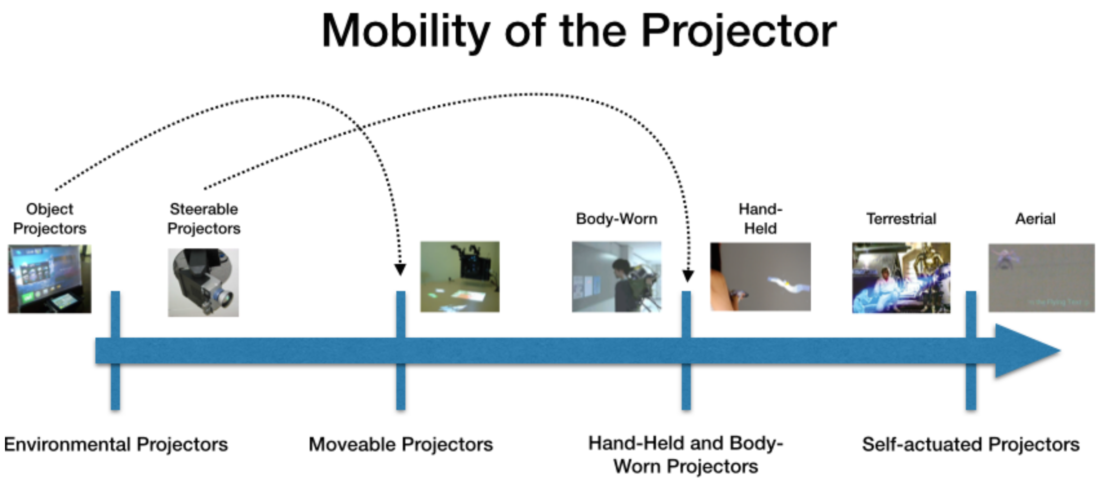


Figure 3.1: Mobility Categories of Projectors - The four different categories aligned, where the x-axis represents the increasing mobility (from left to right). The subclasses of Object- and Steerable-Projectors are hybrid classes indicated by the arrows.

projected display – is usually fixed as well. Thereby *Environmental Projectors* provide the advantage of a large undistorted projected display.

Still two subcategories exist that are a hybrid between *Environmental Projectors* and more mobile classes. The first class is represented by Steerable Projectors where even though the projector is fixed in the environment, the projected display can be created at different positions in the environment. This subclass represents a hybrid between the *Environmental Projectors* and the *Hand-Held and Body-Worn Projectors*. The second subclass are Object Projectors, where the projector is mounted on a moveable object but the created projected display is limited to this specific object. This subclass is a hybrid between *Environmental Projectors* and the *Moveable Projectors* class.

3.1.1.1 Steerable Projectors

Steerable Projectors allow, even though the projection unit is mounted in a fixed position in the environment, to create projected displays at nearly arbitrary positions in the environment. This capability of Steerable Projectors makes them nearly as powerful as *Hand-Held and Body-Worn Projectors* in their dimensions of freedom to create displays. Therefore the sub-class of Steerable Projectors represents a hybrid between

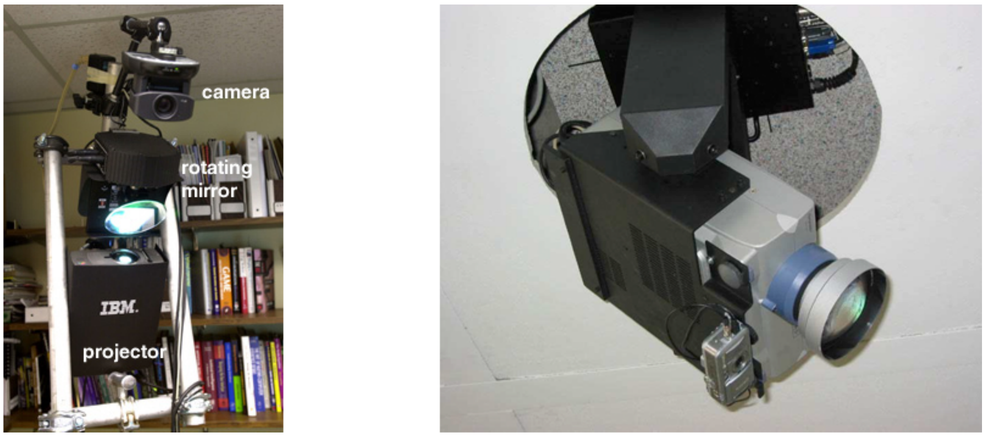


Figure 3.2: Two solutions to create Steerable Projectors - On the left side the Everywhere Displays approach by Pinhanez where a steerable mirror is positioned on top of the projector to deflect the projection ray [164]. On the right side the beaMover approach, where a projector is rested on a moving head to redirection it's output [35].

the *Environmental Projectors* and the *Hand-Held and Body-Worn Projectors*.

Two general solutions exist to create steerable projections, the first one would be to attach a steerable mirror on top of the projection lens. This has been done by Pinhanez with the Everywhere Displays [164]. The steerable mirror allowed him to deflect the projection into the direction of the demanded display. This can be seen in Figure 3.2. The second possibility is to make the whole projection unit moveable as it is done e.g. for the beaMover projectors [15]. Similar to a Pan-Tilt-Zoom camera the projector is mounted on a controllable moving head that can rotate and tilt the projector to a certain direction controlled through a standard DMX interface (compare Figure 3.2 right).

But both approaches suffer from the same drawback. Since in most cases the projection is not orthogonally aligned with the projection surface, keystone effects occur. To circumvent these drawbacks Pinhanez created a method to remove perspective distortion based on a three dimensional model of the environment. In this three dimensional model the image that should be projected is placed as a texture on the surface where the projected display should be presented in the real world. Afterwards a virtual camera that shares the same optical parameters as the projector is placed in the model at the point where the projector is mounted in reality. This camera then takes an image of the placed textures. Finally this image is projected. Due to the shared optical parameters, the image that gets projected appears to be undistorted in the real world. This approach has two major drawbacks, the first one would be the loss of resolution as the projection of the undistorted image does not take full advantage of the whole possible projection size. And the second drawback is the need for a three dimensional model of the environment. As this is often not available for all environments it makes this approach less suited for a mobile projector in an unknown environment. Still, Molyneaux et al. applied a similar approach for *Hand-Held and Body-Worn Projectors* based on an ad-hoc created three dimensional model constructed from Kinect depth data [149].

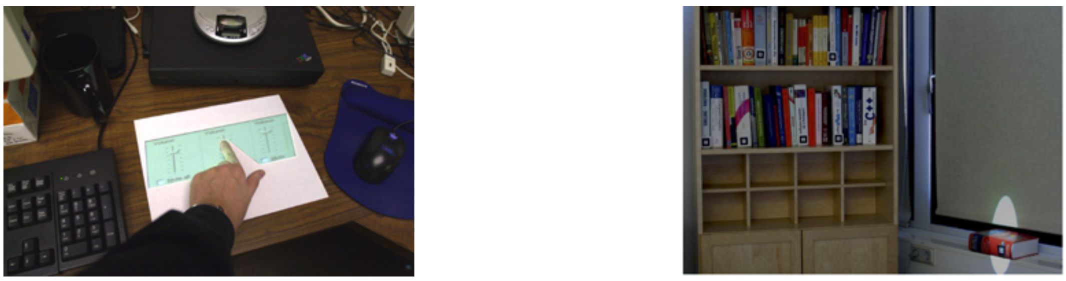


Figure 3.3: Two applications for Steerable Projectors - On the left side the Everywhere Displays projects standard desktop interfaces [164]. On the right side the SearchLight highlighting objects [35].

A variety of applications can be enabled through this techniques, apart form the projection of standard desktop user interfaces [164] (compare Figure 3.3 left), such projectors can also aid as help for finding objects in the environment [35] (compare Figure 3.3 right). In the SearchLight example of Butz et al. the three dimensional model is additionally used to create a map of all objects that can be searched for. A good overview of possible applications for such steerable projections can be found in [163]. As for input on the projection these approaches normally rely on computer vision methods to reduce the instrumentation needed. An additional camera installed on the projection unit, allows to detect direct touch input by comparing successive image frames [164].

3.1.1.2 Object Projectors

The second subclass of the *Environmental Projectors* are Object Projectors. This category is characterized by the fact that the projector and the projected display are contained completely in one object. This object can be moveable but the projector and the projected display create an environment and reference system on their own. With this preconditions, it represents a hybrid class between the *Environmental Projectors* and the *Moveable Projectors*. In contrast to the *Moveable Projectors* this class does not require any kind of set up and the projection is not created in the surrounding environment. Such Object Projectors are only possible through the miniaturization process described in the introduction.

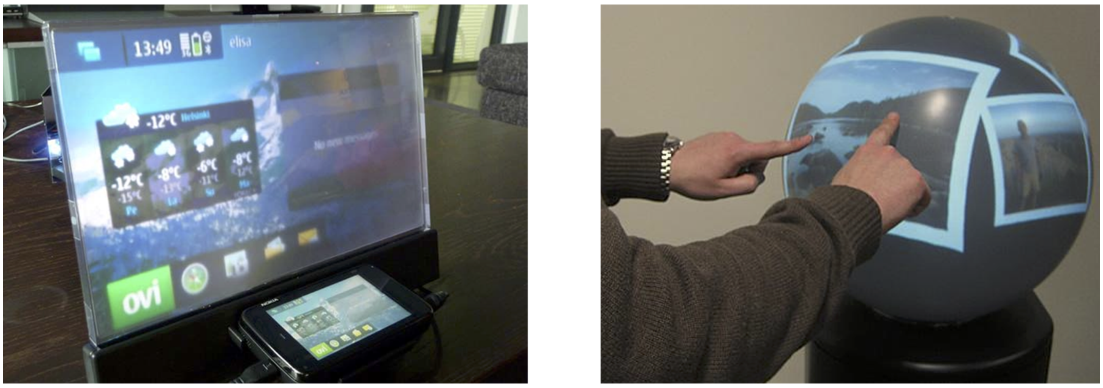


Figure 3.4: Two Object Projectors prototypes - On the left side an an external multi-touch screen for mobile phones [205]. On the right side the Microsoft Sphere prototype [21]

Virolainen et al. presented an external multi-touch screen for mobile phones [205]. They envisioned a home use case where the user can place his phone in an object that contained a projector and a projection screen. This object enables the user to interact with his phone's interface on a large-scale multi-touch screen (compare Figure 3.4 left).

The Microsoft Sphere prototype is another example of such self contained Object Projectors. Benko et al. presented a spherical self-contained projection based display that allows for direct touch interaction on the spheres surface [21]. Even though the mobility of this prototype at that time (2008) was rather limited by the high weight, current mobile projectors would allow for a smaller and much lighter prototype with the same capabilities.

Another example for an Object Projector is the Pileus system by Matsumoto and Hashimoto [138]. The system integrates a projector into an umbrella. The projection surface of this system is locate at the inside of the umbrella. The user was able to interact with the projection through either rotation gestures or through a joystick mounted in the handle of the umbrella. With the PenBook, Winkler et al. presented a tablet with integrated projector and an Anoto based projection surface [224]. The Anoto paper allows for easy pen interaction with the projection (compare Figure 3.13).

In this thesis we also present one prototype that represents the subclass of Object Projectors. Our prototype, called guitAR [122, 123], consists of a guitar and a projector mounted on the headstock. The projection surface is in this case the fretboard. Additionally, we also mention the possibility of creating a similar system using an *Environmental Projector*. Further details can be found in Chapter 4.2.

3.1.2 Moveable Projectors

The class of *Moveable Projectors* is, apart from being mobile, characterized by the requirement of a surface, on which the projector has to be positioned for usage. Furthermore they require a surface in the environment to project on. They significantly differ from the *Environmental Projectors* as they can be set-up ad hoc in many different environments. One of the simplest projectors for this kind would be a peripheral projector for notebooks that can be set-up for meetings spontaneously.

Even more sophisticated examples for such kinds of projectors exist in research as well as in current commercial products. One of the earliest prototypes are the iLamps by Raskar et al. [172]. The camera-projector combination of Raskar et al. allowed to combine a cluster of projectors ad-hoc to one big display using a structured light approach (compare Figure 3.5). The different projectors, which are connected by WiFi,

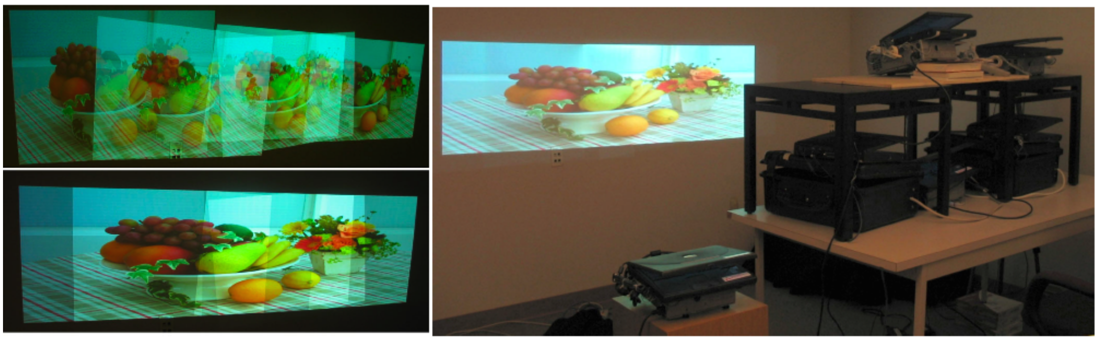


Figure 3.5: iLamps Projection Cluster - A cluster of mobile camera-projector units combined to one big display [172].

alternately project a pattern that is captured by the cameras that are connected to the other projectors. By using this method, the relative position of each projected display can be determined and afterwards combined to one larger display.

With PlayAnywhere, Wilson presented a *Moveable Projector* approach to augment surfaces without instrumentation of the environment [218]. Using a short-throw projector, which allows to create large projected displays from a short distance through a specific lens-mirror-system, Wilson developed a prototype that allowed to create large interactive surfaces that even was capable of augmenting documents placed in the projection area. Through infrared light and two cameras with infrared filters they where able to track the user's fingers touching the surface. The PlayAnywhere prototype can be seen in Figure 3.6.

The PlayAnywhere concept was recently picked up by the Taiwanese company UrRobot and they built a commercial product for kids that incorporates an interactive projection [203]. Their Robii system is an educational system that allows kids to learn through a variety of different techniques such as interactive puzzles. The system employs direct touch interaction as well and can be easily set-up on a variety of different surfaces (compare Figure 3.6). With MobiSpray, Scheible and Ojala presented a system consisting of a *Moveable Projector* and a mobile phone that can be set up spontaneously [185]. After a short calibration phase, multiple users can then utilize their mobile phone to "spray" on the calibrated projection canvas.

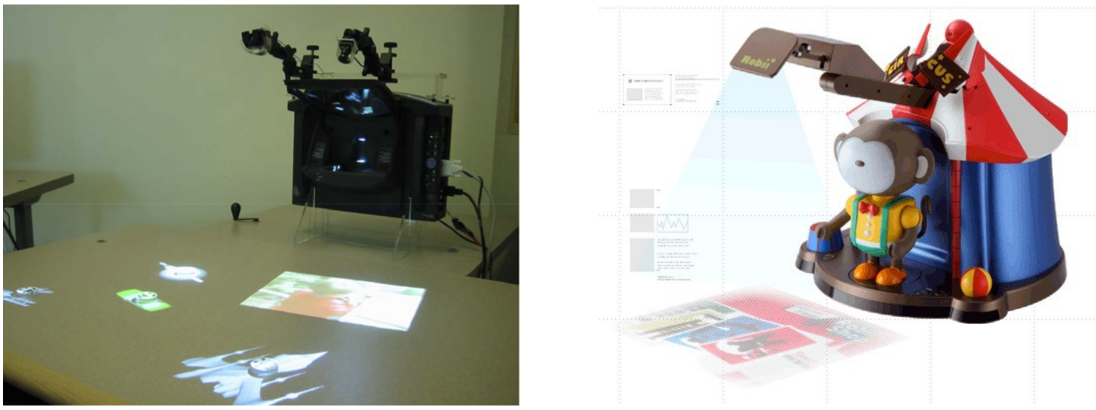


Figure 3.6: PlayAnywhere and UrRobot Robii - On the left side the PlayAnywhere prototype by Wilson [218] and on the right the commercial Robii system based on Wilsons concept [203]

In the course of this thesis we also present a system that falls into the area of *Moveable Projectors* called SurfacePhone [220]. The SurfacePhone is a novel configuration of a projector phone, which aligns the projector to project onto a physical surface to allow tabletop-like interaction. The projection is created behind the upright standing phone and is touch and gesture-enabled. Multiple projections can be merged to create shared spaces for multi-user collaboration. More details can be found in Chapter 5.2.

3.1.3 Hand-Held and Body-Worn Projectors

Most of the currently developed and investigated prototypes fall in the class of the *Hand-Held and Body-Worn Projectors*. This is mainly due to on the evolution of projection technology as described in Chapter 1. This class is characterized through the fact that the projector is fully mobile and creates projected displays on either user chosen surfaces or on the user itself (e.g. on the user's arm). The high mobility of this class comes at the drawback of a perspectively distorted projection. This problem has been solved through a variety of approaches, ranging from the creation of three dimensional models following the Everywhere Displays approach [149, 164] to sensor fusion based approaches [56]. To ensure a more accurate overview, we divide the related work in this class into two sub-classes. On the one hand we will present hand-held projectors and on the other hand body-worn projectors.

3.1.3.1 Hand-Held Projectors

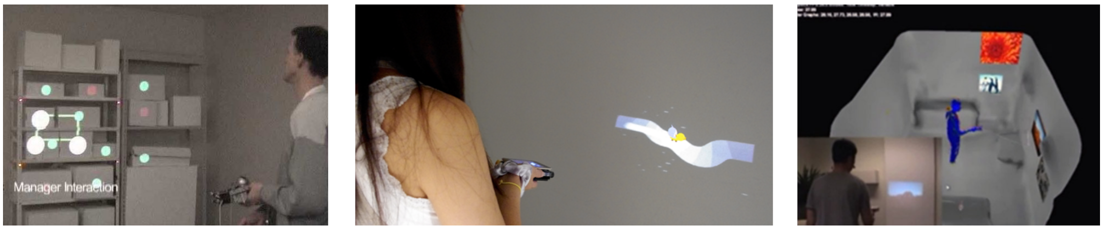


Figure 3.7: Three Applications for Hand-Held Projectors - Left: Object augmentation through the RFIGlamps [16]. Center: Character animation with MotionBeam [216]. Right: Adaptable projections based on ad-hoc generated environmental models [149]

The earliest hand-held prototype was presented by Rapp et al. with their SpotLight system [170, 171]. Albeit their prototype not being fully mobile, they presented a sophisticated system that based on dynamic peephole interaction allowed users to explore large-scale content. One of the earliest investigations of the effectiveness of hand-held projectors was conducted by Hang et al. [82]. With their prototype they were able to show that hand-held projectors can significantly help exploring large-scale content compared to a standard mobile device. Beardsley et al. investigated early AR applications for hand-held projections with the RFIG lamps [16]. The successor of the iLamps [172] enabled augmentation through infrared emitting markers spread throughout the environment. This can be seen in Figure 3.7. These three prototypes where not fully mobile but through the miniaturization more sophisticated and fully mobile prototypes became possible. With MotionBeam, Willis et al. presented a sophisticated mobile system for character interaction with handheld projectors [216] (compare Figure 3.7). Their work is based on the tradition of pre-cinema handheld projectors that they investigated before [214].

Another interesting approach that combined the advantages of *Hand-Held and Body-Worn Projectors* and *Environmental Projectors* was presented by Virolainen et al. [204]. Their Burn-To-Share prototype allowed for easy sharing of pictorial content on public surfaces. Therefore the user simply projected a picture with their mobile projector on to a surface that is equipped with an *Environmental Projector* and a camera. The projected image is captured with the camera and when the user switches off his mobile projector the *Environmental Projector* would start projecting the before projected image.

Molyneaux et al. [149] presented a system that is Based on the KinectFusion algorithms [99]. They create an ad-hoc model of the environment with either a hand-held Kinect-projector unit or several Kinects installed in the environment. This model allows to remove the problem of the distortion and creates a perspectively corrected projection. Through three different prototypes they explore novel interaction and augmentation techniques of mobile projected displays. A screenshot of the model the Kinect creates and the undistorted projection can be seen in Figure 3.7. Even though their prototype represents the technically most sophisticated system today, Molyneaux et al. did not evaluate their system with users.

Figure 3.8: Multi-User Interaction with Hand-Held Projectors - On the left the system of Cao et al. merging multiple projections in one information space [38]. On the right the Side-by-Side system allowing multiple users to interact with multiple projectors without any instrumentation [215].

Through the decrease in price of mobile projection hardware it became possible to investigate multi-user scenarios. Cao et al. tracked multiple hand-held projector in the environment to allow for co-located multi-user interaction [38]. Besides several new interaction techniques they introduced the concept of merging multiple projections in one information space. This can be seen in Figure 3.8. The drawback of their system is the requirement of the high instrumentation through an expensive Vicon tracking system. To circumvent this drawback Willis et al. followed a different approach with their Side-by-Side system. It is designed for ad-hoc multi-user interaction with handheld projectors [215] (compare Figure 3.8). The tracking of multiple independent projected images in relation to each other is accomplished by projecting invisible fiducial markers through infrared light. Therefore Willis et al. replaced two of the color LEDs in the projection unit with infrared LEDs.

As this thesis mainly focusses on hand-held projected displays, several explorations can be found in this thesis as well. While Chapter 5 and 6 present investigations on perception and spatial interaction concepts, in Chapter 7 we present an approach to widen the field of mobile projection to a variety of use cases.

3.1.3.2 Body-Worn Projectors

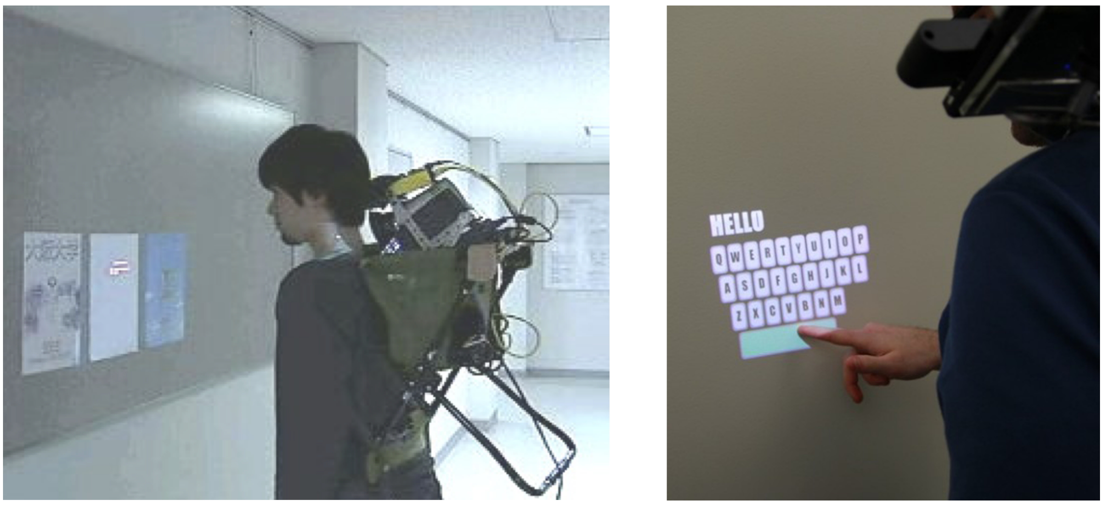


Figure 3.9: Body-Worn Projector Systems - Left: Early investigations of a Body-Worn Projector by Karitsuka and Sato [104]. Right: The OmniTouch system allowing unrestricted touch input for a Body-Worn Projector [84]

Karitsuka and Sato [104] presented to the best of our knowledge the earliest Body-Worn Projector prototype. With a projector mounted over the shoulder of the user the system allowed for spontaneously created projected displays as can be seen in Figure 3.9. Even though at this time, the available projection technology was from a large size, their system was able to project undistorted images onto with infrared reflected markers equipped surfaces. Based on a infrared emitting spotlight and a camera the system was even able to track user input using infrared reflective markers placed on the fingers of the user. A similar approach was taken up by Mistry et al. [146]. The Wear-Ur-World system tracked the fingers in front of a projector necklace using coloured finger caps.

Other possibilities to sense the users input for Body-Worn Projectors have been investigated by Harrisson et al. [84, 85]. Skinput allowed to interact with a shoulder worn projector that created a display on the users forearm. To sense the location of finger taps on the arm and hand Harrisson et al. analysed mechanical vibrations that propagate through the body. The drawback of Skinput was that the users had to wear a sensor on the arm. Therefore Harrisson et al. replaced this sensor through a depth camera that was shoulder worn as well in their OmniTouch system [84]. Besides sensing on-body touch input, the system also is capable of registering touches on arbitrary flat surfaces. This approach was later extended by Winkler et al. with AMP-D [223]. In contrast to OmniTouch, AMP-D is an ambient information display that constantly projects information onto the floor in front of the user. If the user perceives something interesting in the floor projection he can interact with the information through hand-gestures.

An extensive overview of suited positions for projectors worn on the body based on the visibility of the projected display, has been presented by Ota et al. [162]. They conclude that the shoulder is one of the best suited spots for a Body-Worn Projector, this is in line with most of the aforementioned work.

3.1.4 Self-actuated Projectors

The last class represents the most mobile kind of projectors, the *Self-actuated Projectors*. Even though the class of *Hand-Held and Body-Worn Projectors* already represents fully mobile projectors that except of the projected displays are independent from the environment, *Self-actuated Projectors* are going one step further as they are not only fully mobile but actually can move through the environment on their own. Basically this can be done through robotic components that allow the projector to move on its own without human control either through the air, on land or on water. Such independent drones have – especially if they are terrestrial – the advantage that they can carry a larger projector and other bigger sensor that allow to project undistorted bright images. This is for *Hand-Held and Body-Worn Projectors* not always possible as they have size limitations to be still small enough to be hand-held. So far two different kinds have been explored, terrestrial *Self-actuated Projectors* and aerial *Self-actuated Projectors*.

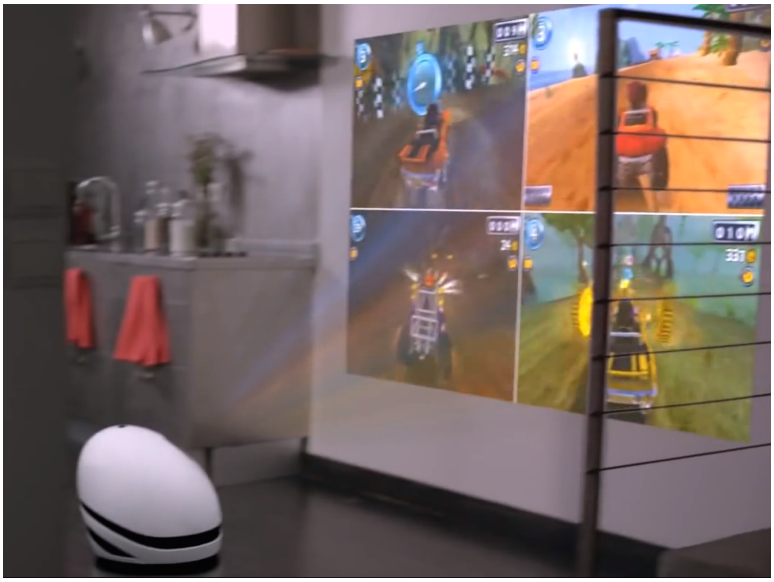


Figure 3.10: Keecker Autonomous Robot Projector - The Keecker product design envisions an autonomous robot that is capable of projecting in-situ needed displays on nearby surfaces [108].

3.1.4.1 Terrestrial Self-actuated Projectors

The vision of projectors being integrated into robots is quite old. R2D2 from the Star Wars movie trilogy was capable of projecting holographic images while at the same time being an autonomous terrestrial robot [210]. The R2D2 idea was recently picked up by Keecker [108]. Even though currently they only have a design vision (which can be seen in Figure 3.10), they want to commercialize an autonomous robot that can be controlled and called using a mobile device.

3.1.4.2 Aerial Self-actuated Projectors

Besides these autonomous driving robots recently also flying drones with projectors have been explored. Scheible et al. presented the Displaydrone a mobile projector mounted on a multicopter [183]. The content of the projection was transferred to a phone that was connected to the projector. As the multicopter needed to be flown manually they only carried out an experimental evaluation and obtained first impressions on how the Displaydrone is perceived by the viewing audience. An example of these projected flying displays can be seen in Figure 3.11. Recently they also replaced the projector with a physical display, in this case a tablet [187]. Floating Avatar by Tobita et al. is a prototype that represents a hybrid class between the *Self-actuated Projectors* and the Object Projectors [201]. Their system consisted of a flying blimp with an integrated projector in the blimp corpus. The projector created a projected display on one surface of the blimp's corpus. The system is meant to be a used as a tele-presence system. As the projector and the projection surface are completely contained in the object, the Flying Avatar system can also be classified as an Object Projector. Since the blimp was capable of autonomously following other users in the environment, it also is an instance of the class of *Self-actuated Projectors*.

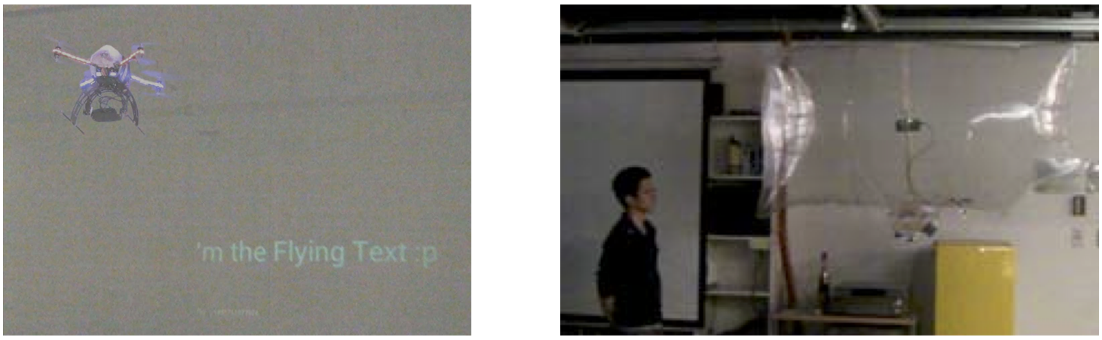


Figure 3.11: Flying *Self-actuated Projectors* - On the left side the Displaydrone of Scheible et al. [183], on the right side the Floating Avatar system of Tobita et al. [201]

3.2 Interaction Techniques for Projected Displays

At this point we want to review existing interaction techniques for mobile projected displays. Therefore we focus only on projected displays that are created by *Moveable*

Projectors or projectors of even higher mobility. We classify the interaction techniques into six different classes. Compared to the classification introduced by Rukzio et al. [182] we present two additional classes (*Bi-Manual manipulation of Surface and Projector* as well as *Around the device interaction*) as through current developments their classification does not cover all possible interaction techniques any more.

When we refer to interaction with mobile projected displays, we consider the projection to be the output, not a way of triggering actions on other devices. One example for such a technique would be the work of Schmidt et al. [186] called PICOntrol. They used light-sensing diodes to remotely control electronic appliances through the light emitted by a mobile projector. Very similar Hosoi et al. presented with CoGame a technique to steer a robot by projecting the path in front of it [91]. For our categorization of interaction techniques, PICOntrol, CoGame and techniques that are alike, are not considered. We only focus on techniques that are used as a mean to interact with the content on the projection itself not the environment or objects placed in it.

For our classification we divide the existing interaction techniques into six categories, which are the following:

- *Direct Interaction on the projection*: This class covers all cases where the user is able to reach the projection with one of his extremities and manipulate it directly.
- *Input on the projection device*: This class represents techniques where the user interacts with the projected display through input on the projecting device.
- *Movement of the projector*: In this case the projected display is controlled through relative- or absolute movement of the projector.
- *Manipulation of the projection surface*: Here all techniques are included where the interaction is triggered through manipulation of the surface that contains the projected display.
- *Bi-Manual manipulation of Surface and Projector*: This class incorporates techniques that rely on simultaneous manipulation of the surface where the projected display is created and the projector itself.

- *Around the device interaction*: The last class covers techniques where the interaction is triggered by movements of the user in the vicinity of the projector, without either touching it or the projected display.

In the following we will present several techniques of the different categories. Some of the presented techniques have been applied and analysed in this thesis as well.

3.2.1 Direct Interaction on the projection

Interaction techniques covered by the class of *Direct Interaction on the projection* are characterized by the fact that the user can reach the projected display with one of his extremities. This fact allows him to directly manipulate and interact with the content in the projection. This comes at the drawback that the user has to stand rather close to the projection. This means in cases of *Hand-Held and Body-Worn Projectors* that the projection is limited in size, especially since for current projection units no short-throw lenses exist.

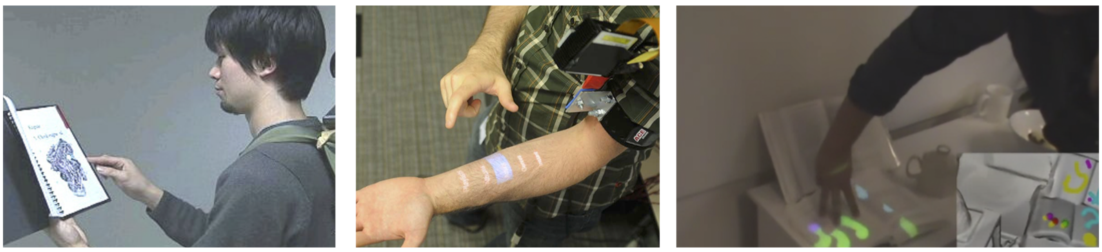


Figure 3.12: Direct Touch Interaction on the Projection - Left: Karitsuka and Sato presented a system that employed a infrared reflective finger cap to register touches [104] Center: Harrison et al. used muscle sensing techniques to detect on body touches [85]. Right: Based on an environmental model Molyneaux et al. identified the position in the projection the user is touching [149]

The most frequently used technique in this category is the interaction with the projected display by touching it. One of the earliest systems utilizing a mobile projector that allowed for touch interaction was presented by Karitsuka and Sato [104]. There implementation employed infrared reflecting finger caps to track the users finger position relative to the projected display (compare Figure 3.12). The PlayAnywhere system employed infrared light as well, but Wilson's finger tracking algorithm is based on shadow tracking and did not require the user to wear anything on his fingers [218]. Another system that allowed for interaction in this way was presented by Winkler et al. [222]. They investigated how a user can interact with the content on his projector phone during a call by touching the projected display. But their system even employed an optical tracking system installed in the environment and required the user to have a marker attached to the finger as well.

With the current advantages in depth camera technologies and the recent commercialization of the Kinect a variety of different algorithms emerged that allow to register the users touch on the on different objects. One of the most sophisticated algorithms employing such a depth camera was presented with OmniTouch by Harrison et al. [84]. In their earlier prototype called Skinput they used a different sensing technique relying on muscle input [85] (as can be seen in Figure 3.12). Of course this sensing technique only works on the body parts that are equipped with such a sensor and therefore does not allow to touch projected displays created in the environment. Very similar to OmniTouch the systems presented by Molyneaux et al. also allowed for touch input based on the data of the Kinect depth camera and their environmental model [149]. Their system can be seen in Figure 3.12.

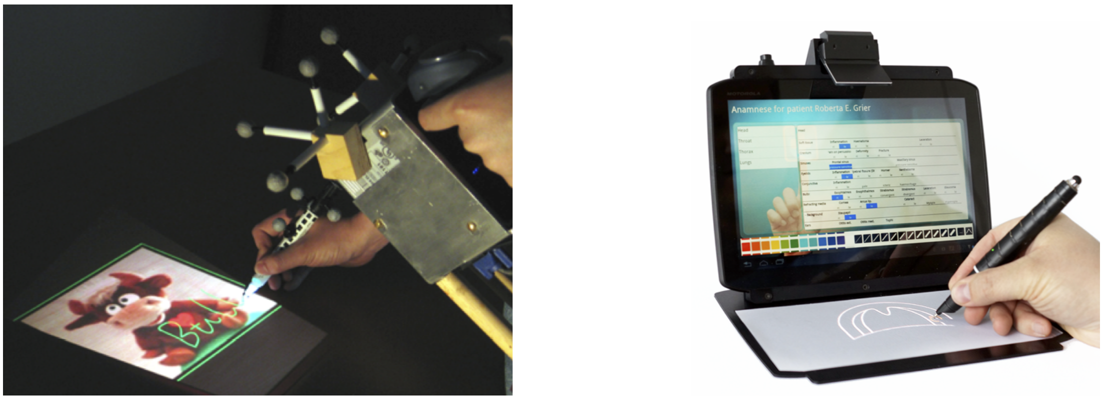


Figure 3.13: Pen Input on the Projection - Left: Cao and Balakrishnan used a Vicon tracking system to track projector and pen to allow for intuitive input [37] Right: The PenBook uses an Anoto pen for interaction on the projection [224]

Another frequently used approach for *Direct Interaction on the projection* is input through a pen. Pens provide the advantage that they can incorporate markers to allow for easy tracking. Additionally writing with a pen is one of the oldest interaction techniques known to man. Cao and Balakrishnan employed this technique using the Vicon system that also tracked their projectors [37]. This can be seen in Figure 3.13.

In Winkler et al.'s PenBook an Anoto pen was used that allow tracking its own position on the Anoto paper based projection surface [224] (compare Figure 3.13 right). Furthermore, with Interactive Dirt McFarlane et al. presented a body-worn system that allowed for input through a telescopic metal pointer [141]. By equipping the telescopic pointer with an infrared diode it was possible to track its position on the projection using a WiiMote.

Even though direct touch and pen input is one of the most compelling interaction techniques as it is quite well known from current devices such as smartphones or tablets, other direct interaction techniques have been investigated as well. Cauchard et al. presented a foot interaction technique that allows the user to step on the interface elements [42]. An example of this interaction can be seen in Figure 4.4. This technique was enabled through a steerable projection that allowed to project onto the floor in front of the user while comfortably holding the device in the hand. The tracking of the foot is done using the camera of the projector phone. Due to the size of the foot, this technique requires large sized user interface elements for precise input. But for ambient information displays or quick interactions with a body-worn projector such as the AMP-D system [223] this technique would be well suited.

In this thesis two different prototypes were developed that rely on direct touch interaction on the projection. We present two novel algorithms to register touches on the projected display. The one of the SurfacePhone [220] is only based on the standard camera and the accelerometer incorporated in today's smartphones and can be found in Chapter 5.2. The second algorithm is following the depth camera based approach, but our system follows a more light-weight approach to ensure the easy adoption for mobile devices. The detailed description can be found in Chapter 7.

Figure 3.14: Input on the projection device - Left: Current projectorphones and detachable projection cases rely only on input on the touch-screen of the device [165]. Right: An interaction technique to allow in-call interaction directly on the projecting device [129].

3.2.2 Input on the projection device

The class of *Input on the projection device* represents all interaction and manipulation techniques with which the user interacts with the projection device itself. Of course this class requires the projector to have input modalities. Even though most pico-projectors do not provide this feature, projectorphones such as the Samsung Galaxy Beam [76] fall into this category as they only rely on input on the touch-screen (compare Figure 3.14). *Input on the projection device* has the advantage that the sensors, such as a touch-screen, a joystick or keys and buttons can be integrated into the device and that they do not rely on complicated sensing techniques. Furthermore, the user can be further away from the projection as they don't have to be able to reach it as opposed to the *Direct Interaction on the projection* class. This allows for larger projected displays.

Most *Hand-Held and Body-Worn Projectors* allow for this kind of interaction as well but normally in combination with another technique. The system developed by Cao and Balakrishnan featured button input on the projector as well as *Movement of the projector* [37]. SideBySide follows a similar approach [215]. On the one side the user was able to move the projection but to trigger actions on the other side he had to press a button on the projector. We investigated input on a projectorphone with a back touch screen [129]. Very similar to the work of Winkler et al. [222] this system allowed to interact during a phone call when the device is held next to the ear. In contrast to

Winkler et al., their system did not require an optical tracking system. This approach can be seen in Figure 3.14.

3.2.3 Movement of the projector

Due to the increasing mobility of the projection technology, *Movement of the projector* as a mean of manipulating the content in the projected display became possible. In this category all techniques that rely on the movement of the projector are covered. Due to the spatial coupling of the projector and the projected display, the movement of the projection unit naturally comes with a movement of the projected display. This approach allows for larger projected displays as the user can stand further away from the projection surface compared to *Direct Interaction on the projection*. Still, if the content on the projection is manipulated through movement, the size of the projected display has to be significantly smaller then the projection surface, as the movement will need space as well. Additionally the movement comes, as discussed above, at the cost of a distorted display which is often not coped for in current prototypes.

The concept of moving a projector to create life like animations has already been employed with the *laterna magica* [214]. A brief overview of such methods can also be found in Chapter 1 of this thesis. One particular concept that often depends on movements when applied to a mobile projection approach is augmented reality. As the content that one wants to augment is often larger then the projected display, the user needs to relocate the display to fully explore the augmentation. To cope for this distinction we will first give an overview of general user interfaces and then present AR prototypes relying on mobile projectors.

3.2.3.1 General Interfaces

To the best of our knowledge Rapp et al. presented with Spotlight the earliest form of such a movement based interfaces for mobile projectors [170, 171]. Their early hand-held prototype allowed to explore large scale content such as a calendar, by moving the projection unit. They employed the so called dynamic peephole interaction that has been quite extensively investigated for mobile devices [39, 92, 179, 230]. In this technique, the viewport is physically moved above the virtual workspace that is static with respect to an external frame of reference (compare Figure 6.11). This requires a spatially-aware display to detect which part of the virtual workspace has to be shown. One of the advantages of this technique is that it exploits the micro- and in larger environments also the macro-spatial memory of the user. Therefore this particular technique has also been employed by Cao et al. for multiple projectors that shared one large virtual workspace [38]. Their optically tracked projectors shared one workspace that was distributed in a larger room, in which objects such as tables could be exploited as memory aids. Instead of a hand-held projector, Blasko et al. adopted the dynamic peephole technique in their studies of a wrist-worn mobile projector [25]. Even though their prototype was only simulated by a short throw beamer, their investigations showed a lot of suitable interaction techniques for such a projection unit mounted in a smartwatch. Beside the peephole technique they also investigated scrolling and other navigation techniques for websites that are based on rotating and moving the wrist where the future projector might be mounted.

As the dynamic peephole techique is often seen as one of the most promising interfaces for mobile projectors Kaufman and Ahlström developed a model for target acquisition using a projector and a dynamic peephole [106]. Their model was based on prior findings for mobile devices. In later studies they also showed that the technique can aid macro spatial memory for the figural space [107]. Even though their studies reveal several advantages of such a dynamic peephole interface, until today no comparison of a dynamic peephole interface and other interaction techniques exist. In this thesis we present the design and evaluation of a sophisticated peephole interface and compare it against a touch-based interaction. In Chapter 6.2 you can find evidence that even though this technique is very promising, it is not the most effective technique for mobile projection units in terms of task completion time.

When applying the dynamic peephole metaphor, one has to take into consideration that this technique is only meant to be used for navigational tasks. While the panning and zooming of the content can be managed by the spatially-aware display, selection normally is done using a *Input on the projection device* technique. All aforementioned prototypes either used buttons or a touch screen on the device for input, even though gesture tracking of the movement might be a suitable solution for this problem as well.

Interfaces where the *Movement of the projector* is the mean of input have not only focussed on the dynamic peephole so far. With Motionbeam, Willis et al. presented a technique where a character that is contained in the projection could be animated by movement [216]. As already mentioned, their technique is based on pre-cinematic techniques of moved projectors [214]. The SideBySide prototype is another example for such interactions [215]. Through the movement of the projection Willis et al. allowed multiple projectors to interact with each other. The same concept was also applied by Shilkrot et al. [189]. But instead of the invisible infrared markers of SideBySide, their prototype used visible markers in the projection to retrieve the relative positions of the projectors.

Besides the already mentioned dynamic peephole investigation we also present a technique that is exploiting the users micro spatial memory through movement [41]. Instead of using one virtual workspace that is placed in the environment we allow the user to distribute several workspaces relative to him that are always shown completely. This technique is shown to be highly effective and reduces the temporal demand that exist for switching between several applications on current smartphones. More details can be found in Chapter 6.

3.2.3.2 Augmented Reality Interfaces

In most cases where mobile projection is used for AR applications, the area that needs to be augmented is much larger then the projected display. Therefore the display needs to be moved to reveal the whole augmentation. With this, the AR overlay is revealed similar to the dynamic peephole interaction discussed above. One of the earliest cases where a mobile projection unit was used to create an AR overlay are the RFIGlamps by Beardsley et al. [16]. Through infrared emitting markers in the environment the system was able to augment real world objects as can be seen in Figure 3.7. A far more sophisticated prototype was shown by Ni et al. with AnatOnMe [158]. Their system

Figure 3.15: Augmented Reality Interfaces - Left: The AnatOnMe system allows for better doctor-patient communication through on-body projection [158] Right: Twinkle is a system that reacts on users motion and collisions with real world objects [231]

tried to ease doctor-patient communication by allowing to project detailed medical information directly onto the body of the patient, which is illustrated in Figure 3.15. This helps for example to communicate the procedure of an invasive operation. Yoshida et al. presented with Twinkle a mobile camera-projector unit that allowed to augment an arbitrary physical surface [231]. Through a variety of computer vision algorithms, the system recognizes the features of the physical environment and displays images and sounds that are generated based on the user's motion and collisions of projected images with objects. This concept is based on findings that we will also present later on in this thesis as we developed a very similar system [131]. An example for this can be seen in Figure 3.15. A body-worn AR system that relied on motion for interaction was presented by Krum et al. [114]. By moving through the environment the user is able to reveal different information that is only visible to the user through retro reflective material.

In this thesis we present not only an in-depth analysis of requirements for mobile projection based AR applications, but we also several prototypes that employ motion as a mean of interacting with the projected content [128, 131, 132, 133, 188]. More details can be found in the next chapter.

Figure 3.16: Manipulation of the projection surface - The Cobra system that allows to interact with the projection content by manipulating the projection surface [229].

3.2.4 Manipulation of the projection surface

The interaction class where the user is the closest to the projected display is represented through the *Manipulation of the projection surface*. As the user does not only need to carry the projection device but also the projection surface, this class effectively can be seen as an Object Projector with the user being the object. Here all techniques are included where the interaction is triggered through manipulation of the whole surface. Especially with recent developments of deformable or organic user interfaces, many prototypes rely on such a setting to explore possible interaction techniques [181]. With increased sensing capabilities different manipulation techniques became possible for such settings [193]. Nevertheless, in terms of prototypes that envision the projector as the main output and not only as a prototyping device only one system exists that uses *Manipulation of the projection surface* as a mean of input. The Cobra system presented by Ye and Khalid consists of a shoulder worn projector that allows for input through deforming and moving the a projection surface relative to it [229]. Cobra sensed the position of the projection surface relative to the projector by infrared LEDs in the surface and a WiiMote on the shoulder. Additionally, the system is able to sense deformations of the surface by flexsensors integrated into the surface. This allows for a very flexible input with a huge expressiveness. Ye and Khalid prototypically implemented an ego shooter for this setting which can be seen in Figure 3.16.

3.2.5 Bi-Manual manipulation of Surface and Projector

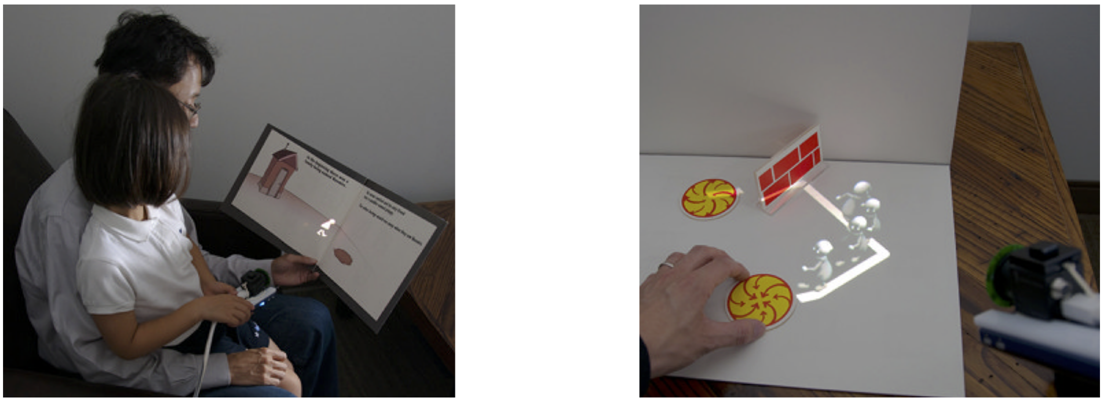


Figure 3.17: Bi-Manual manipulation of Surface and Projector - HideOut is a system that allows to augment surfaces with invisible infrared reflective markers. This enable tangible interaction with the surface while providing also the possibility to move the projector [217]

The class of *Bi-Manual manipulation of Surface and Projector* represents a combination of the classes *Manipulation of the projection surface* and *Movement of the projector*. It incorporates techniques that rely on simultaneous manipulation of the surface where the projected display is created as well as the projector itself. Even though this requires the user to be very close to the projected display in order to be able to touch the surface, this combination increases the expressiveness of the interaction drastically. Nevertheless, it has often been neglected in current investigations.

Willis et al. presented another system of this class, called HideOut [217]. As one of their main use cases, they present the possibility of interactive story telling. To enable an expressive and intuitive way, they incorporate tangible interaction with real world objects into the system. The tracking of these objects is accomplished by infrared reflective markers that are invisibly painted onto the objects. The projector is equipped with a camera and an infrared spotlight. By moving the projector, or the objects the user can manipulate and develop his own story. The working system can be seen in Figure 3.17.

In this thesis, we also present a system that requires the user to interact with the projector as well as the projection surface. In Chapter 4 we present LittleProjectedPlanet, a system that allows to project physical simulations on the real world. By manipulating the properties of the real world, the user can interact and change the simulation. The concept and implementation is further described later on and is illustrated in Figure 4.7.

3.2.6 Around the device interaction

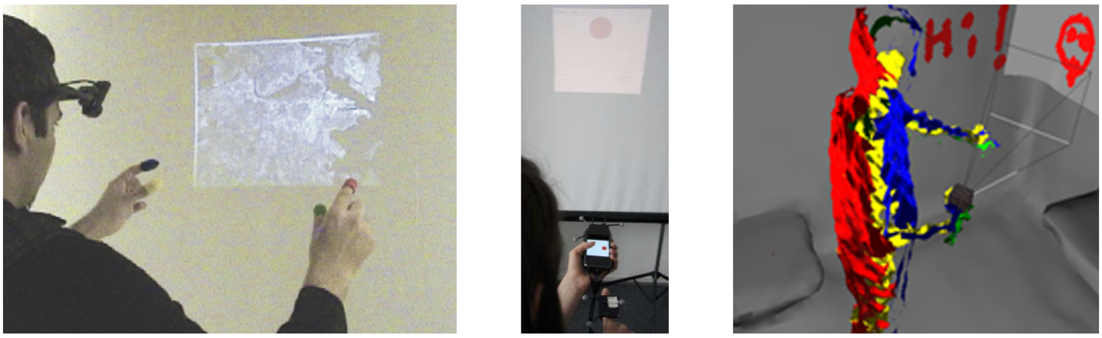


Figure 3.18: Interactions around the projecting device - Left: The SixthSense system allowed for gestures executed between the projector and the projection surface [146]. Center: Winkler et al. analysed different mid-air pointing techniques using an optical tracking system [221] Right: Such mid-air pointing techniques have been implemented by Molyneaux et al. [149]

The last class, *Around the device interaction*, is characterized by the user neither interacting on the projected display nor the projecting device. The user can stand at arbitrary distances for this interaction as there are no requirements in terms of reachability or needed surface size. The techniques in this class make use of the space around the projector through mid-air interactions. This allows for a variety of different techniques but all of them require sophisticated sensing techniques. But with the recent developments and devices such as the Kinect and Leap Motion, this sensing became possible even though these technologies have not yet been incorporated into mobile devices. On of the earliest systems that represents this class would be the SixthSense system by Mistry et al. [146]. A body-worn projector could be controlled by conducting gestures between the projector and the projected display as can be seen in Figure 3.18. A very similar technique was explored by Cowan and Li, but their ShadowPuppets system relied on casting shadows onto the projected image [53]. These shadows were detected by a camera through a comparison of the projected and the captured image. This approach has the advantage that several users can cast a shadow and with that multi-user interaction is naturally supported. The disadvantage of this technique is that the casted shadow can be quite big and with that the input is very imprecise. Therefore Cowan and Li argued that only gestures for e.g. panning and zooming should be supported through this technique [53].

The aforementioned system of Molyneaux et al. also allowed for the shadow casting technique of Cowan and Li, but they also incorporated mid-air pointing as one possible interaction [149]. By tracking the user with a Kinect installed in the environment they were able to sense mid-air pointing. Their system was capable of transforming gestures and pointing direction into input for a painting application. This input was even possible outside the current projected canvas that could be revealed afterwards by pointing the projector in that direction. An in-depth analysis of such mid-air pointing techniques has been conducted by Winkler et al. [221]. By comparing different positions for such pointing, they found that interacting behind the phone yields the highest performance, albeit showing a error rate that was twice as high. In further investigations they were able to show that such techniques can have a significant advantage for different mobile applications.

3.3 Summary

In this chapter we presented prior work that is related to the work of this thesis. Firstly, we categorized different projected displays by their level of mobility. Additionally, we classified existing interaction techniques for mobile projected displays. We hope that these categorizations can aid other researchers firstly as a way of distinguishing their own work from already existing approaches and secondly help them finding areas that are rather unexplored yet. Thereby we contribute with a novel way of characterising interactions with mobile projected displays as well as the kind of projected display itself. In the following, we will keep these categorizations for our own techniques and prototypes to give an immediate understanding of the setting.

4

Augmented Reality Application Concepts for Mobile Projected Displays

One of the main goals of this thesis is to create an insight on applications that enable the mobile projected display to seamlessly blend in with the spatial features of the environment of the user. The concept of augmented reality allows for a direct combination of digital content with real world objects. As discussed in Chapter 2, projection based AR displays provide one of the highest rates of immersion as the AR layer is directly created on the object. Therefore AR applications for mobile projected displays present the highest rate of blending between the real world and a mobile projected display possible. At the same time such AR applications also present the most challenging type of applications. The registration of the projected display with the real world has to be as accurate as possible to ensure a good user experience. This requires the highest amount of spatial awareness of the projection unit. Therefore we will now focus on hardware - and interface designs for AR applications instead of other more general applications that require a lower level of spatial awareness.

The designs and results of this chapter have been published previously in the following publications: [122, 123, 128, 130, 131, 132, 133, 188].

4. AUGMENTED REALITY APPLICATION CONCEPTS FOR MOBILE PROJECTED DISPLAYS

Mobile projectors can facilitate AR applications in a variety of different scenarios such as pointing one towards the object one is searching for. The new output capabilities of these pico projectors equipped devices provide a rich design-space for AR applications and games as we already discussed in Chapters 2 and 3. They have the ability to augment objects through overlaying projection and with that, they can overcome the problems that we are facing today when creating AR applications. Projecting a dynamic overlay directly onto a surface of the real world may enhance the possibilities, even though it can be hard to identify the projected overlay in bright light. The usage of a head-mounted-display (HMD) provides excellent results in terms of augmentation but it is also cumbersome to use and sometimes straining the users. Furthermore as a consequence of the display being attached to a single user, applications using a head mounted display can only be used in multi-user scenarios, when a large amount of hardware is used. Another common technique for dynamic AR overlays is to use the screen of the mobile device as a magic lens [23]. But in these scenarios one has to struggle again with the small size of the device. Moreover, such a magic lens display is not really enjoyable to use with more than one user, as the device has to be passed around to allow the overlay to be well perceivable.

To empower pico-projectors to augment the real world around them, the projector needs a concrete model of its surroundings. For this several techniques exist. The most suited one for the form factor phone, is the usage of computer vision based methods, since most modern mobile phones are already equipped with powerful cameras. When using computer vision based approaches, different possible spatial layouts of camera and projector unit unfold, and with that different possible AR applications. In this chapter we identify different application types based on the spatial configuration of the camera and the projector for hand-held projector phones. As part of this design considerations, we derive different application types using different spatial layouts of cameras and projectors: *congruent setups, partially intersecting setups and disjunct setups.* Such a classification is useful to structure the design space of possible AR applications for projector phones.

With mobile projection, also the problems of distorted projection caused by non-orthogonal projection angles or hand jitter arise. Furthermore everyday life objects often have non-planar surfaces that have to be taken into account as well. These problems are out of the scope of this thesis, but plenty of research has already focused on these problems, e.g. [56].

The remainder of this chapter is structured as follows. Subsection 4.1 describes the different hardware design considerations of these AR interfaces. In this conceptual section we also discuss how the spatial layout of the camera relative to the mobile projection unit can affect the characteristics of applications for this new sort of hardware. In Subsection 4.2 we describe several applications that evolve from the developed concepts.

4.1 Design Considerations

As discussed in the introduction, today's projector phones are very limited in terms of suited applications and interaction techniques. To fully exploit the potential of mobile projection, we classify different spatial layouts of the camera position relative to the projector unit and discuss the impact on the AR interaction techniques facilitated by these layouts.

We first want to define the terms *camera field of view* (FoV) and the term *field of projection* (FoP) for easier discussion of different layouts. The FoV of the camera is defined as the area the camera is able to capture. The FoP is the area the projector is able to project on.

Generally one can distinguish between three different spatial layouts: First setups where the FoV and the FoP do not overlap are categorized as disjunct since the projection is aligned in a different direction than the FoV of the camera. Setups where the FoV of the camera and the FoP overlap are categorized in two different classes, partially intersecting and congruent. In both setups the direction of camera and projector is the same, they only differ from the configuration of the lens of camera and projector or their distance to each other. If the FoV of the camera overlaps partially with the projected field, then it is categorized as an intersecting projection. In the third category

– congruent – the entire projected field is situated within the image produced by the camera. Due to different hardware specifications of cameras and projectors (different projection angles, aperture, and others properties) the actual spatial setups could be very different. Today, due to the technical limitations, just disjunct setups exist. This is mostly due to the needed size of the projection unit. but keeping in mind that current 3 megapixel camera modules are approximately 1.5mm high an integration next to a projector is possible. We are convinced that the partially intersecting and congruent layout provide a lot more potential for new interactions as we illustrate in the following paragraphs.

4.1.1 Disjunct Alignment

Figure 4.1: Disjunct alignment of camera and projector. - The FoV and the FoP have no shared space.

In the case of disjunct alignment, camera and projector are often attached to two different sides of the mobile device. As a result, the visual field of the camera and the projected image are not overlapping. This setup is the most common in todays projector phones such as the Samsung Galaxy Beam [76] (compare Figure 4.1).

The alignment described is rather unsuitable for the augmentation of physical objects. While the camera tracking could identify the objects located in the FoV, the FoP is directed towards a different angle, therefore direct projection onto the objects in the FoV is not possible. There are two ways of overcoming this problem. The first possibility is determining one's own position in relation to the position of an object by means of a digital model of the environment and subsequently being able to augment the whole environment. This approach, however, requires the availability of a spatial model of the environment at all times. Furthermore, this procedure causes a considerable restriction of mobility.

The second possibility is to create AR applications using a camera and projector setup as described above, is adapting objects or taking advantage of the physical structure of the object in order to augment it. For example, an optical marker, which can be identified and interpreted by the camera, could be attached to the first page of a book, resulting in the projection of additional information onto the open cover of the book. This would enable users to quickly and easily access reviews, summaries, and other services.

A benefit of systems which use disjunct projection is that they allow for optical flow tracking. This would enable *Movement of the projector* based interaction techniques. For the other classes this would only be possible by sacrificing projection space, as projection within the FoV would impair the optical flow tracking process. Additionally also alternating between a frame where the projection is active and a frame that is captured by the camera would be possible. But this technology would require a projector that has a high frame rate as well as the possibility to synchronize camera and projector. Current pico projectors only have up to 30FPS and can not be reliably synchronized. Therefore such a multiplexing of camera and projector frames are unsuitable for current hardware but can provide a possibility in the future. How such movement sensing techniques can be exploited is described in Chapter 2.2.

4.1.2 Partially Intersecting Alignment

Figure 4.2: Partially intersecting alignment of camera and projector. - The FoV and FoP are partially overlapping each other.

In the case of partially intersecting projection the FoV and the FoP are situated on the same level, partially overlapping each other as shown on Figure 4.2. By knowing the angle of aperture of the camera and the projector's lens, the size of the FoV and PoV, as well as its misalignment, the overlapping area can be calculated. This kind of projection is the most suitable for the augmentation of real world objects. The fact that the FoP might just minimally affect the visual characteristics of the object, the image produced by the camera makes the stable use of visual trackers impossible. However, this is only suitable for the augmentation of bigger-sized objects. With smaller-sized objects the area that can be used for augmentation can be too small for augmentation and tracking at the same time. The projection would change the appearance of the object to radically such that there are not enough features for the optical tracking.

The Map-Torchlight application uses a partial intersecting setup for the projection of additional Points of Interest (POIs) on a large paper map and Shelf-Torchlight for projecting additional information to products in a retail environment (compare Figure 6). An application which assists someone in fixing, e.g. the engine of a car, could be realized in a very similar way. By attaching visual markers to the engine compartment the it is possible to determine the position of the projector phone relative to the engine, so that it can mark, for example, the screws which shall be removed in a particular step of a certain task.

The advantages (or differences) of a partial intersecting setup compared to a congruent setup is that the FoV and the FoP are areas which have just a small effect on each other, so that the benefits of both can be exploited. For example the non-overlapped area in the camera FoV can be used to allow gesture-based interaction (as proposed by Mistry et al. [146]) without interfering with the projection.

4.1.3 Congruent Alignment

A congruent setup is given when camera and projector are attached on the same side of the mobile phone, and the entire FoP is contained in the FoV. This is the difference to the case of the partially intersecting alignment (see Figure 4.3). A disadvantage of this spatial configuration is that the projection could influence the processing of the camera image. When an object is augmented the projection changes the visual appearance of the object and by this it can interfere with the optical tracking. However, the congruent projection enables the user to interact directly with the projection without any limitation. The application LittleProjectedPlanet, as described in section 5, uses this spatial configuration. It allows for *Direct Interaction on the projection* and with that it enables the user to operate the projection through the modification of physical objects. Another domain for a synchronous projection setup could be an OCR application which recognizes and marks spelling mistakes. School children would be

Figure 4.3: Congruent alignment of camera and projector - The FoP is completly included in the FoV

able to control their homework by holding their mobile phones with the integrated projector upon their exercise books on which the projector could mark the mistakes and give additional information about them. However, the realization of such a system as an end product for costumers will take considerable research. The main problem that hinders such an application is the current lack of robustness of handwriting recognition.

Generally the effects of the congruent and the partially intersecting alignment can be simulated with both of the hardware types by only using a small part of the FoV and the FoP. In the congruent alignment parts of the camera image and parts of the projection have to be ignored and in the partially intersecting alignment only the part where the FoV and FoP overlap are used. This would result in a loss of resolution and size of FoV and FoP in both cases. But since cameras already provide a high enough resolution for optical trackers and projection units are expected to increase in resolution, these workarounds should be kept in mind for future setups.

4.1.4 Other Alignments and Design Issues

Figure 4.4: Steerable mobile projection - Interaction with a steerable mobile projector.

Besides the described alignments no fixed orientations between FoV and FoP are possible. However the design of a projector phone should not be limited to fixed setups. With the FB-04V NTT Docomo presented a projector phone where the projection unit can be removed from the device and is controllable via Bluetooth. Even though such a setup seems promising, it is not suitable for vision-based AR applications since the orientation of camera and projector is unknown. Therefore such a system would require a possibility to determine the position of the components relative to each other.

Another possibility is the usage of a mobile steerable projection unit, which was done by Cauchard et al. [42]. A motorized mirror which could be controlled by a mobile phone, was attached to a mobile projector (compare Figure 4.4). The prototype has also been discussed in Chapter 2.1. This setup allows to imitate all alignments mentioned before, which would make it the most versatile.

Besides this classification, other issues should be taken into account when designing applications for mobile camera projector systems. The related work section provides an overview on the latest research done in this field.

Not only the spatial configuration of the mobile device camera and the projector play a role when discussing the potential and limitations of mobile camera-projector units. Other limitations, such as the physical nature of objects and the projection onto the objects cluttered appearance are still not discussed or investigated. Am I allowed to project on a stranger passing by? Many technical challenges still remain and have to be solved. Effects of hand tremor may be overcome utilizing accelerometers and gyroscopes. Moreover, camera-tracking methods have to be improved. All these factors currently have a big impact on the user experience and have to be taken into account when designing applications for the mass market of projector phones.

4.2 Prototypes

In this section we show the potential of projector phone interfaces for AR applications. On the basis of fully implemented prototypes, covering different alignments of camera and projector and employing different interaction techniques, we present how projector phones can cover a wide range of applications. These reach from mobile recommender systems to mobile games. The first two prototypes - Map Torchlight and Shelf Torchlight - represent the class of a *Partially Intersecting Alignment.* With LittleProjectedPlanet we investigate a congruent alignment. All three of these follow the type of *Hand-Held Projectors.* The last prototype of an AR system that we present here - guitAR - has not only another alignment but also represents an *Object Projector.*

4.2.1 Map Torchlight

4.2.1.1 Idea

The advantages of paper-based maps have been utilized in the field of mobile AR applications in the last few years. Traditional paper-based maps provide high-resolution, large-scale information with zero power consumption. There are numerous implementations of magic lens interfaces that combine high-resolution paper maps with dynamic handheld displays [127, 179]. From a perception perspective, the main challenge of magic lens interfaces is that users have to switch their attention between the magic lens and the information in the background. With the Map Torchlight application we attempt to overcome this problem by augmenting the paper map directly with additional information. The Map Torchlight is an example for a partially intersecting projection that is tracked over a paper map and can precisely highlight points of interest, streets, and areas to give directions.

4.2.1.2 Interaction Techniques

Figure 4.5: Map Torchlight - The hardware prototype used for Map Torchlight (left) and the augmentation of a paper map using Map Torchlight (right).

The general advances of a mobile projection system are also present in our Map Torchlight system: The projection area is larger compared to a standard phone display and simultaneously the mobile projection can overcome the switching cost of magic lens interfaces. The basic interaction pattern is still similar to magic lens interfaces. Sweeping the projector phone over the paper map, the projector will highlight, for instance, POIs on the map. Therefore, this prototype facilitates an interaction technique from the *Movement of the projector* class. Since the projection is significantly larger than the device display (around 8 times in our setup) more dynamic information can be directly presented on the map (as can be seen in Figure 4.5). As shown in Figure 4.5, larger objects can be highlighted compared to a traditional magic lens interfaces. Additionally, the map can be used as a shared projection screen by multiple users. For instance, one user can communicate a route to another user through the city by moving a projected crosshair over the map. The waypoints are stored in a Keyhole Markup Language file (KML is a is an XML-based language schema for expressing geographic annotation and visualization) and transferred via Bluetooth to the second user's mobile device. All AR overlays are contained in such KML files, as it makes it easier to generate and exchange them. A downside of the projection is that the real-world appearance of the map cannot be completely overlayed, as it is possible with (video see-through) magic lens interfaces.

4.2.1.3 Implementation

The Map Torchlight is fully implemented for Nokia mobile camera phones (S60 3rd edition). We use the tracking toolkit by Rohs et al. [178] to track the mobile device with the attached projector in real time relative to the map (6 DoF). The tracking algorithm processes about 22 frames per second. Our actual prototype is based on a Nokia N95 mobile phone coupled with an AIPTEK V10 Mobile Projector (640x480 pixel) attached to the phone using a standard AV cable. The whole setup weighs about 360 grams. Due to technical limitations, the mobile phone's screen can only be mirrored and not be extended on the projector. Therefore, the projector always shows the mobile screen content. As the focus of the projector can only be adjusted manually, a first calibration step is needed before interacting with the map.

4.2.2 Shelf Torchlight

4.2.2.1 Idea

The search for a certain book in a library, which contains many books can be a time-consuming task. Even if one finds the right shelf, one still has to browse an often huge area in the shelf. The same problem occurs when searching for a specific product in a supermarket shelf that fits ones personal preferences (e.g. an allergic or diet profile). With Shelf Torchlight we present a prototype that aims to overcome the problems when searching for a book or a product in a shelf, using a projector phone. Furthermore, Shelf Torchlight can also act as a mobile recommender system, taking the personal profile of the user into account.

4.2.2.2 Interaction Techniques

The basic interaction concept we apply is similar to the torchlight metaphor that was used as well in Map Torchlight. By sweeping the projector phone over the shelf additional information is projected onto the objects and next to them. By integrating proximity awareness into the system, we extend this interaction technique with a semantic zoom and with that making the proximity towards the object one of the most important factors. Modjeska describes a semantic zoom in contrast to a physical zoom as follows: "A physical zoom, on the one hand, changes the size and visible detail of objects. A semantic zoom, in the other hand, changes the type and meaning of information displayed by the object." [147]. In our case a physical movement closer to, or away from the object changes the kind of information that gets projected. The closer the user is to the intimate region of the object, the more detailed the information becomes. To illustrate the function of the semantic zoom we picked two scenarios for our applications. on the one hand, the search for a specific book in a library on the other hand the search for a product that matches the user's needs.

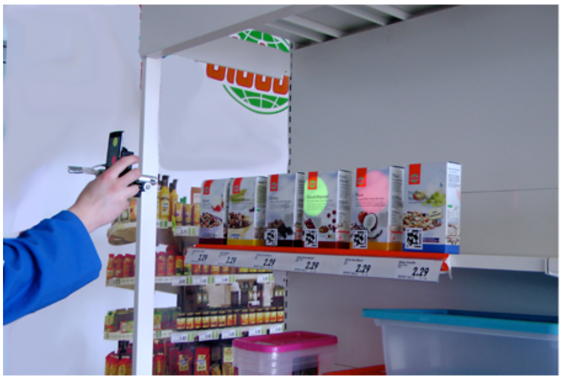


Figure 4.6: Shelf Torchlight in a Retail Scenario - Projecting dots onto products indicating how suited the product is for the user (left) and the semantic zoom revealing detailed information (right).

In the library scenario the system knows which book the user is looking for and thereby supporting the navigation task at the shelf. When one moves the projector phone over the shelf the desired books are getting highlighted with a white rectangle that matches the spine of the book. If the user draws closer to the shelf and with that activating the semantic zoom, he gets additional information like the average user rating retrieved from amazon.com, which gets projected onto the spine. If the user goes one step closer, the complete reviews for the book get projected.

In the retail scenario the products get compared to the personal profile of the user, that contains all his allergies, gusto, the shopping list, etc.. Not only the profile of the user but also of his whole family when he does the family shopping. Standing farer away from the shelf and moving the projector phone over the products, Shelf Torchlight projects green, yellow or red circles indicating how suited the product is (see Figure 4.6). For example a product that contains an ingredient that leads to an allergic reaction by the user or one of his family members, a red circle is projected onto the packaging. The semantic zoom will then reveal an explanation why the product got categorized in this way. In this example it will tell the user that the product contains a specific ingredient such as nuts. Since allergies are private information the semantic zoom shows this information only when the user is close to the shelf and maybe able to shield the projection. While the projection of the red circle only indicates that the user should not buy this product, uninvolved customers can not draw conclusions what reason leads to this advice, since it could also be a personal preference.

4.2.2.3 Implementation

The hardware of the prototype is based on the Map Torchlight prototype and with that a partially intersecting alignment of camera and projector. To track and identify the products and books, we use computer vision based methods. In a first attempt we tried to use the feature tracking of the metaio Unifeye Mobile SDK. After the first tests we experienced that the SDK is unsuitable for more than 3 different markers since the memory of the mobile device that was used was to small to process the image data. Therefore, we used the Visual Codes by Rohs [177] instead. In our examples the codes contain the ISBN of the books respectively, the EAN of the products, and were positioned on the spine of the books, respectively the facing of the product. The projection is aligned on the books and products in such a way that it does not interfere with the markers.

4.2.3 LittleProjectedPlanet

4.2.3.1 Idea

With the LittleProjectedPlanet prototype we explore the possibilities of projector phones in a gaming scenario, which was inspired by the Playstation 3 game LittleBig-Planet. The projector phone is used to augment the hand drawings of a user with an overlay, displaying physical interaction of virtual objects with the real world. Therefore, a congruent alignment setup is needed. Players can sketch a 2D world on a sheet of paper or use an existing physical configuration of objects and then simulate physical procedures in this world to achieve game goals.

We propose a mobile game combining hand drawn sketches of a user in combination with objects following a physics engine to achieve game goals. Enriching sketches in combination with physical simulation was presented by Davis et al. [57]. The ASSIST system, was a sketch recognition system that allows e.g. an engineer to sketch a mechanical system as he would on paper, and then allows him to interact with the design as a mechanical system, for example by seeing a simulation of his drawing. Interestingly the creators of the game LittleBigPlanet bought the ASSIST system and parts of it were integrated into the game play.

In contrast to the ASSIST system, we present a game that is designed for mobile projector phones, combing real world objects and projected ones, but utilizing a physics engine. We think that such a projector phone can also be utilized to improve the learning and collaboration in small groups of pupils. Because of the mobile setup of our prototype it provides a higher degree of freedom in contrast to a more teacher-centered system on an interactive white board (as shown by Davis et al. [57]).

4.2.3.2 Game Concept and Interaction Techniques

The slogan of the popular game LittleBigPlanet is "play with everything" and that can be taken literally. The player controls a little character that can run, jump and manipulate objects in several ways. A large diversity of pre-build objects is in the game to interact with, and each modification on such an item let them act in a manner physically similar to those they represent. The goal of each level is to bring the character

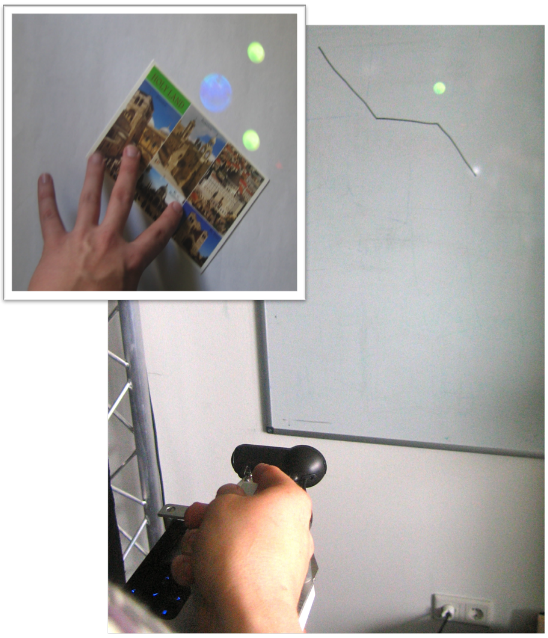


Figure 4.7: LittleProjectedPlanet Gameplay - A user playing the game with a postcard (upper left corner). User is sketching a marble run and projected tennis balls are bouncing on it (center).

from a starting point to the finish. Therefore, it has to overcome several barriers by triggering physical actions. But the main fascination and potential of the game is the feasibility to customize and create levels. Creating new objects is done by starting with a number of basic shapes, such as circles, stars and squares, modifying them and then placing them in the level. Having done so, the user can decide on how these objects should be connected mechanically.

We took this designing approach as an entry point for a mobile AR game. It allows the user to design a 2D world in reality, which is then detected by a camera. Out of this detection, a physical model is being calculated. In this model, the user can place several virtual objects representing items like tennis or bowling balls. These virtual objects then get projected into the real world by the projection unit. When starting the physics engine, the application simulates the interaction of the virtual and the real world objects. Afterwards the results of the virtual objects are projected onto the real world surface.

Just like in LittleBigPlanet our application offers the user different ways of playing: One is like the level designer in LittleBigPlanet; the user can freely manipulate the 2D World within the projected area and place virtual objects in it. Similar to children building tracks for marbles in a sandpit, the player can specify a route and then let the virtual marbles run along it. A different gaming mode is a level based modus, but instead of steering a character as in LittleBigPlanet, the user designs the world. As a goal the user has to steer a virtual object e.g. a tennis ball from its starting point to a given finish. The game concept uses a direct manipulation approach. Enabling the player to modify the world at runtime, lets the real world objects become the user's tangible interface. But not only the objects are used for the interface, by changing the orientation and position of the projector the user can also modify the physical procedures (e.g. gravity by turning the mobile camera projector unit around).

For designing this 2D world the players can use several methods. Basically they have to generate enough contrast so that the 2D world can be detected by a standard edge recognition algorithm (utilizing the Sobel operator). Sketches on a piece of paper or a white board could for example be used for this, but simply every corner or edge of a real world object could generate a useful representation in the physics engine. There is no need for an extra tracking of a sketching device such as an IR LED equipped pen. Just requiring the projector phone itself the game is playable nearly anywhere with nearly everything and it is easy to set up. Figure 4.7 shows how a user is projecting virtual marbles on a track he sketched on a whiteboard. An important problem to allow a smooth and seamless interaction for the user is that the gravity in the projection is aligned with the real worlds gravity.

This prototype represents the category of *Bi-Manual manipulation of Surface and Projector*, as the user can not only control the game by changing the projections surface content, but also by changing position and orientation of the projector.

4.2.3.3 Implementation

Due to the unavailability of sophisticated projector phones (with an optimal alignment of camera and built-in projector and e.g. a CPU that is able to process the physics simulation) we used a Dell M109S, a mobile projector with a maximum resolution of

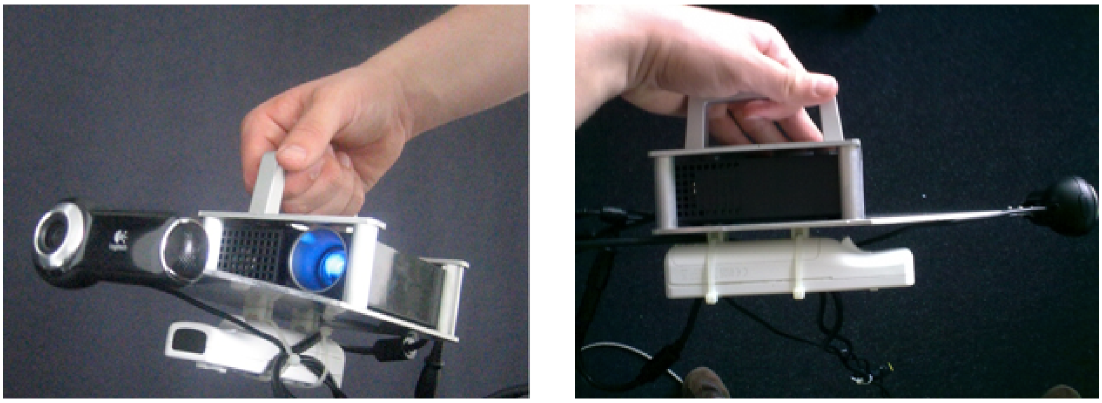


Figure 4.8: LittleProjectedPlanet Hardware Prototype - The hardware prototype consists of a Logitech Quickcam 9000 Pro, a Dell M109s and a WiiMote all mounted on a aluminium construction.

800 by 600 pixels and a weight of 360g. We attached a Logitech QuickCam 9000 Pro webcam to it. All together our prototype weighs around 500g and is therefore okay to handle (compared to the prototype used in Map Torchlight our prototype is 240g heavier, but the projector has 50 lumen instead of just 10 and also has a higher resolution). We think this prototype provides a good trade-off between mobility and sophisticated projection quality. In contrast to the few mobile devices with built-in projectors, our projector and camera are mounted in such a way that the camera FoV fits the projected area (congruent alignment). But because of the different focal lengths of camera and projector in this setup, the camera image is always wider than the projected image (which can be seen in Figure 4.8). For controlling the application and to determine the orientation (to set the gravity) a Nintendo Wii remote is attached to the prototype. Today's mobile phones are already equipped with an accelerometer or an electronic compass, so the functionality of the Wii remote can easily be covered using a future projector phone. The application is fully implemented in Java using the QuickTime API to obtain a camera image. As a physics engine Phys2D, an open source Java based engine, is used. WiiRemoteJ handles the communication with the Wii remote. Connected to a standard laptop or PC the camera projector unit has a refresh rate of approximately 25fps when running the application.

The area of the camera image containing the projected image is processed via an edge recognition algorithm. Every pixel of a detected edge gets a representation as a fixed block in the physics engine. That gives the user total freedom in designing the world. The update of the world in the physics engine is done every 300ms but the user, for example, can pause this update for editing. Adapting the gravity of the physical model to the actual orientation of the projector phone is done through calculating the roll (this denotes the angular deviation along the longest axis of the Wii remote) of the Wii remote.

4.2.4 GuitAR

4.2.4.1 Background and Idea

Even though the guitar is one of the most popular musical instruments among autodidacts it is still hard and for many people frustrating to learn. Traditionally, musical teaching is a one-to-one situation where a student performs and the teacher gives feedback or demonstrates how to play. For autodidacts this one-to-one situation is often changed to a one-to-many situation in which the student performs while the instructions are coming from many different resources. Whilst traditional resources were books with play-along CDs, nowadays, resources span from free video lessons to online communities where novices can upload videos of their playing efforts and get valuable feedback and advices. But still many people leave the instrument solitary in the corner caused by the absence of success.

For ages people used different visualization aids to ease guitar learning. To carve the notes into the fretboard is one prominent example, which can be found on many old guitars (compare Figure 4.9). The most popular and easiest accessible information source of learning materials for autodidacts today is the Internet. Due to the increasing popularity of video-sharing portals like YouTube, guitar video lessons are becoming widespread. Besides the lack of interaction and feedback, the major disadvantage of video clips is that the fingering in the video is inverted, based on the frontal view on the instructor. Guitar teachers report that the different point of view is a big problem for almost all novices. Furthermore the student has to switch his focus from the screen to the fretboard of the guitar frequently.

In this subsection we want to present guitAR, a system for projector phones that can overcome the problems that autodidactic guitar students have to face. With a projector phone mounted onto the headstock of a guitar, it is possible to project instructions directly onto the appropriate position of the fretboard. The projected information includes fingering and phrasing instructions for chords and melody sequences. We present two different approaches either applying an *Object Projector* or a *Moveable Projector*.

Figure 4.9: Guitar out of the bequest of Franz Schubert - Notes are carved into the fretboard. (Today in possession of the Haus der Musik, Vienna)

4.2.4.2 Music Learning Interfaces

Many HCI approaches for learning a musical instrument exist. Especially for the piano a wide variety of commercially available, products as well as research projects exist. Piano learning interfaces range from keyboards with keys that can light up to indicate what should be played (for example manufactured by Yamaha as well as Casio) to the Moog PianoBar, which is an LED bar that can be attached to any standard piano. Yamahas Disklavier takes this one step further and actuates the keys of the piano that needs to be played as well. The possibility to actuate the keys was picked up by the MusicPath project, which allows piano teaching from remote locations through the connection of two Disklaviers [30]. Both, the teacher and the student can see what the other is playing through the actuated keys and with this, it also allows to communicate the strength that is used in the keystroke. All the above-mentioned interfaces have one drawback and that is the lack of information about the hand gestures. It is not obvious which finger is used to play which note. This problem was addressed by Xiao and Ishii with MirrorFugue [226], which allows visualizing hand gestures of a remote collaborative piano player.

Prior HCI approaches to alleviate the learning of the guitar mainly focused on using AR displays overlaying a camera image with the instructions on how to play a distinct chord or which notes to play next [36, 152]. These approaches, which are based on optical markers, have the same disadvantages as video lessons: the student sees the instructions in an inverted view and has to switch his view permanently between the display and the guitar. Besides this, the student has to manage to keep the optical marker, which is attached to the guitar, inside the video image. This retrenches the student further, since he is not free to move the guitar. Even though these markers can be replaced by a markerless tracking - since guitar-necks normally provide a rich amount of features that could be tracked - the area in which the student can move the guitar around would still be limited. The approach presented lets the students move their guitar freely and the instructions that are given are presented directly on the fretboard of the guitar in such a way that the student‘s focus of attention can stay on the guitar the whole time.

The usage of stereo cameras to track the fingers of students was presented by Kerdvibulvech [109]. Burns et al. created a system that tracks the fingering with just a normal Webcam using a circular Hough transformation [33]. Both approaches where able to determine the position and check if a chord was played right but not able to give any instructions.

There are commercially available guitars that are especially made for guitar novices as well, such as the Yamaha EZ EG [97], which is a MIDI guitar without real strings, where a button, which can be lit up, replaces the note on the guitar neck. By using colour patterns, the students can learn chords and songs. However, the Yamaha EZ

EG has the disadvantage that it does not provide the flexibility and the feeling of a real guitar. Once the student has learned to master this instrument he has to start again getting used to real strings. Besides the Yamaha EZ EG, there is Fretlight [95], a fretboard with an integrated LED for each note that can be controlled via a computer. Fretlight has several disadvantages. It is not applicable on a standard guitar, the guitar needs to be connected to computer, and the content for Fretlight is not freely available. In contrast to that, the concept presented here can be used with every guitar without changing or damaging the instrument.

Besides the correct playing of notes learning a musical instrument requires also to gain continuous expressivity on the instrument. Johnson et al. used other output modalities to ease the learning of a music instrument [100]. Their prototype MusicJacket was able to give vibro-tactile feedback to the arms to indicate to a novice player how to correctly hold the violin and how to bow in a straight manner. A similar approach for fine-tuning of the body expression was presented by Ng with i-maestro [157]. By using a motion capturing system they created a 3D augmented mirror that gives interactive multimodal feedback on the playing and body pose of the student. A drawback of these two approaches is the need for a huge amount of instrumentalization of the player. Furthermore, an adaption of these approaches for guitar players would hardly be feasible, since sensors and actuators would need to be attached to the students fingers and this most certainly would negatively influence the students playing.

4.2.4.3 Concept

Our approach to ease guitar learning is to project information on fingering of chords or songs directly onto the fretboard of the guitar using a mobile projector. We propose two different possible AR approaches. The first approach utilizes a projector phone mounted at the headstock of the guitar (see Figure 4.10 right). The second one is based on a tablet computer with integrated projector (see Figure 4.10 left). The first approach would allow the student to move freely around and would only need the mount with projector phones becoming ubiquitous. This approach represents a projector from the type of *Object Projectors*. The tablet approach on the other side represents a *Moveable Projector*. However to realize this type of projection for guitAR first of all a technique that is able to track the guitar neck with high precision has to

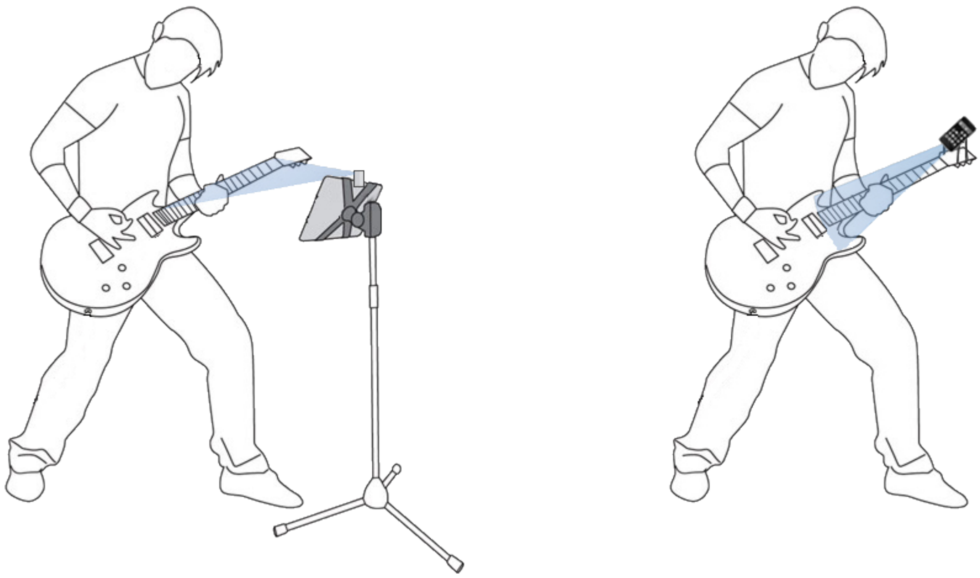


Figure 4.10: GuitAR Concept - Using a projector phone mounted on the headstock (right) or a tablet computer with integrated projector (left)

be established. This again would limit the radius of movement (which is what existing approaches suffer from [36, 152]) to make sure that the guitar's neck is in the field of projection, as well as in the camera image if a computer vision based approach is used.

Regardless if either a tablet or phone is used, the most important factor for such a system is the visualization. The guitar allows more versatile phrasing of a note than a piano, especially complex playing techniques like string bendings or Hammer-On‘s that are often used, need to be distinctively but easily recognizable visualized.

To indicate which finger the student should use to fret a certain note often numbers are used in todays chord diagrams. The correct fingering is essential for fast and clean guitar playing and therefore a factor that should be taken into account when designing a guitar learning application. Using the wrong finger for a certain chord can lead to slower playing or even worse, when a chord variation should be played it may not be possible since the finger that should play the variation is blocked. The use of numbers is difficult to realize in the projection since the space on the fretboard is limited and it may be cumbersome to read. Thus, we choose to use different colours to indicate what finger the student should use. In general, detailed symbols or characters are hardly recognizable on the fretboard. Therefore, we propose to use only basic shapes like circles or squares that are distinguishable when being projected on the fretboard.

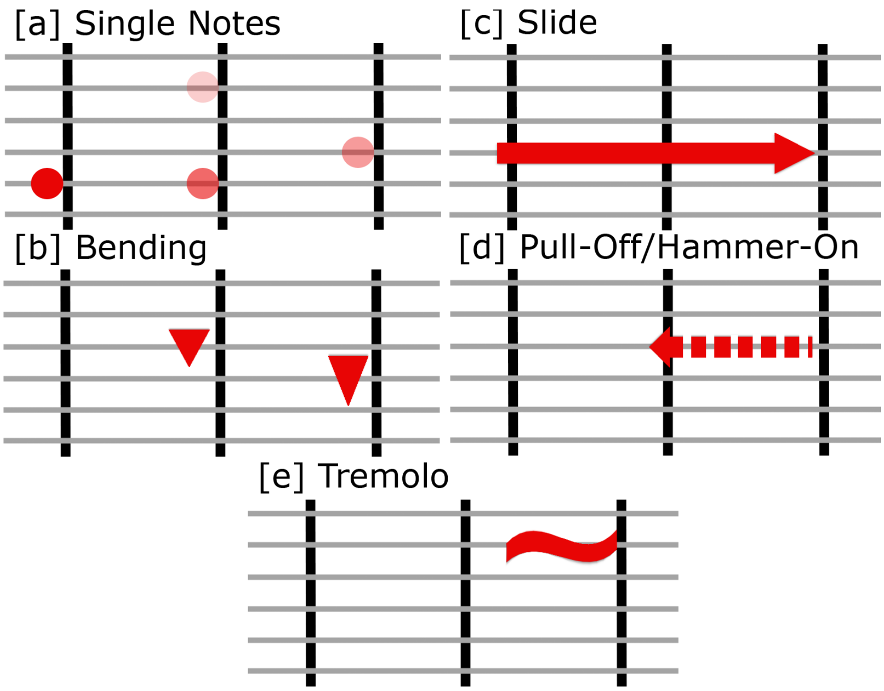


Figure 4.11: GuitAR Visualizations - Visualizations for different playing techniques

The following visualizations (Figure 4.11) have been developed in collaboration with a guitar teacher and a more advanced guitar student:

- When a single note has to be played, a coloured dot is projected onto the fret. If more than one note has to be played at the same time different colours indicate which finger the student should use to play which note. Such information is indeed contained in chord diagrams, but usually not in guitar tabs for a whole song. For a blank string a white dot is projected on the nut of the guitar. Since musicians typically are more concerned about which notes or chords they have to play next than what they are playing at the moment, upcoming notes have to be visualized as well. We choose to fade out the colour to visualize this, so that the next notes are the brightest and the following are fading out slowly (compare Figure 4.11[a]). The student can adjust how many notes are shown and how far from the current beat they should be, to adapt the technique to his learning performance.

- To visualize that a string has to be bend, a triangle is projected on the fret and the size indicates the pitch to which the string has to be bended. The direction of the head of the triangle indicates if an up- or down-bending should be performed (Figure 4.11 [b]).

- A slide from one note to another - which is performed on only one string - is indicated with an arrow on the fretboard. The origin of the arrow indicates the note on which the sliding starts and the arrowhead indicates the destination note to which the student has to slide to with his finger (Figure 4.11[c]).

- To indicate a Hammer-On or Pull-Off, a dotted arrow is projected again with the arrowhead indicating the destination note (Figure 4.11 [d]). The dotted arrow reflects the movement which the player performs compared to the complete arrow of a slide. When performing a Hammer-On or Pull-Off, the notes between the start and the end are skipped, which is reflected by the gaps in the arrow.

- A finger Tremolo on a specific note is visualized through a curled line (Figure 4.11[e]). Whereat the curliness of the line can indicate how articulated the tremolo needs to be played.

The advantage of the described visualizations is their unambiguousness. They are well distinguishable even with a distorted projection. Their form originates from the movements of the fingers of the guitar player and the visualizations that are used in today's guitar tabs. With that they are easy to learn and to recognize, even for people who are not familiar with the system.

One problem of video lessons is that when a student wants to figure out a certain part, he has to repeat this part in the video over and over again. Therefore he has to take his hands from the guitar away to control the video. The Yamaha EZ-EG and Fretlight suffer from the same problem. Therefore we propose to integrate different input techniques to control the projection without the need to remove the hands from the guitar. Three different modalities would be feasible - speech, gesture and sound input. Speech input is a reliable technology in todays mobile phone but it always requires pressing one button to trigger the recording. To control the projection easy and short commands like rewind or play would be enough but a continuous recording and processing of audio data would be needed. Another possible input technology would be to detect if the student plays a predefined sequence of notes. But this again would require continuous recording and processing. The third feasible technique would be gesture recognition. Most phones today contain sensors like accelerometers or gyroscopes, and with them gestural detection is easily possible. When a projector phone is attached to the headstock of the guitar, swings with the guitar neck could be interpreted as gestures. This approach seems to be the most promising since the computational overhead is comparably small and movements normally are not that fast so that an acceleration threshold can be used to distinguish the normal small movements from an intentional gesture.

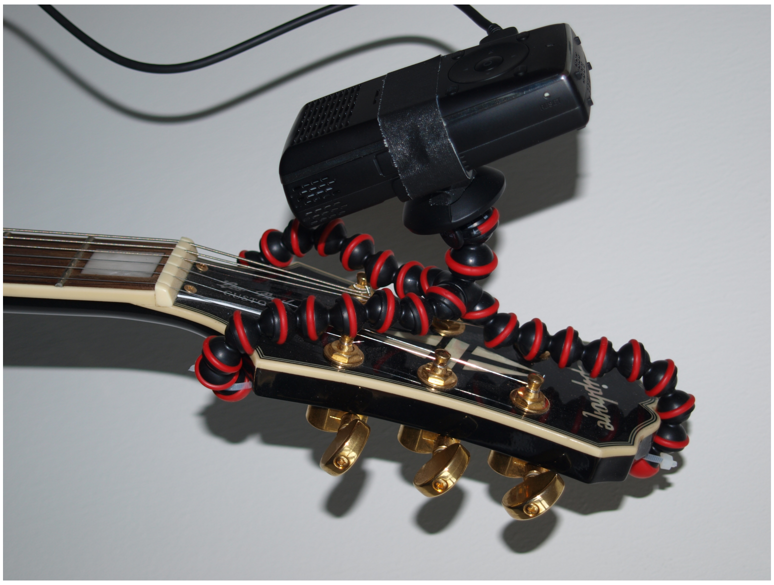


Figure 4.12: GuitAR Prototype - The prototype consists of a Gorillapod and an AAXA L1.v2 Laser Projector. On the back of the headstock (not visible in the image) a Phidget Accelerometer was attached.

4.2.4.4 Implementation

We created a prototype of the described concept using an AAXA L1 laser projector that is mounted on the headstock of an Epiphone SG guitar (see Figure 4.12). The mount consists of a Joby Gorillapod that was fixed to the neck using cable straps. On top of the Gorillapod the laser projector was mounted. With this mount, all guitar tuners are accessible and normally functionally while at the same time the projector is easily adjustable. The projector weighs 122gramms and including the mount, the whole prototype weighs 210gramms. Mounted to the headstock there is no adverse effect on the playability of the guitar even though it adds a little bit more weight.

To control the projection we implemented a Qt application running on an Apple MacBook. The application is capable of projecting 25 different chords and also able to read tabs for complete songs in ASCII format and project the notes onto the fretboard. When projecting a complete song the tempo in which the notes are shown can be adjusted individually to the learning speed of the student. Unfortunately ASCII tabs contain no information about the correct fingering, therefore we will use another standard in the next iteration. The alignment of the projected image to the fretboard was done manually. In future implementations we would aim for automatic vision based recognition of the fretboard and automatic alignment.

For the gestural input we attached a Phidget accelerometer to the headstock of the guitar. For the recognition of the gestures, we used the One Dollar Unistroke Recognizer by Wobbrock et al. [225]. The recording of the gesture starts when the acceleration of the headstock reaches a certain threshold. From the three dimensional data that the accelerometer provides, only two dimensions are used since a movement along the axis of the guitar neck is not feasible when playing the guitar seated. The two remaining axes are mapped to the One Dollar Unistroke Recognizer x- and y-axis.

4.2.4.5 User Evaluation

The visualizations for different notes and playing techniques were demonstrated to two advanced guitar players and one guitar novice. They rated the visualization as straight forward and easy to learn. First tests with the prototype showed that the different shapes and colours are easily distinguishable and referable to the strings they should belong to (compare Figure 4.13). Also the mount was proven to be robust and stable enough to keep the projection aligned with the fretboard even when the guitar was heavily moved. The only thing that the testers stated was that they were not able to determine which note to play when the projection was blocked through their hand, which happened when the note lies behind the hand. With the Yamaha EZ EG and

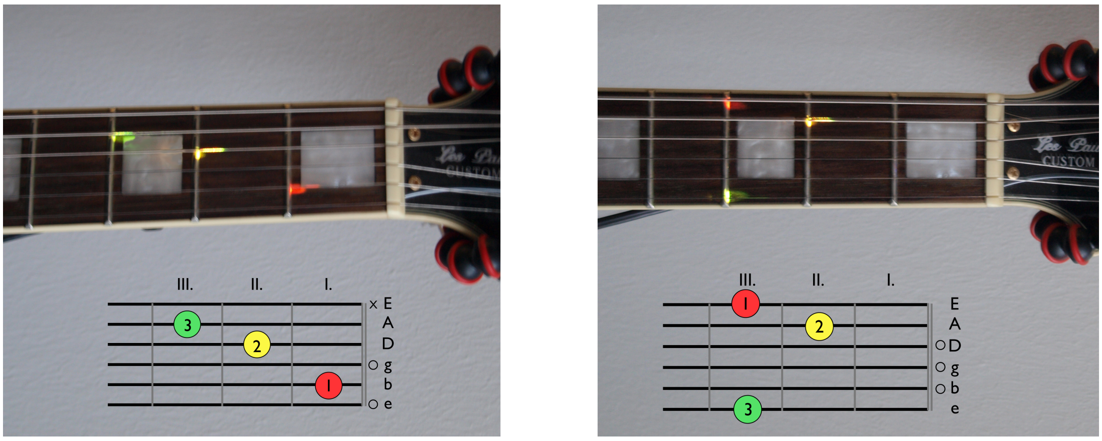


Figure 4.13: GuitAR Projection - The prototype projecting a chord, C-major (left) and G-major (right). Beneath are the corresponding standard chord diagrams.

Fretlight the problem is analogue because the fingers cover the light that shows the position of the note on the fretboard. When using a *Moveable Projector* such as a tablet with integrated projector, projected from the front onto the fretboard, the light would only be blocked if the angle between the tablet and the guitar was precipitous. Otherwise, the instructions would simply be projected onto the fingers and the student could estimate the exact position. But nevertheless, all users rated this to be a minor problem.

4.3 Summary

To answer the question **How have future projector phones to be designed to enable spatially aware applications** we presented in this chapter different interface designs for mobile projection-based AR applications taking hardware issues of future projector phones into account. All the interfaces focus on the augmentation of real world objects in the environment and with that, the most sophisticated form of spatially aware applications. We showed how the different spatial setups of camera and projector units effect the possible applications and the physical interaction space. Our presented prototypes highlight these challenges. The presented classification can help to structure the design space of mobile projection applications. Of course many open issues still remain. As discussed earlier not only the spatial configuration of the mobile device camera and the projector play a role when discussing the potential and limitations of mobile camera-projector units. Today, hardware issues still hinder the exploitation of the full impact of mobile camera-projector units. Our research tries to make a contribution into the direction assuming that we will have better hardware of mobile-camera projector units, we will have more powerful applications that go beyond projecting only content like images or videos. The implementations show how researchers can overcome the current hardware problems and investigate the area of mobile camera-projector systems more deeply. With our categorization using different classes based on the spatial configuration we want to establish a framework for AR applications using projector phones. We think that they have a big potential to enrich the usability of mobile devices. They enable larger presentation sizes and are well suited for multi-user settings. With all this we demonstrate that AR applications using mobile projected displays are not only already possible, but also that they can provide a level of spatial integration that on current device is not yet possible. Therefore we argue that making this integration of projected displays with the real world should be a major consideration in the design of future mobile projection devices.

5

Display Alignments in Mobile Multi-Display Environments

While the effects of heterogeneous multiple displays in classic desktop settings have been explored extensively, this is not the case in mobile settings. To support the user when interacting with mobile projected displays the spatial alignment of these need to be analysed to find the most suited ones. In this chapter we compare different alignments of mobile projected displays and smaller physical displays. Using mobile eye-tracking we evaluate these alignments in terms of visual separation effects. Afterwards we present an application case of the best suited alignment with the SurfacePhone. This novel configuration of a projector phone, aligns the projector to project onto a physical surface to allow tabletop-like interaction in a mobile setup. The projection is created behind the upright standing phone and is touch and gesture-enabled. Multiple projections can be merged to create shared spaces for multi-user collaboration.

This chapter is based on the following previously published publications: [43, 220].

5.1 Visual Separation in MMDE

Clamshell phones, handheld dual-display game consoles, projector-enhanced tablet PCs and cameras are steadily increasing the number of multi-display mobile devices. In the case of projector phones,the larger projected display can be viewed by more than one person at a time unlike traditional handsets.

Such mobile multi-display environments (MMDE) operate by providing visual information on different screens. They create larger screen estate, which can be used for partitioning information. Often the larger display allows sharing public information, while private information can be kept on the devices screen for the owners eyes only, or to principally support input feedback. While MMDEs can be used both in single-user contexts (using multiple displays to partition tasks) and multi-user contexts (e.g. for privacy setting), the design space of MMDEs needs to be defined for single-user contexts before design considerations can be applied to multi-user environments.

The existing MDE literature shows how such device ecologies are affected by concerns of visual separation [196, 197]. Visual separation is the division of information across space in MDEs. The fact that the information is not continuous can create difficulties for the user when handling information and interacting across display spaces.

Research in fixed multi-display environments has shown that visual separation of content can affect performance [22, 136]. Tan and Czerwinski [197] found a significant detrimental effect when dividing information across multiple displays at different depths for the same separation angle. Likewise, Su and Bailey [196] found that when positioning large displays through workspaces, the relative depth between displays can affect users performance.

While it could be argued that users direct control of the projection space and the closeness of the phone display could reduce the effects of visual separation on mobile MDEs, it could equally be argued that the mobility of displays could accentuate static MDE problems. Moreover, mobile projector phones have an inherent depth differential between the phones screen and the projection. Prior work in MDEs would suggest negative visual separation effects due to this depth gap. With a lack of understanding of how visual separation affects usability and performance, it is hard to identify appropriate designs, suitable interaction techniques or adapt these devices to specific applications.

The principal contributions of this section are that:

1. Visual separation does not affect the viability of MMDEs

2. Displays must be positioned in the same field of view (Figure 5.1)

3. Mobility factors do not exacerbate visual separation. 4. We present some implications of our study on MMDE design

We now will review existing investigations of the field of MMDE as well as MDE to lay the foundations needed for this section.

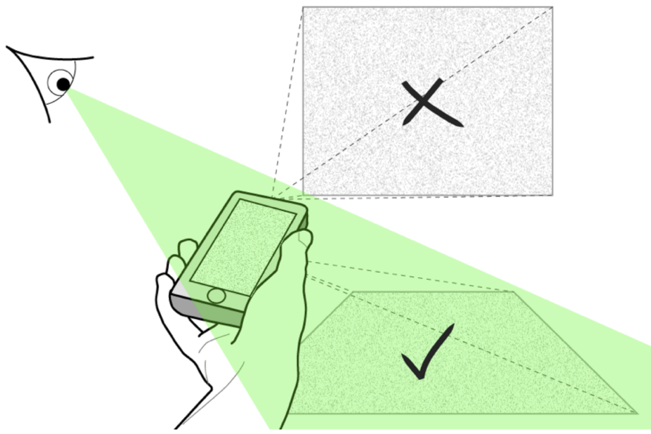


Figure 5.1: MMDE Design Consideration - In a mobile multi-display environment, displays need to be positioned in the same field of view.

5.1.1 Types of Mobile Multi-Display Environments

In this section we explore existing work in MMDEs. MMDEs are either partially mobile (i.e. a mobile component imported in a traditional MDE) or fully mobile (i.e. a mobile device that supports more than one display). Partially mobile MDEs include environments where a mobile device is imported inside a traditional MDE, for example Greenberg et al. [75] present an environment in which a PDA is used in conjunction with shared public displays. Since the mobile component can be flexibly reoriented relative to the existing MDE, the visual separation effects of this component in the overall environment may be mitigated by the ability to easily reorient the device. This could minimise visual separation between displays and thus current research in fixed MDEs is likely to hold in the partially mobile MDE case.

There are many existing examples of fully mobile MDEs in the literature, which can be divided into two categories, multi-device-single-display and single-device-multi-display. Mobile multi-device-single-display environments are created when individual single-display mobile devices are brought together to create one display. For example Lyons et al. [134] present techniques using a network to link multiple single displays in order to share co-located display spaces. Finally, Siftables [144] provide a set of tangible interactive objects, each equipped with a single display that can be combined in order to manipulate data and information. They support tangible interaction effects, such as removing a physical item from a pile to delete associated virtual data. In all the described cases, each individual display can easily be moved and re-oriented depending on the desired situation. The users can then intuitively reduce visual separation effects.

Single-device-MMDEs provide more than one display on a single mobile device. This type of environment has gained a lot of popularity with the growth of embedded pico-projectors in existing devices such as phones, cameras, camcorders and even tablet PCs. Traditionally these displays have been fixed relative to one another, such as with a mobile projector phone where the projection lens is normally fixed at an orthogonal angle to the mobile phones screen. Some devices, such as the Nintendo DS present re-configurable hardware capabilities in between the two screens. Unfortunately, these capabilities are not currently exploited by software applications. Nonetheless, increasing numbers of single-device-MMDEs exploit a reconfigurable multi-display layout, as the Codex [89] where two screens are hinged and can be rearranged into different positions. Despite the possibility of re-orienting these devices, many single-device-multi-display environments do not allow the user to rearrange displays in order to simultaneously visualise information. This bears the question of whether we can immediately transfer guidelines from research on fixed multi-display contexts to mobile single-device-multi-display environments.

5.1.2 Factors of Influence of Visual Separation

Having reviewed existing MMDEs, in this section we review existing work on visual separation in MDEs and the visual separation challenges for single-device-MMDEs. Factors amplifying the effects of visual separation have been studied for a range of multi-display configurations including when displays are of different sizes, when placed at different distances from the user, if oriented at different relative angles and when separated by surrounding bezels or frames.

5.1.2.1 Size and Depth

Mandryk et al. [136] show that users are faster at interacting between two identical and continuous monitors compared to using a secondary monitor of smaller size placed with a small gap to the primary screen. Pointer warping techniques such as Mouse Ether [13] and frame memory pointer [20] propose cursor movement techniques that can help reduce the effects of visual separation across displays of different sizes in heterogeneous MDEs.

Early literature in ergonomics [3] advises that documents and screen are kept at the same distance from the user for data-entry tasks that require rapid shifts between both elements, to reduce costs in switching views. Recently, Tan and Czerwinski [197] show a detrimental effect due to visual separation when a screen and a projector are placed at different depths within the same visual field. These negative effects can be reduced with techniques such as the Perspective Cursor [155], that remaps the ordinary mouse cursor in a complex heterogeneous MDE depending on the perspective of each user regardless of their position.

In most single-device-MMDEs, the screens used are set to have similar characteristics, dimensions and are often at the same distance from the user (i.e. where the device is held). However, in projector-enhanced mobile devices, screens and projections vary in size and distance depending on the proximity to the projection surface. Although absolute size and distance can be configured by manipulating the device, relative size and distance between displays are typically fixed and may cause visual separation effects due to angular or focal displacement.

5.1.2.2 Angular separation and Field of view

Tan and Czerwinski [197] show greater visual separation effects of depth when the data is separated by a 55 angle (i.e. outside the useful FOV) compared to a 27 angle (i.e. inside the useful FOV). Su and Bailey [196] studied visual separation for multiple large displays and found negative effects when the secondary screen is situated on the same horizontal plane as the primary screen but at an angle of 70 relative to the user, at the periphery of their field of view. Their study also showed a negative effect when the second screen was completely behind the user (i.e. in a completely separate FOV); however, they found no effect when the secondary screen was oriented at an angle from the first screen and were both at the same distance from the user. Following their experiment, they presented a set of guidelines on how to position two large displays relative to each other: the displays should stay on the same horizontal plane, at no more than a 45 subtended visual angle and should not be placed behind a user; in other words both displays should stay within the users FOV.

Some single-device-multi-display environments are designed with the displays in different fields of view. For example, some clamshell phones are equipped with both an internal and an external display, such as the Samsung Alias 2. With this configuration, the screens are on different sides of the phone (i.e. in a different FOV) and cannot be used simultaneously. Codex [89] is a dual-screen device that works with a hinge between the screens and offers different functionalities for different rotational postures of the screens, that can be in same or different FOV depending on context.

Z-Agon [139] is another example of single-device-multi-display with 6 screens fitted in a cubic arrangement. Held in the palm, it can be moved to explore content on the 2 or 3 faces in front of the user while other faces remain hidden at the back of the cube.

5.1.2.3 Bezels

In MDEs, Tan and Czerwinski [197] found no effects of visual separation due to bezels and physical distance between screens alone. Yang et al. [227] found minimal visual separation effects between Lens-Mouse (a mouse with screen on top) and the monitor. Task performance in Yang et al.s study [227] degraded in their dual-monitor condition attributed to distance and not bezels. Contrarily, Bi et al. [22] found that splitting symbols across two displays with a bezel in the middle was detrimental in a search task. Bi et al. [22] also found that interacting with data was faster with no bezel compared to a tiled screen. Forlines et al. [71] show that for an individual user; having information split across multiple vertical screens is detrimental in terms of reaction time to accomplish a visual search task compared to a single vertical screen. Stitching [90] is an interaction technique designed to reduce visual separation effects by using a pen interface to draw interaction lines across multiple displays.

Chen et al. [46] present a dual-display e-book reader and shows advantages of using multiple screens for reading. For example, information can be separated on both screens through the bezel for multi-document reading. Moreover, the device supports interaction techniques that draw on real books, such as moving one screen towards the other to turn pages. In addition, the screens can be detached and reassembled for different modes of use. Devices with dual screens separated by a bezel already exist, such as phones, laptops or even game consoles as the dual-screen Nintendo DSi or dual-touch-screen Toshiba Libretto laptop.

5.1.2.4 Mobility

In all the above designs, MMDEs have very different characteristics to traditional MDEs. We have identified inherent size and depth gaps which create potential angular and focal separation in the case of projector-enhanced mobile devices or individual displays placed in separate fields of view. Previous research in MDEs shows that multiple screens need to be placed within the same useful FOV of the user to avoid negative effects of visual separation [196] and also that specific interaction techniques need to be applied if the size of the displays differs. Yet, MMDE designs do not necessarily follow these guidelines because the studies presume a fixed position and orientation and no or limited control over changing display placement during the task. It is therefore essential to determine whether visual separation effects previously demonstrated in fixed MDEs translate to MMDEs. Therefore in the following we explore the design space for MMDEs and we determine if the negative effects of visual separation in MMDEs can be reduced by aligning displays within the same field of view.

5.1.3 Design Facotrs for MMDEs

There is a fixed number of ways to position displays together in a single-device-MMDE. When such devices possess more than two displays, these design considerations apply to each pair of displays individually. The displays can either be separated by: distance vertically (Figure 5.2 a,b,c) or horizontally (Figure 5.2 d,e,f), an angle (Figure 5.2 g,h,i),

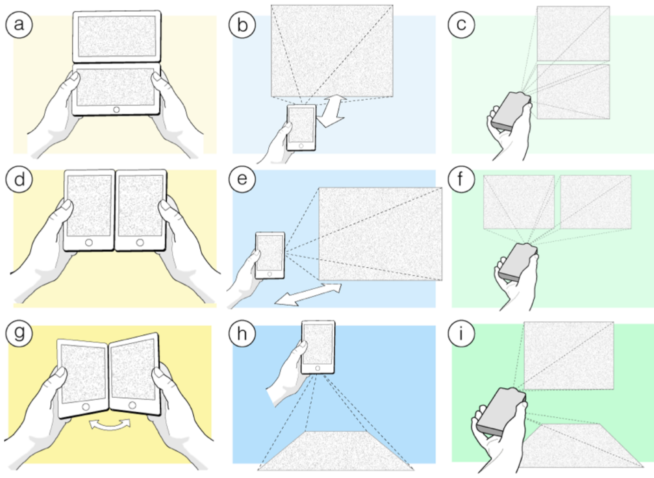


Figure 5.2: Possible layouts for two displays on a mobile device for different types of displays - screen-screen (a,d,g); screen-projector (b,e,h) and projector-projector (c,f,i). Displays can be separated horizontally, (a,b,c), vertically (d,e,f), by an angle across any plane (g,h,i) or any combination of the above positions. In the screen-projector cases (b,e,h), the displays are further separated by depth due to the inherent properties of each display

or any combination of those conditions. The displays can be separated by any distance δ that will vary depending on the devices design: from a few centimeters wide such as the size of a bezel or a hinge (Figure 5.2 left column) up to a few meters wide in the case of a projector enhanced device (Figure 5.2 middle and right columns). When the displays are separated by an angle α (Figure 5.2 g,h,i), α can be of any value (0-360) along any axis in the cartesian space.

When the displays are close to each other or at a small angle from each other, they are in the same field of view. However when δ or α have high values, the displays are in different fields of view. In fixed MDEs, displays tend to be in the same field of view, which is not the case in current MMDEs. In our user study, we will determine whether placing the displays in different fields of view increase visual separation effects.

Depending on the design of the device, the displays are either relatively fixed: always at the same distance and angle from each other or reconfigurable: the distance and angle between the displays is context-dependent such as in Codex [89].

In Figure 5.2, we present possible layouts of two displays: two screens (left column), a screen and a projector (middle column) and two projectors (right column). In the two screens case, the screens are unlikely to be more than a few centimetres apart in order for the device to be handheld; the design is therefore similar to traditional screens in MDEs separated by a bezel. The visual separation effects are then likely to be similar to the effects of bezels in MDEs. However, bezels do not affect visual separation as long as information is not cut across the bezel [22] and appropriate interaction techniques are implemented [90]. We have therefore chosen not to explore visual separation effects for this configuration.

In the case of a screen and a projector, the displays have by default heterogeneous characteristics, such as different sizes and resolutions and are moreover separated by depth. The literature on MDEs shows that depth can be an important factor when managing visual separation effects. Moreover, the position of the projector lens on the device itself will determine if both displays will be in the same field of view or not. We believe that visual separation effects will be at their strongest in this type of environment, hence our decision to run the user study with a projector enhanced mobile device.

The two projectors case is similar in characteristics to traditional large displays MDEs, such as two projection spaces that will display either on the same, on an orthogonal or on opposite planes, characteristics that have already been explored in the MDE literature. Yet, dual-projectors mobile devices present some interesting features such as the ability to display at different depths depending on the surrounding environment, as when displaying on an uneven wall. Nonetheless, in most multi-projector cases, the projections will either be separated in distance (depth), in plane or in size of projection. We believe that those issues are similar to the ones encountered by a screen and a projector case and that any experimental results obtained for the former configuration will apply to this category too.

5.1.4 User Study

The purpose of this study is to identify the effects of visual separation on single-device-MMDE when the multiple displays are in the same field of view and when they are not, as well as when the device is fixed or mobile. We run the study using a projector-enhanced mobile device since the embedded displays are by design of different sizes and displaying at different depths. We decided to study the combination of a screen and a projector instead of the two-projector case since this configuration is more prominent in current devices. We expect that the lack of physical connection between displays will generate greater effects of visual separation.

Our experimental setup includes the following aspects of mobility: handheld (participant can hold the device as they feel comfortable), portability (implies that the size and distance of the projection will vary), unsteadiness (jitter is not compensated for) but not actually moving between rooms in order to allow comparison between results in the fixed and the mobile settings.

5.1.4.1 Task

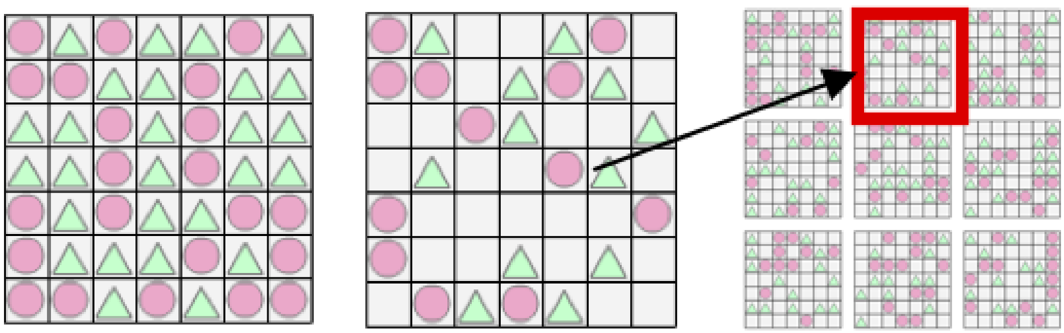


Figure 5.3: User Study Task - a. Left: Example of pattern displayed on the screen b. Centre: Matching sparse version c. Right: 3x3 grid displayed on the projection

The task chosen for this experiment is a visual search task. Visual search is a typical task for analysing visual separation [71]. Tan et al. [197] use different types of task including text comparison as it is representative of tasks in which the user must cross reference and compare content displayed in multiple locations. Chen and Chien [45] use a similar task when looking at effects on visual performance on small screens. In our experiment, we chose an image comparison over a text comparison task, since the laser projectors resolution could affect reading accuracy.

The task chosen consisted of matching a pattern on the screen (Figure 5.3 a) with a sparse version of the same pattern (Figure 5.3 b) positioned inside a projected 3x3 grid of competing matches (Figure 5.3 c). This makes use of the different display sizes, showing the initial pattern only and a keypad on the small display and the 9-pattern grid on the larger projected display. The sparse versions are randomly created by deleting half of the items from the initial pattern and replacing them with blank cases. The competing patterns in the grid are other sparse versions of the initial pattern for which 5 items are permuted in order to look similar but not match the initial pattern.

The participant would select a matching pattern on the projection by pressing the corresponding number on the numeric keypad on the screen below the initial pattern. Depending on the answer, the participant could receive positive audio-feedback and continue to the next trial or receive negative audio-feedback and would have to repeat the same trial until the correct matching pattern was found.

5.1.4.2 Experimental Design

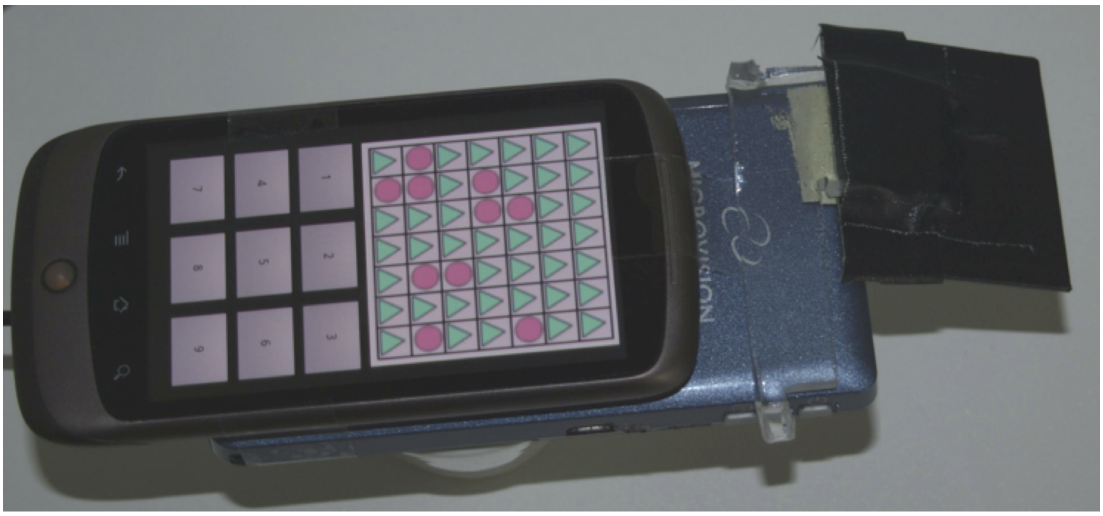


Figure 5.4: User Study Hardware - Phone and projector used for the user study fitted with the Floor setting mirror

For the study we used a Google Nexus One with a touch screen combined to a Microvision ShowWX laser pico-projector (Figure 5.4). Additionally the study makes use of a portable eye tracker to analyse the switching behaviour between the two displays.

The experimental room was darkened to optimize the projector viewing conditions. The independent variables were:

- Position of the projection relative to the phones screen: in the same FOV (Floor), in different FOV separated by one angular plane (Front) or by two angular planes (Side)
- Mobility: whether the device is fixed on a tripod or handheld by the user: mobile setting.

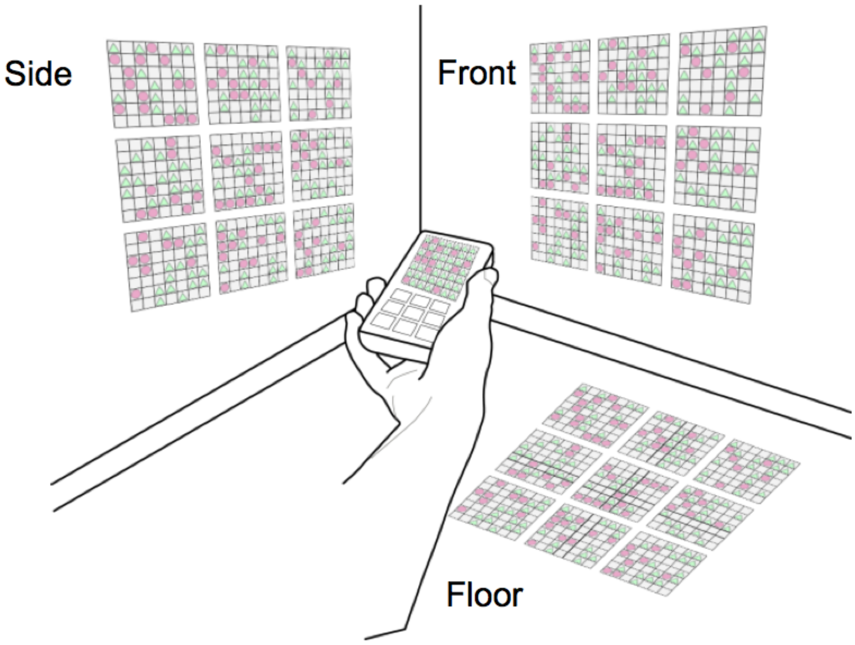


Figure 5.5: User Study Task Pattern - Example of task pattern and grid of sparse patterns in the three positions in fixed setting: Front, Floor and Side. For each trial, the grid is only displayed in one setting only.

Position The projection spaces relative to the screen are described on Figure 5.5: The Front projection corresponds to the alignment of the phone and the projector. A mirror is placed at the top of the projector lens and oriented at a 60 ° angle downwards for the Floor condition, as shown in Figure 5.3, and a 40 ° angle sideways for the Side condition. In order to reduce the keystone effect introduced by the mirror, we projected at a resolution smaller than the projectors maximum one.

Mobility In terms of mobility, the device was either set at a fixed position on a tripod (fixed setting) or held by the user (mobile setting). In the fixed setting, the assembly phone-projector is placed on top of a tripod and the participant had to stand on predefined markers by the tripod. The position of each projection space (on the floor, front and side walls) was predefined in order to set a constant position and aspect ratio of the projection for all participants. We ensured that all three projection spaces were the same distance from the device (110cm) and would therefore always have the same size (middle of the projected grid fixed at 60cm wide).

In the mobile setting, the user is holding the device and can use any projection surface at any distance or size that they are comfortable with. The user was free to move around the room with the device. The distance to the wall and the size of the projection would then vary depending on users movements. We did not impose any restriction on how the user would hold the device. Nonetheless, we observed that most users held the device in the non-dominant hand and touched the screen with the dominant hand, while other users held the device in both hands and used their thumbs to touch the screen. None held the device with their dominant hand.

5.1.4.3 Eye tracking procedure

The context switches were measured using a mobile eye tracker: Tobii Glasses that recorded eye movements at 30 Hz. This eye tracker is non-intrusive as it is low weight (75 grams glasses) and fully mobile so participants could roam freely. Some IR markers were positioned around the various display spaces (in the fixed setting) to allow automatic data mapping and help repositioning the projected image at the same place for each participant. The eye tracker records both a video of the scene and where the user is looking in the scene.

5.1.4.4 Hypothesis

Based on the literature review and our preliminary exploration of the issues, we expected display configurations (relative positions of displays within same or in different FOV) and whether the device is being held (mobility) to significantly affect performance and produce visual separation effects.

We presumed visual separation effects to be less important when the screen and the projection are in the same FOV (floor setting) than when the projection is in a different FOV than the screen (front and side settings). We also expected for participants to compensate visual separation effects when holding the device since they could themselves reconfigure the display areas adaptively.

5.1.4.5 Procedure

Twelve volunteers (5 men) aged between 24 and 35 years old (avg. age 28.6 years) were recruited from within one of our universities. All our participants were familiar with touch-screen technology and all had normal colour vision. We used a within-subjects design where position and mobility were counterbalanced across participants.

We explained the task to each participant individually. To start a trial the user pressed the Start button whenever they felt ready. There were 8 trials for each experimental condition. Participants were also told that they should say aloud if they pressed the wrong button in order to identify false negatives. After the experiment, users filled out a NASA TLX satisfaction survey.

In summary the experimental design was: 12 participants x 2 mobility factors x 3 positions x 8 trials = 576 data points.

5.1.4.6 Measures

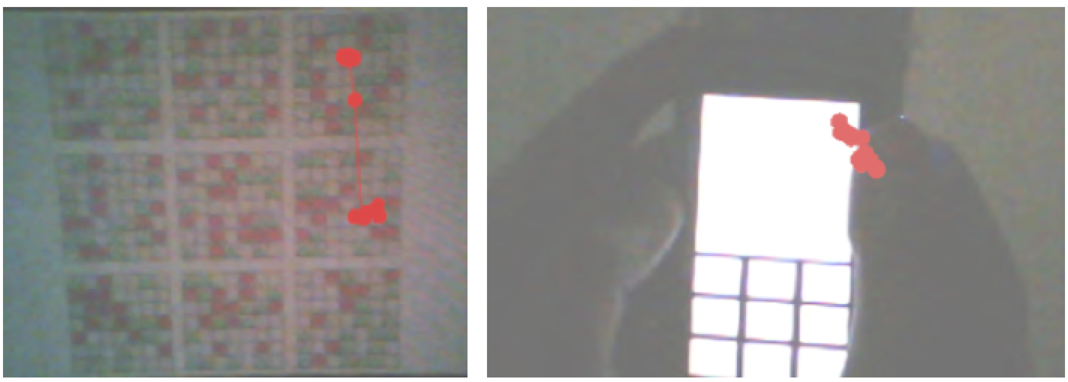


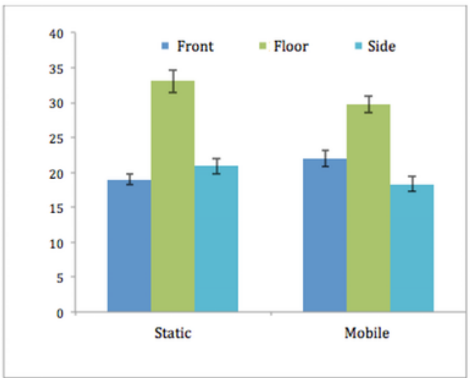
Figure 5.6: User Study Eye Tracking Data - Snapshots from the eye tracker video Left: User is looking at the projection. Right: User is looking at the phone

- Number of context switches between the screen and the projected display. Previous studies on visual separation do not measure the number of context switches. However the problems induced by context switches are quantified in mobile projector-phone studies [82] as well as in some MDE studies [19, 60]. This is measured by the portable eye tracker. The number of context switches is computed by the eye tracking software in the fixed setting using the IR markers and is then manually verified through analysis of the eye tracker video. In the mobile setting, the switches are manually counted at the video analysis stage (Figure 5.6) since the position of the projection space is not constrained in this setting.

- Completion time and number of errors in performing each trial, including number of false positives. These are typical measures in visual separation studies [71, 196, 197] and allow comparing participants efficiency for different experimental settings. The completion time is timed between the start of the task to its successful completion.

- Position preferred - NASA TLX: This test assesses subjective information on a 7-point scale for mental, physical and temporal demand; performance; effort and frustration. We have combined this traditional subjective workload questionnaire with some personalised questions aimed at gathering user preference data.

5.1.4.7 Results



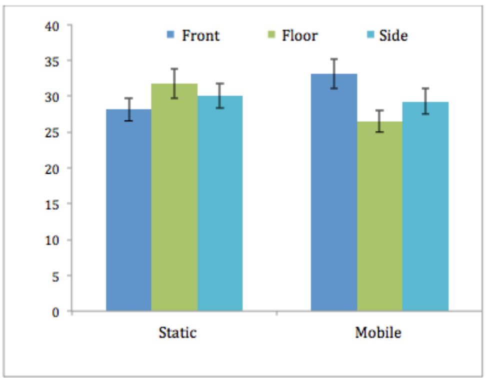


Figure 5.7: Context Switches and Task Completion Time - Average of context switches (left) and average task completion time (right) for each conditions: Front, Floor, Side in the two mobility settings.

We used a repeated measures ANOVA test for the number of context switches, completion time and number of errors. We used the univariate ANOVA test for the NASA-TLX results analyses with subject as a random factor.

- *Number of context switches*: We found a main effect for position ($F(2,94)=62.817$, $p<0.001$), pairwise post-hoc comparison showed significant differences between the positions: Front and Floor ($p<0.001$), and Side and Floor ($p<0.001$) and no significant differences between Mobile and Static conditions ($F(1,95)=1.034$, $p>0.05$). The mean for Front and Side were respectively 20.49 and 19.62 context switches, compare to 31.41 for the Floor conditions as shown on Figure 5.7(left).
- *Task completion time and number of errors*: Our findings showed no significant difference in trial completion time for position ($F(2,94)=0.390$, $p>0.05$) and mobility ($F(1,95)=0.057$, $p>0.05$), as well as no significant difference in error-rates for the different positions ($F(2,94)=1.049$, $p>0.05$) and mobility ($F(1,95)=1.143$, $p>0.05$). The average error rate across all conditions was 8.9%. Figure 5.7 right shows the average trial completion times across all conditions.
- *NASA TLX*: We only found a significant difference in temporal demand for position ($F(2,22)=4.086$, $p<0.05$). Floor is perceived as faster than Front and both are perceived as faster than Side (means for temporal demand for Floor is 3.67, Front is 3.83 and Side is 4.33 on a 7-likert scale). For all other variables no significant effect was found.
- *Position preferred*: In the fixed setting, 75% of participants chose the Front position with the remaining participants preferring the Floor. In the mobile setting, half of the participants preferred the Floor, 42% the Front and 8% the Side. When asked what their favourite condition was overall, 75% favoured a mobile position compared to a fixed one (Figure 5.8).

5.1.5 Discussion

At this point, we discuss the above results on four related themes: Viability of mobile MDEs, Dual-display configurations, Substantiation of mobile uses and Design implications

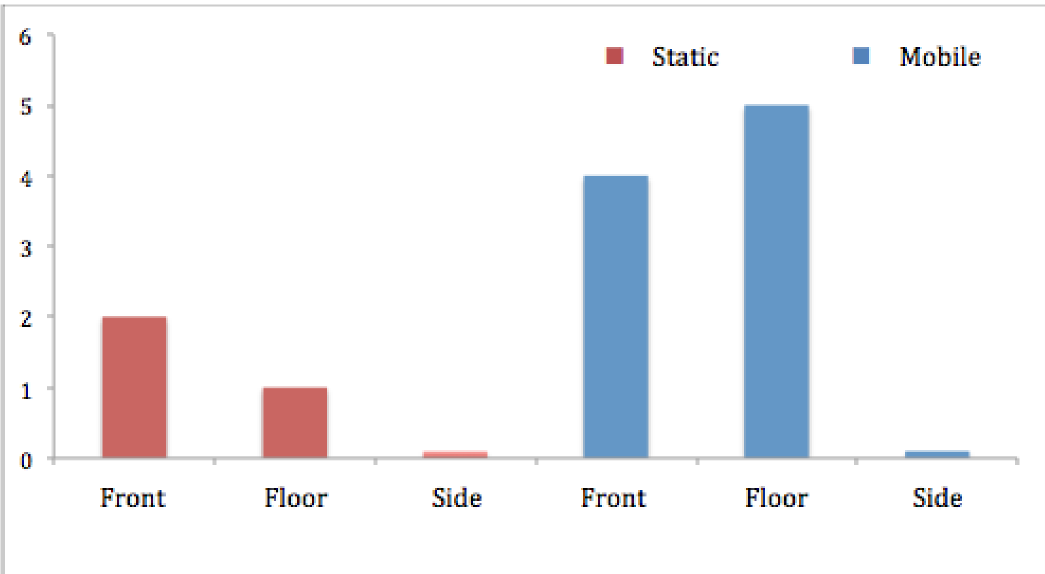


Figure 5.8: User Preference of Position - Overall preferred position for each participant

Visual separation does not affect the viability of MMDEs The results of the study show that visual separation effects did not prevent users from carrying out the task, which is reflected through the low error rate of only 8.9% over all tasks and conditions. This result is valid for both the static and the mobile conditions.

During the experiment, participants had no problems using a mobile dual-display device even with very heterogeneous displays in terms of size, resolution and depth of the displays. In a case where tasks are divided across displays, single-device-MMDEs can outperform todays single display devices. Moreover, the tasks can make use of the different displays characteristics, such as our experiment uses the phones screen to display a single pattern and a keypad and the large projection space to display a large grid of 9 patterns.

We conclude that MMDEs with heterogeneous displays are a viable solution.

Displays must be positioned in the same field of view The eye tracker recorded significantly more eye context switches in the floor condition: over 30% more than in the other positions; whereas we found no significance in completion time and error rate across the three positions in both fixed and mobile settings. This important result would have been overlooked should we have used traditional task performance measures only.

The results show that the number of eye context switches does not affect task performance and that there is a higher number of context switches when both displays are in the same field of view. This suggests context switches are a lot cheaper to perform when both displays are in the same field of view (Floor setting) as they only require a simple eye movement and little or no head and neck movements, unlike the Side and Front conditions where participants reported discomfort. One participant said about the Side setting: "It was very uncomfortable to constantly turn my head during the experiment". We also believe that the higher number of context switches in the Floor condition is due to the fact that context switches can be considered as epistemic [112], using the active memory to store the position of the geometric shapes in the pattern. Instead of having to remember the positions in the pattern, users could externalise their thought processes by switching context more often. This is also the reason why this setting appeared as being faster paced to the participants. We recommend that multiple displays in MMDEs should have the displays aligned in the same field of view.

Mobility factors do not exacerbate visual separation Since we found no significant difference between mobile and static setting in terms of error rate, task completion time or context switches, we believe that the following mobility factors: handheld, portability and unsteadiness, have no effect on visual separation. Participants wrist and hand movements in the mobile setting did not help compensate the effects of visual separation. A possible reason could be that they were already compensating for the jitter of the projection resulting from the participants holding the device in their hands. Since none of the participants mentioned jitter as a problem during the experiment and in the post-study questionnaire, we conclude that they instinctively compensated for any mobility-induced jitter effects. Our experiment showed no more visual separation effects between mobile and static settings, even though the projection space and size were varying; and since participants showed a strong preference for the mobile setting, the investigation of mobile scenarios is justified. Since scenarios involving the aforementioned characteristics of mobility are justified, we recommend the investigation of other mobility factors.

5.1.6 Design implications

In the following sections, we present some design implications for future MDEs that emerge from our discussion in terms of type of displays, display physical arrangements, flexibility of design and mobility.

Type of displays for MMDEs Our experiment showed that MMDEs are viable, which includes heterogeneous dual-display solutions. Although dual-display solutions for mobile devices are technically possible, they are currently under-exploited by manufacturers. Our study demonstrates that these solutions should be envisaged more often since visual separation effects do not present issues for carrying out activities where tasks are distributed across displays, such as in our experiment. This is also valid for activities wherein the user chooses which display to use depending on application and context needs. Those scenarios of use are consistent with most common uses of MDEs as described by Grudin [78].

Additionally, most existing dual-display mobile devices are designed with multiple displays of similar types, whereas heterogeneous displays offer more potential, such as the ability to choose where to display depending on the context without generating negative visual separation effects. While current usage of heterogeneous dual-display mobile devices is often limited to one display at a time, we encourage designers to consider exploiting both displays simultaneously. This would also allow more flexibility in the choice of interaction technique.

Physical arrangements of displays Our experiment shows that having both displays in the same field of view is paramount for applications that make use of both displays. Users can reduce the amount of information they have to remember and can instead use active memory to recall information by switching gaze between displays more frequently. This is particularly important for applications that suffer from heavily cluttered displays, such as map applications. This pattern of increased context switches to alleviate cognitive load is equally important when one display is also used to facilitate input to the other display. This especially facilitate input techniques that require *Input on the projection device*. In this situation, the displays must be arranged within the same FOV.

However, arranging displays in the same FOV is not trivial in a mobile environment where external factors influence how the user holds the device and on which surfaces content can be displayed. These external factors range from luminosity and glare to the available projection spaces, number of users viewing the content and the type of information being displayed. The usage of a steerable projection could overcome these environmental issues, as presented by Cauchard et al. [42], a steerable projection can reduce visual separation effects in MMDEs by automatically reconfiguring the alignment of the displays according to the context the device is used in.

Flexibility of design Prior research conducted in MDE suggests that displays arranged on different planes or separated by more than 45 angle result in lower task performance and provide negative visual separation effects [196]. However, in our study, we find no significant task performance differences, whether in time completion or error rate, across the different settings. These results show that guidelines for MDEs are not directly applicable to MMDEs. One explanation could be the use of a very small display close to the user. This shows that although it is preferable for the user to have both displays in the same field of view, there is more flexibility in the alignment of displays in MMDEs than in MDEs.

This is especially the case for applications that do not require epistemic actions from the user, and for which the need for rapid context switching is not crucial. For those applications, manufacturers have more freedom to position the projection unit wherever it best suits the device ecology. This could result in smaller devices since the projection unit could be placed where it fits best without generating visual separation effects on performance. In this case, a wide range of interaction techniques can be supported for which displays do not need to be aligned, such as foot interaction on the floor [42] or even shadows on the projection [53] for any other projection setting.

Mobility In our study we find no more visual separation effects when the device is held than when the device is fixed on a tripod. Most current single-device-MMDEs are built for scenarios of use in which the device is placed on a surface. Our study shows that all factors of mobility are worth investigating, such as when the user is walking while holding the device; or stopping by to obtain contextual information about the area they are walking by. Many contextual applications could benefit from true mobility and new interaction paradigms could be envisaged, such as the use of haptic while on the move.

5.2 Application Case - SurfacePhone

As we learnt form these design implications the combination of a mobile projected and a physical display offer a very flexible MMDE space. At this point we want to make use of this knowledge and present a prototype that incorporate these findings.

Current projector phones such as the Samsung Galaxy Beam [76] precludes many of the prevalent sharing and collaboration techniques that are well known and investigated for example in today's tabletop systems. Therefore we here present a setup of a physical and a projected display that takes into account the aforementioned findings and allows to recreate such tabletop-like interactions in mobile scenarios with a private and a public display. The so called SurfacePhone is a novel configuration of a MMDE that consists of a physical and a projected display. It is able to project a second display right behind itself, while it is standing on a surface.

The investigated setup allows for collaboration and sharing as well as advanced single-user interactions. The projected display is touch- and gesture-enabled and additionally spatially aware. This allows connecting multiple projected displays into one combined display. Here the design process of the SurfacePhone as well as our implementation are presented. Starting with the considerations for such a system and the envisioned usage concepts, we then present two prototypes supporting aforementioned affordances. The initial concept prototype allowed us to easily evaluate the previously discussed concepts and ideas. The technical mobile phone case prototype (hard- and software open sourced) was developed to show the technical feasibility of the SurfacePhone concept.

5.2.1 SurfacePhone Concept

The design of the SurfacePhone concept encompasses the position of the projected surface in relation to the phone, the position and orientation of one SurfacePhone to other SurfacePhones in the environment, and the modalities to interact with screen and projected display in either scenarios. Further, we distinguish between *single device/single user*(SDSU), *single device/multi user* (SDMU), and *multi device/multi user* (MDMU) scenarios.

5.2.1.1 Position and size of projection

Hinckley et al. [89] showed that a range of very private to very public and collaborative application scenarios can be supported, depending on the spatial relation of dual-screen postures. The projection in front of the mobile device would resemble the laptop display configuration, it is a very private setup. This is because the projection is mainly visible to the user facing the device. In such a configuration the projection could show a soft keyboard. A projection to either sides of the phone would imitate the setup of Bonfire [102]. In this semi-private setup, the projected surface is still within easy reach of the user, but more public than in the laptop scenario.

These two configurations have been explored intensively but a projection behind an upright standing phone has been neglected so far. This setup – the SurfacePhone – consists of a public projected display and a private display (as can be seen in Figure 5.9) and presents a more collaboration oriented setup. To some extent, it resembles the Battleship setup of Codex [89], albeit the difference that the primary user is able to see both the phone display and the projected display. In this setup, there is a clear separation between the private phone and the public projection that is visible and within reach to people in the near vicinity. This comes at the expense of a slightly more difficult interaction with the projection as the user has to circumvent the phone to touch the projection.

Additionally, this MMDE setup is in line with the findings above. When the user is sitting in front of the upright standing phone, the phone's display as well as the projection are in the same field of view. This allows the SurfacePhone to split the information between these two displays without risking visual separation effects.

To give an idea of the size and position of the projection early-on, we experimented with different alignments we found the optimal size and position of an (undistorted) projection behind the phone that we could achieve to be around 17cm $\times$ 14cm in size, 14cm behind the phone and 4cm to the left of the center of the device. The projection, thus, is three times as big as the 4" screen of the iPhone 5.

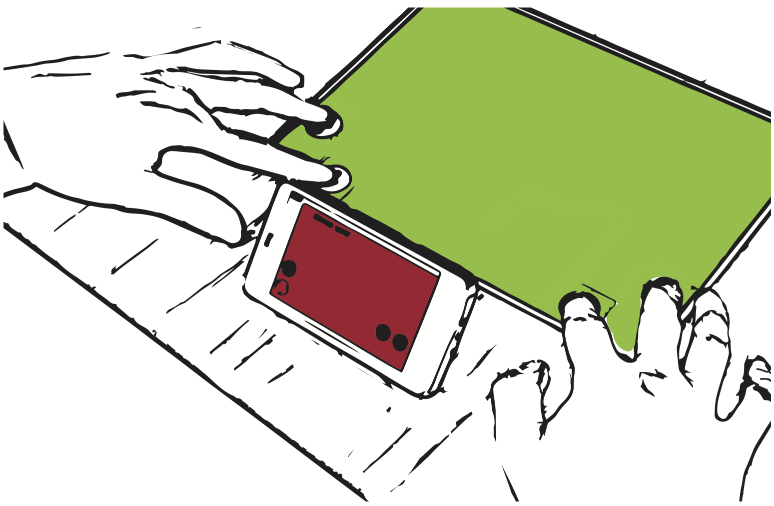


Figure 5.9: Single Device Single User (SDSU) - In this setup the screen space of the user is enhanced to e.g. solve the fat-finger problem using the projection behind the device.

5.2.1.2 Configurations

The SurfacePhone can be used alone, or by multiple users using one or multiple SurfacePhones. Specific to this setup is that the user is interacting with two displays simultaneously that have quite different affordances. The mobile display is brighter and more touch-responsive than the projected display. However, the projected display is bigger – at least in terms of size, not necessarily resolution – and supports *Direct Interaction on the projection* as well as *Around the device interaction.*

Single-device, single user (SDSU) This configuration can be used, for instance, to overcome the fat-finger problem on mobile devices by outsourcing e.g. controls of a game (compare Figure 5.9) to the projection or showing the main view of the game on the projection. Apart from that, the projected display could be used as a general secondary display, for instance, showing a task manager or notifications of applications currently running on the device. Furthermore, an "Ambient-aura" around the phone that is coupled to the user's perspective on the device could enrich the immersion of multimedia and games such as has been shown for TVs with IllumiRoom [101]. Finally, phone screens are very useful for augmenting the reality of the user, but cannot serve publicly visible augmentation (as discussed in detail in Chapter 4). The projection on the other hand could be used to augment a real playboard with projected tokens. For example it could project chess tokens on a real board to play against the computer or a human opponent. As the projected display is fully covered in the camera image, the SurfacePhone represents an congruent alignment.

Single-device, multi-user interaction (SDMU) Leveraging the inherent differences in publicity of the displays the SurfacePhone can be used for several sharing tasks in small groups (compare Figure 5.10). For instance, the projection of the phone can be used to present pictures or slides to a small group of people. The screen of the SurfacePhone can be used to browse the content and decide which content should be shown on the projection. Advantages of using SurfacePhone in this scenario include that users do not have to hand over their phone to other people; the content can be presented to all people simultaneously; only specific pictures or slides for presentation

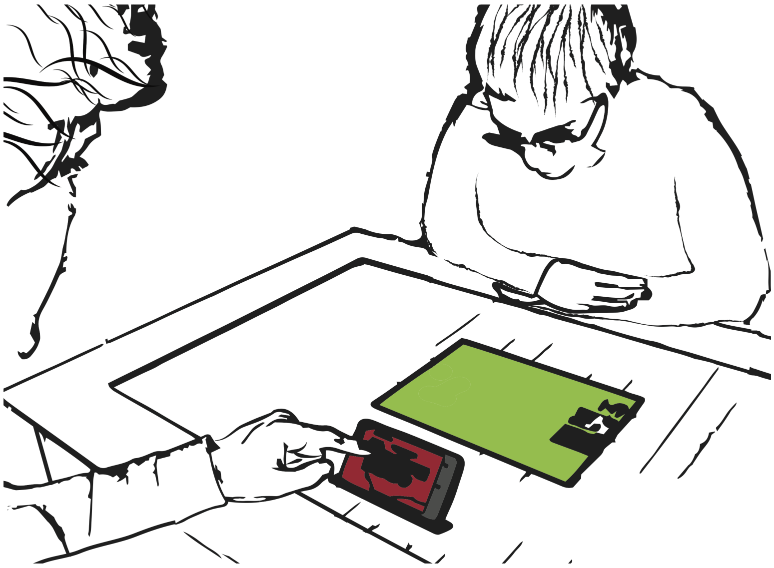


Figure 5.10: Single Device Multi User (SDMU) - In this setup the additional screen space is used to share information with small groups e.g. one user is presenting pictures to another user.

can be selected to address time or privacy constraints. Finally, the projection can also be touch- or gesture enabled, giving the viewers the possibility to interact with the pictures or slides. Similarly, the setup is also suitable for games such as blackjack: The person playing the bank controls the game from the screen. Other players sit in front of the projection and use *Direct Interaction on the projection.*

Multi-device, multi user interaction (MDMU) Finally, when more than one user brings their SurfacePhone to the table, projections can be merged at different sides forming larger shared surfaces. These can be used for collaboration, e.g. data sharing, as well as competitive scenarios such as gaming. Depending on the scenario and the familiarity of the participants, different setups support different degrees of collaboration.

Sitting next to each other on the *same side* (Figure 5.11 left) is the most intimate setup as both the projections as well as the phone screens are visible to both users. This setup, for example, could be useful to collaboratively search for holiday trips. Users

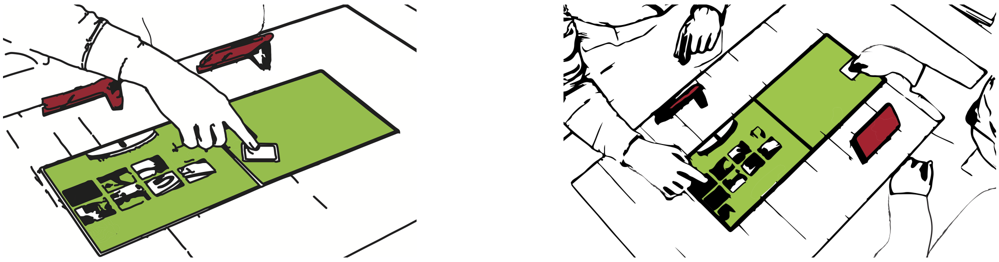


Figure 5.11: Multi Device Multi User (SDMU) - Short Edge Merge - In this setup the projections are merge on the short side. The users can either sit on the same side generating a very intimate setting (left), or sitting opposite to each other so that both projections are easily touchable (right).

can first explore offers on their personal devices, then share it to the surface. The setup also supports collaborative gaming, where the projection shows a shared game view and users control their characters or army on their mobile devices. Being able to also see other users' phone screens may significantly improve communication in collaborative planning.

On the opposite, sitting *face to face* (Figure 5.12) merging the long side of the projections is the most distant setup. It suits users unfamiliar with each other, as well as competing opponents in a game for instance. In both cases, users have private interaction on their mobile display, using it to selectively share content on the projected surface. Also, the own projected display is likely not within easy reach of other parties making it more personal for each user.

Sitting face to face, but at the same time *next to* each other (Figure 5.11 right) combines properties of both aforementioned setups. In this setup, users keep their private view on their mobile screens, but expose their projected surface to be easily reachable by the other party. Therefore, the setup particularly emphasizes familiar use cases of interactive surfaces, encouraging participants to manipulate all objects on the surface. Two users may also sit round the corner of the table which is in general equivalent to the previous case, but allows more easily to come round and take a look on the other user's private display when both users desire so.

Finally, groups with more than two devices merge projections at *arbitrary sides* in their center. Obviously, no general rule for the visibility of phone screens or reachability of projections can be determined. However, like people do when playing games involving hand cards, users can arrange to ensure the required visibility and privacy. Projections can be merged to a central surface for a few participants. Larger groups can also merge a ring of connected projections. Although these cannot visualize content in the middle, conceptually this area can still exist. For instance, in a card playing game, when making a trick, cards can move from one surface to another "through" the blank center, thereby incorporating the unprojected area into the interactive surface.

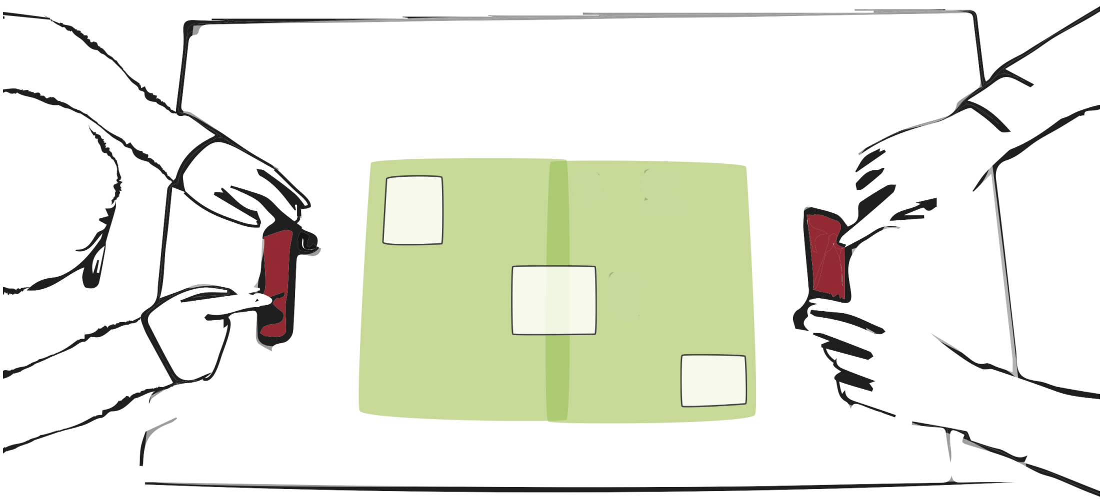


Figure 5.12: Multi Device Multi User (SDMU) - Long Edge Merge - Two users sitting face to face with the projections merged at the long side

5.2.1.3 Interaction Techniques

In the following we will discuss required interaction techniques for the SurfacePhone that suit aforementioned application and usage scenarios. Here we draw from users' experience and familiarity with smartphones and tabletop systems. The technical feasibility of the here described techniques will be addressed in the implementation section of the technical prototype. While in theory a projector phone could offer various additional hardware that could be used for sensing, constraints on size, battery time, etc. made manufacturers of projector phones settle to the current set of hardware which mimics that of ordinary smartphones besides the projector. Thus our exploration, to be realistic, should not involve additional hardware. Henceforth, the only difference between the SurfacePhone and other projector phones should be the different position and orientation of the projector. The setup could be advanced to a manually rotatable projector to fit standard wall projection scenarios as well.

With today's prevalence of multi-touch interaction users would expect to be able to interact with the projected content using *direct touch*. This includes long touches and double-touches, to allow for a richer input set through different touch modalities. Furthermore, gestures like directional swipes are common on tabletops and should be supported as well.Besides these *Direct Interaction on the projection* due to the phone cameras wide angle also other *Around the device interaction* techniques such mid-air gestures can also be considered. Furthermore with the front facing camera another opportunity for these *Around the device interaction* techniques exist.

Another interesting space of interaction lies *around* the projection. As the phone camera is seeing an up to ten times larger space around the projection, invisible buttons around the projection are possible. Similarly, gestures that cross the edges of the projection could be supported, for example, to move content to another user's projection that is currently not merged.

When many SurfacePhones merge so that they form a ring rather than a central space, not *all* other projections are inside the camera viewport. In this scenario the phone camera's flash LED could come to the aid. Research from Shirmohammadi and Taylor [190] and our own exploration suggest that the enabled flash LED of one phone can be clearly identified in the camera image of another phone and used to infer each other's orientation and distance using Lambertian reflectance. When there is a continuous surface between the devices, the flash LED can be easily identified up to 10 meters. By letting each device blink a unique pattern every now and then, devices could be uniquely and spatially detected.

As the SurfacePhone is a mobile device,interaction through *Movement of the projector* can be measured using the built-in motion sensors and the optical flow of the camera's video stream. The projection could, for example, be changed from showing display-fixed content that moves with the device to showing a dynamic peephole into world-fixed content. This interaction technique is discussed in a different setting in the next chapter. Any table could thus become the personal virtual desktop that is explorable by moving the SurfacePhone like a flashlight over the table.

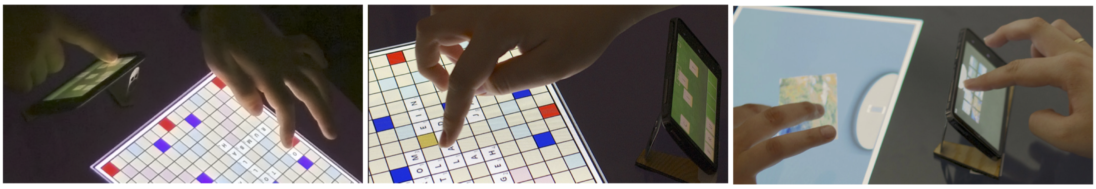


Figure 5.13: Transfer Interaction Techniques - The three different transfer techniques. (left) Direct transfer: A user places a scrabble piece at a precise position on the board through simultaneous touch. (middle) Binned transfer: Elements from the bin element (here the bench) on the phone are placed on the projection using touch-swipe. (right) Mediated transfer: The presenter drags another picture on the proxy element at the top of the phone.

A regularly occurring task when using the SurfacePhone is to transfer content from the screen to the projection and vice-versa. Following on [6] we can distinguish between three main categories of transfer-techniques that can be supported: *direct*, *binned*, and *mediated* transfer. These techniques are all based on *Direct Interaction on the projection*, *Input on the projection device* or a combination of both.

- *Direct transfer* is used to transfer an item from a specific position on the phone to a specific position on the projection or vice-versa. For this category we propose to use *Human Link*, as in [219]. The body of the user is conceptually used as a medium to transfer the content between the two displays. The user touches the content that they want to transfer on the phone and then, simultaneously or in quick succession, touches the point in the projection where they want to place it or vice versa (Figure 5.13 left). This represents a combination of *Direct Interaction on the projection* and *Input on the projection device*.

- *Binned transfer* uses a bin element on one or either displays that is used to place content items in the bin that then can be transferred using a form of *direct transfer*. For instance, to place a whole word in the scrabble game, users can position the letters on the bench (the bin) on their phone screen in correct order and then transfer them altogether by swiping over the target positions on the projection (Figure 5.13 middle). Similarly, users could select pictures on their phone to a bin and then fan them out on the projection with a finger swipe.

- *Mediated transfer* uses a proxy or gate element through which content is transfered. To transfer a object simply drag and drop it on the *proxy* (Figure 5.13 right).

5.2.2 Concept Prototype

To explore and evaluate the SurfacePhone concept, we built a concept prototype to validate that the proposed display configuration is actually desirable and usable. Through the placement of a standard mobile phone on a multi-touch surface it is easily possible to simulate the projection behind the phone. This allows us to test users' experiences providing a more robust, responsive, and clearer multi-touch surface than would have been possible through developing a technical prototype in the same time (which we present later on).

5.2.2.1 Implementation and Applications

The hardware setup consists of a Samsung PixelSense table running Microsoft's Windows 7 and Surface SDK; further two HTC HD 7 running Windows Phone 7.5 which offer a stand to arrange the phone on a table more easily. Markers placed below the phones allow them to be tracked by the table. Our software framework creates a 23cm $\times$ 18.5cm sized virtual projection 9cm behind and 3cm to the left of the phone. This size exceeds the projection size that is supported by our technical prototype by 33%. As phone manufacturers surely are able to build devices that support projections of these dimensions by using short-throw lenses or curved mirrors, we assume the projection size fits a realistic usage scenario. The devices communicate over Wi-Fi. As soon as phones are moved such that projections intersect, a merged projection is created. This merged projection can either be a graphically highlighted union of the individual projections, or something different like a shared playboard within the concave hull of the projections' corners.

Applications Based on this prototype we developed several applications that will aid in validating the SurfacePhone concept. In the following we will now present these applications.

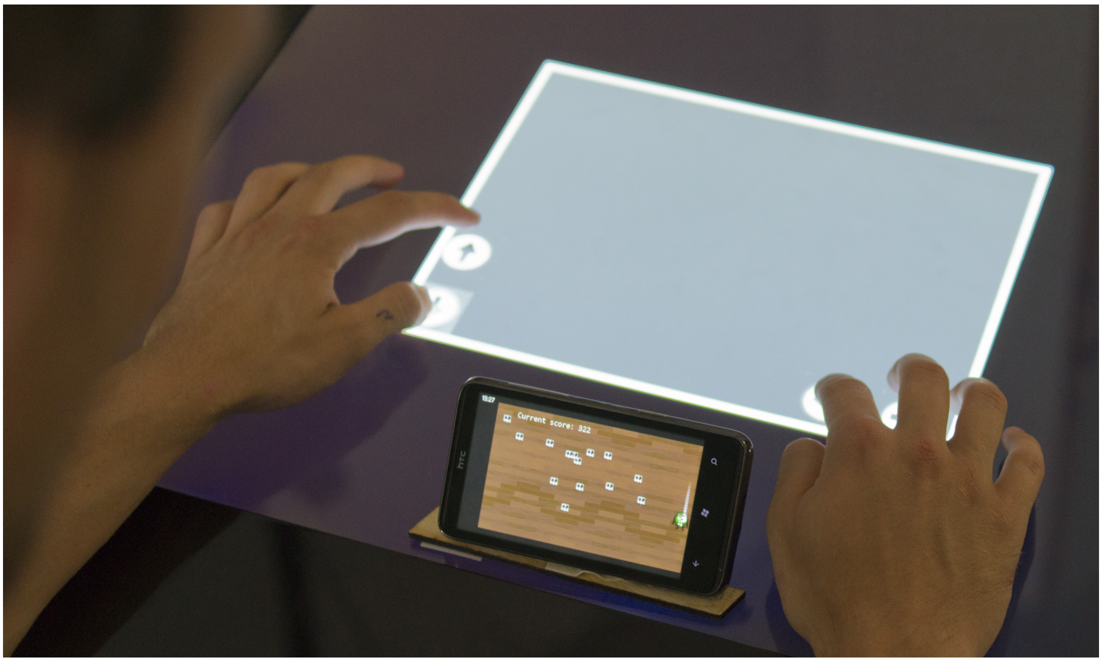


Figure 5.14: Concept Prototype SDSU "escape" game - Single-user "escape" game: The game controls are "outsourced" to address the fat-finger problem with on-screen controls.

Single-user game "escape" (SDSU) The "escape" game represents the SDSU category by supporting external controls on the projection in a single-user game. The task of the game is to escape monsters by moving the character horizontally and vertically on a play field without other obstacles. When playing the game on the mobile phone, the on-screen controls and finger of the user cover parts of the play field on the phone. By "outsourcing" the controls to the projection behind the phone, thus providing free sight on the whole play field, we assume users will perform better in the projected mode (Figure 5.14).

Multi-user presentation (SDMU) In this application the SurfacePhone is used to present pictures or slides to a small group of people in two different ways: Either the user publishes thumbnails to the projection by dragging the thumbnail on the *proxy* at the top of the phone screen. The audience can then use standard multi-touch techniques for rotating and enlarging the pictures to their will (Figure 5.13 right). The other possibility is that users browse their content on the projection and present items fullscreen on the phone by double tapping them (Figure 5.15). Different to the first way, the user gives up their privacy for the benefit of having a larger space themselves that can be quicker explored. Additionally the higher brightness and resolution of the phone might provide a better viewing experience compared to the projection.

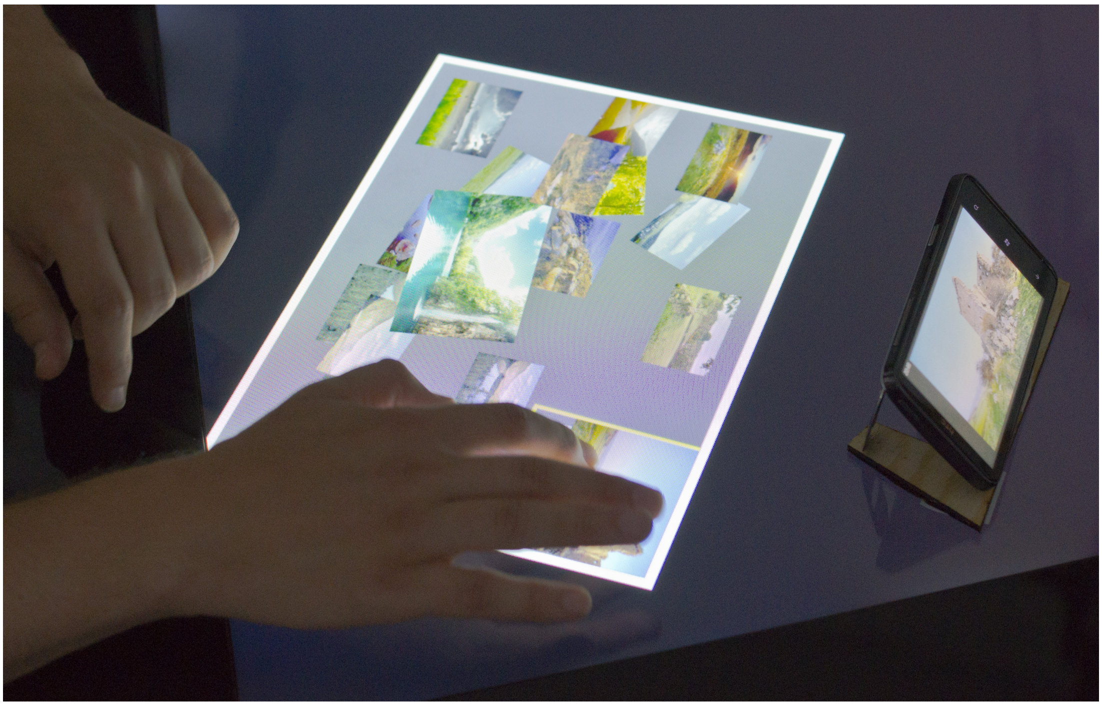


Figure 5.15: Concept Prototype SDMU Picture Presentation - Multi-user sharing from projection: the presenter selects images to be displayed on the phone screen.

Multi-device picture sharing (MDMU) The exemplary picture sharing – which would similarly work with other content types – is comparable to the SDMU presentation application. Users publish their thumbnails to the surface by using the *proxy* or *Human Link* techniques as in the presentation application. As soon as more than one device and user merge their projections by intersecting them, the merged space can be used to share all sorts of personal data. Thumbnails then belong to the joint surface, allowing all participants to explore pictures through multitouch operations and transfer them to their phone using one of the aforementioned techniques. When one of the participating users withdraws from the merged state, the view is split and the separate projections retain prior items and positions on their side. If items have not been moved to the phone, these items are moved back to the projection of the owner. This feature shall give users a simple means of privacy control as they can withdraw with items that they only want to present but not give away. This application is well suited for the setup where users sit next to each other (cf. previous section) on the same side, especially when images are part of a puzzle, whenever privacy is not an issue or otherwise next to each other on opposite sides.

Multi-device scrabble game (MDMU) The scrabble application (Figures 5.13 left) particularly emphasizes the private display on the mobile phone. It shows a standard scrabble playboard on the merged projections. The phone screen shows the letters available to the user and a virtual bench on the bottom where words can be arranged with the letters using drag and drop. On their turn, users either use the *Human Link* technique to place any letter, no matter if on the bench or not, by touching the letter and the target position on the playboard. Alternatively, they first put the letters to place in correct order on the bench and then swipe over the empty fields on the board to place these letters. Depending on whose turn it currently is, the board changes its orientation to face the corresponding user. Letters can be taken back to a precise position on the phone using *Human Link* or to a random position by double tapping them.

5.2.2.2 User Study

With this user study we assess the quality of the overall SurfacePhone concept and its several components using our concept prototype. Using the four aforementioned applications we assess input techniques (e.g. *Human Link* and *proxy*), output (e.g. size and visibility of displays) and possibly occurring problems such as undesired occlusions of the projection and physical demands of the MMDE.

We follow a qualitative approach using think aloud, structured interviews, and video analysis as no similar system is available for comparison. We had 16 participants who took part in pairs to create a more realistic collaborative environment. Their average age was 26 years, (ranging from 23 to 31 years) and 6 of them were female. All participants except one owned a smartphone and 3 of the participants had prior experience with multi-touch tables.

Procedure First we explained the concept of the SurfacePhone by showing them a concept design of the technical prototype and also to convince them that these devices can be built we demonstrated a Samsung Galaxy Beam projector phone. Finally, the experimenter briefly explained the prototype, how it works, and the different configurations (SDSU, SDMU, MDMU) which also represented the different phases of the study.

After that, both participants tried all four applications (one each for SDSU and SDMU, two for MDMU) for approximately eight minutes each. Before each application participants were given time to test the concepts relevant in that phase, for instance, merging of projections and different transfer techniques, until they had no further questions. In single-device applications they took turns in acting as user or audience/spectator. In multi-device applications both users operated their own device simultaneously. To ensure a constant learning curve, the order of applications was always the same, going from single-device and single-user to multi-device and multi-user applications, thereby constantly gaining in complexity. Before each multi-user application, users were allowed to choose device positions (see MDMU before) that fit the task according to their opinion.

For the study the participants had to use all aforementioned applications. For the picture presentation applications (SDMU) both participants acted as presenter and observer in turns. For the picture presentation in MDMU mode we added two tasks. One task was to share pictures that contained Waldo with your partner and the other was to solve a 3×3 puzzle collaboratively on the merged projection space.

While participants were continuously motivated to share their experiences aloud, after each configuration (SDSU, SDMU, MDMU) they filled out a questionnaire regarding the configuration and contained tasks. The questionnaire asked for experience with the applications as well as physical demand, fatigue, visibility of content, feelings regarding privacy, etc. After the study we let participants fill out the Post-Study System Usability Questionnaire (PSSUQ [118]). The study was video captured and later qualitatively analyzed. The textual answers were later analyzed using axial coding.

Results In terms of preferred interaction technique, to transfer information between the two screens, most participants favored *touch-swipe*. Comparing *touch-swipe* and *Human Link* in the scrabble game, 15 of the 16 participants preferred *touch-swipe*. When comparing *touch-swipe*, *Human Link* and *proxy*, nine participants preferred *touch-swipe* and five would rather use the *proxy* technique. This is also reflected in the physical demand. Ten participants stated that *Human Link* has the highest physical demand and four found *proxy* to have the highest demand. The same was reflected in their rating of success, 12 participants said they were most successful with the *touch-swipe* and 2 thought they would be better with either *Human Link* or *proxy*. In their comments the poor performance of *Human Link* became pretty obvious. A majority (9 out of 16) stated that when using the *Human Link* they had the problem that when they touched the phone's screen and tried to touch the projection the phone would slide away. This was especially seen as an advantage in favour of *touch-swipe*.

We asked the participants whether they developed a strategy to solve the puzzle and image tasks. All participants agreed that they followed a certain strategy. From the video analysis and the comments two strategies were particular promising. Three couples would actually change their original sitting position so that they would sit next to each other, allowing both participants to view each others phone displays, helping them to identify the correct pictures before putting them on the surface. Three other couples divided the work between each other so that one participant would move the pictures from the phone to the projection and the other would arrange the puzzle parts in the projection.

To evaluate a possible adoption of such a device we asked the participants whether they would recommend it to their friends and if so what would be the necessary circumstances. All but one participant answered that they would recommend it. Most of the participants found fulfilment of hardware constrains such as reasonable size and battery life to be mandatory. Privacy was desired by all participants as a needed feature and most of them thought that presenting pictures by selecting them on the phone screen and have the projection show them is more reasonable than the other way round. Ten participants stated that they like that they don't need to pass their device around when presenting to groups.

The preferred combination of devices and users was MDMU. Participants found the possibilities that arise from having a mobile device that ad-hoc can create a complex mobile multi-display environment very attracting. Besides games such as Battle Ships, Poker and Black Jack, the collaborative editing of documents, e.g. layouts of newspapers was seen as possible application scenarios. Two participants mentioned the case of ad-hoc meetings for example to collaboratively investsigate construction plans on a construction site.

The results of the PSSUQ are underlining these results. In the overall usability rating the SurfacePhone scored 84.8% and in the system usefulness the user rated it with 87.3%. The interface quality did only reach 81.9% but this might be because of the design of the developed applications. Overall the results of the PSSUQ indicate that the SurfacePhone is a useful new device to extend screen space in single- and multi-user applications.

5.2.3 Technical Prototype

The positive results of the first study motivated us to build a technical prototype of the SurfacePhone that can support aforementioned interactions. Our aim with this prototype is to investigate the technical requirements and challenges for such a device and find solutions for them.

5.2.3.1 Hardware Design

As no similar device configuration has been presented so far, we started from scratch. We chose the iPhone platform, since it was the only mobile platform that allowed two different outputs on screen and projection at that time. After we tested several different projection engines (e.g. TI DLP 2 or Microvision PicoP) attached to the backside of the device, it became obvious that without a fitting short-throw lens that is not

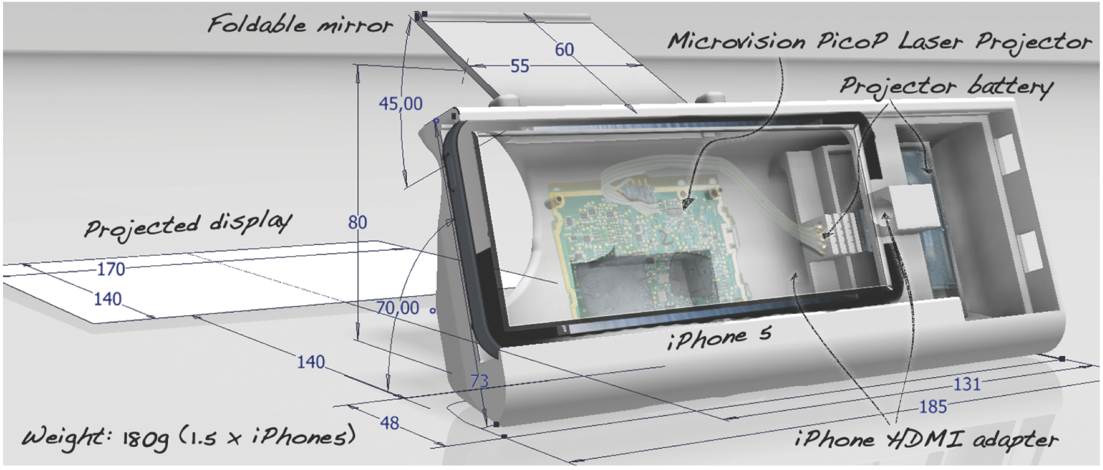


Figure 5.16: Design of the casing for the SurfacePhone Prototype - Design and instantiation of the SurfacePhone prototype that tracks finger touches using in-built camera and accelerometer. (dimensions in mm)

available, the size of the projection would become too small. We solved the problem by attaching a mirror to the top of the phone and the projection unit to its bottom. This way the distance from projector to surface is more than doubled and sufficient to create a projection much larger than the phone screen.

Our prototype consists of an iPhone 5 plus projector case as depicted in Figure 5.16. Both projector and mirror are 4cm to the right of the iPhone camera which is the minimum distance required for projector and mirror not to appear in the wide-angle view of the phone camera.

5.2.3.2 Implementation

Following our previous design considerations, the prototype should support *Direct Interaction on the projection* including different touch modalities, gestures, and tracking of other nearby SurfacePhones.

The software of the SurfacePhone is implemented in Objective-C and C++ on iOS with the help of OpenCV and openFrameworks modules (compare Figure 5.17). First we calibrate intrinsic and extrinsic camera parameters using a printed chessboard and projected chessboard patterns respectively. Having the parameters we can map the projected area from object space to an interpolated orthogonal view of the projected

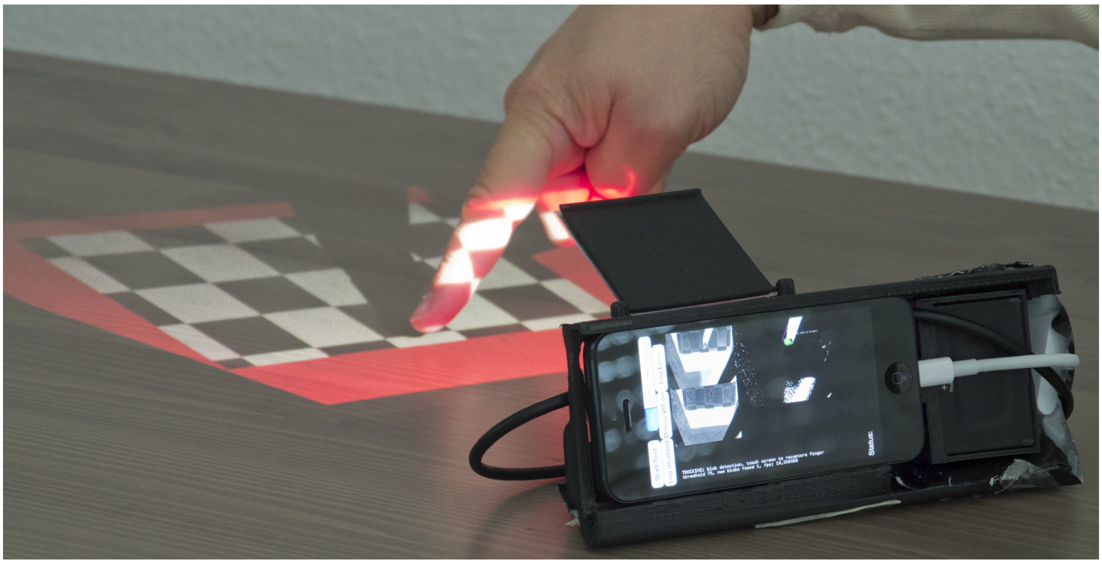


Figure 5.17: Working SurfacePhone Prototype during calibration - The red border (for illustration) in is the relic of perspective counter distortion. The phone shows raw camera image (top-left), background-image (bottom-left) and finger tracking with green dot at recognized fingertip (top-right).

region and use this for tracking. In a final SurfacePhone product this would only have to be performed once.

For the tracking to work robustly at arbitrary locations we must make sure that different lighting conditions are handled. We can let the iPhone automatically adjust exposure and focus of the camera to the center of the image to adapt to different conditions. However, we need the user's finger for a correct estimation. Therefore, in step 3, we ask the user to present their finger for 2 seconds to the center of the camera while we lock correct exposure and focus for future interaction.

In step 4, we capture a still frame for subsequent background subtraction. As the background of interaction can be arbitrary we use background subtraction to separate moving fingers from the background. This step is automatically performed whenever the device comes to rest on a plain surface. Since we constantly measure the accelerometer at 100Hz we can quickly recognize whenever the user starts and stops moving the device.

To eliminate shadows as best as possible we first convert the image to the HSV space and then work on the saturation channel. The literature recommends working on the hue channel to eliminate shadows, but we found that desk colors are often very similar to skin colors which is why we use the saturation channel that works more reliably in our scenario. After background subtraction we find blobs using openCV's contour finding algorithm.

Finally we have to decide which blob represents the primary finger, which one is a possible second finger and which ones are not of interest. Our blob sorting and filtering algorithm favors blobs with fingers, high finger probabilities, less circular shape (to filter hand areas) and lower Y position (to filter shadows appearing below fingers).

To detect the moment the user touches the projection we use the accelerometer of the device. The vibration the finger causes when touching the projection can be sensed by the iPhones accelerometer since the phone is rested on the surface. Finger trajectories that do not end in touches are simply analyzed for long directional movements or otherwise handed to the $1 gesture recognizer by Wobbrock et al. [225]. To track several projections we used the Qualcomm Vuforia SDK. Although not fully integrated into our prototype yet, we evaluated the use of Qualcomm's Vuforia on the SurfacePhone using projected frame- and image markers. Image targets are recognized even up to a 1/50 of the projection.

5.2.4 Design Guidelines

Based on our experiences and studies with the SurfacePhone, we can conclude the following guidelines for designing interfaces for the SurfacePhone:

- In our study all participants favoured having the private display. Therefore applications designed for the SurfacePhone should always consider how to use it properly especially since the location of use will seldom be a private one.
- Where possible, use interaction techniques that do not require simultaneous interaction on both displays. As the user has no hand available to keep the phone in place while touching the phone and the surface.

- The *proxy* was seen as an advantage in the SDMU case where the screens are divided between the users and nobody was able to intervene on the display of the other. Especially for use cases that follow a certain hierarchy between the users, we would recommend to use such mediated transfer techniques to share data between the different displays.

- The flexibility that the SurfacePhone provides in terms of collaboration and usage scenarios was appreciated by all participants. Therefore applications designed for this device should remain flexible to the location as well as the number of users and devices.

- The color of the projection surface as well as lighting conditions have a big influence on finger recognition. An infrared LED might help homogenizing the environment. This was already successfully done by Wilson for PlayAnywhere [218].

- Users should receive feedback about their touch intensity (e.g., a color meter around their touch) in order to be able to maintain a mental model of the strength of their touch. Further, touch thresholds should be personally adaptable to account for anatomic differences.

- The proposed adaptation of touch thresholds only works up to the physical limits of surface vibrations. For thicker surfaces this especially means that light and strong touch thresholds move closer together, possibly resulting in more falsely recognized strong touches. Thus, strong surface materials like granite should be avoided.

- Our first evaluation of the technical prototype revealed that the touch accuracy is not as accurate as on current touch devices. Especially overshooting was a problem. Therefore we would argue to include a personalized touch correction model such as [211].

- If possible, interface elements should be placed to the top left side of the projected display as the fat-finger problem is less of an issue here. When developing applications, interactive elements should have a radius of at least 50pixel to ensure a good accuracy.

5.3 Summary

In this chapter we addressed the research question of **How to align the projected display to support the user**. We investigated visual separation effects for MMDEs compared to the current literature of visual separation in fixed MDEs. Using a mobile eye tracking methodology, we compared different angular separations of a projection and a screen: two displays of different sizes and at different distances from the user. We determined that although task performance was not affected by the displays being in the same or in different fields of view, the number of eye context switches was over 30% higher in the condition where both displays were in the same field of view. We also tested various factors of mobility in our experiment and concluded that they did not affect visual separation. We finally present design implications in terms of types of displays used in MMDEs, physical arrangements of the displays, flexibility of design of MMDEs and mobility.

Additionally, we establish that through the use of an eye tracker, we were able to highlight interesting differences between different physical arrangements of displays; that may not have been revealed with trial completion time and accuracy alone. This further suggests that eye tracking is an interesting way to investigate visual separation issues.

We applied these findings in the design of the SurfacePhone and built a variety of possible application and interaction scenarios around it. It allows for ad-hoc Tabletop interaction in a mobile setup. We also learned which setups and transfer techniques users preferred in different scenarios. Later we demonstrated how the SurfacePhone can be built with only today's commodity phone hardware and the help of a specialized case and state-of-the art techniques for finger tracking and multi-modal touch recognition. Overall we can conclude that the SurfacePhone represents a new and interesting device for ad-hoc collaboration.

6

Spatially-aware Interaction Techniques for mobile projection devices

So far we presented designs for spatially-aware applications and investigated suited alignments of mobile projected displays and physical display that support the user. To turn mobile projected displays into interactive and useful output, suited interaction techniques are needed. Even though as presented in Chapter 3 a variety of different techniques exist, we want to leverage the spatial memory of the user to a higher degree as prior work. Therefore we at this point present our investigation of such two interaction techniques. First we analyse an interaction form that allows effective workspace switches on a mobile device as well as on a mobile projected display. Afterwards we discuss a concept for a dynamic peephole based interface for mobile projected displays.

The results of this chapter have been partially presented in the following publications [41, 124, 125].

6.1 Virtual Workspaces for Mobile Projection Devices

The use of multiple workspaces is common in desktop computing and integrated into most Operating Systems (OS). Known as workspaces on Linux, spaces in Mac OS or virtual desktop in Windows OS, they allow users to de-clutter their principal workspace and mitigate physical display size limitations by adding virtual display real-estate. For mobile platforms, the concept of virtual display space is quite different; it is often used to store icons but rarely to switch between tasks or activities that have already been started. Users instead go back to the main menu and select an icon to access the corresponding application. Since the use of multiple virtual workspaces was initially recommended as a technique to alleviate some of the mental workload in limited display real-estate [40], it seems natural to use this concept for mobile phone technologies.

Current desktop environments represent virtual workspaces using thumbnails spatially arranged in a line or grid. In mobile environments, since the devices themselves have intrinsic spatial properties, we propose to use these spatial properties instead of using a graphical representation. This will give users a more tangible and direct interaction with virtual workspaces that does not exist in desktop computing thus expanding the capabilities of current systems.

The concept of virtual workspaces as it exists in desktop computing can be translated to the mobile environment. Yet, this is not a straightforward process and the results obtained would be suboptimal if careful consideration was not given to the design, as the two categories of devices have very different characteristics and capabilities.

We explore the use of virtual workspaces in mobile environments. We present mSpaces, pSpaces and m+pSpaces, three (projector-)phone-based user interface prototypes, which allow simple and fast access to multiple virtual workspaces located around the user, drawing on their micro spatial memory. With mSpaces, the user accesses different workspaces by moving the phone to different physical locations, while pSpaces (Figure 6.1) gives access to multiple virtual workspaces by moving the mobile projected display to the various locations.

This interface makes full use of the users spatial awareness. Finally m+pSpaces makes use of both the screen and pico-projector. The mobile virtual workspaces are always linked to physical locations relative to the user. In two experiments we compare mSpaces to the current use of multiple concurrent applications on a mobile phone and

Figure 6.1: pSpaces - User accesses various virtual workspaces by moving the pico-projector to the associated physical space.

to p- and m+pSpaces. While we explore interaction techniques to access the various workspaces, we at this point do not investigate new interaction techniques for input.

6.1.1 Background

At this point we want to discuss the background of existing literature in the field of virtual workspaces.

6.1.1.1 Desktop Virtual Workspaces

Virtual workspaces are first introduced by Card and Henderson [40, 87], who initially proposed Rooms, a system for managing multiple virtual workspaces as a way to cope with limited display real-estate by expanding it into virtual workspaces. Multiple workspaces lower the cognitive overhead created by trying to switch tasks and move windows across a limited physical display real-estate. One of the authors' arguments for virtual workspaces is that they help overcome the limitations of small screen size, which is a highly relevant issue for todays mobile technologies. Ringel [173] proposes a taxonomy of organization strategies for users of multiple virtual workspaces. Five organization strategies emerged from a field study. Participants would consistently use the different workspaces to either: divide tasks, divide subtasks, change context between personal and professional usage, use multiple OS, and classify applications. They also show that virtual workspaces have different uses to multiple displays. It is therefore likely that adding an auxiliary display to a mobile environment would not replace the need for multiple workspaces. Finally, users with smaller displays used more virtual desktops, on average, [173] . This proves that multiple workspaces will be well suited to mobile environments that traditionally afford smaller displays.

6.1.1.2 Mobile Virtual Workspaces

We use mobile devices to refer to handheld devices such as mobile phones and tablets. Laptops are not considered to be part of this category as they are traditionally used sat down on a flat surface and not handheld. Fitzmaurice [70] presented the Chameleon system in 1993, offering spatially aware interactions with the environment using a palmtop computer. Despite, the concept of virtual workspaces as we know it in desktop environments is seldom implemented in current mobile devices and many mobile phones only offer the possibility of displaying static menu icons on one or more virtual desktops. Moreover while some mobile phones have the capability to display multiple applications at once, they do not exploit advantages offered by virtual workspaces as they exist in the desktop environment. Adapting multiple virtual workspaces to mobile devices will extend the current range of possibilities offered, such as providing users with a task partitioning tool.

While the concept of multiple virtual workspaces can be translated to the mobile domain, the interaction techniques need to be adapted. This is particularly evident since display sizes and interaction techniques are inherently disparate between desktop and mobile environments.

6.1.1.3 Representation of Virtual Workspaces

In the desktop environment, multiple workspaces are often displayed as a group of thumbnails, each representative of one workspace. They are traditionally organized as a line of thumbnails such as when pressing ALT+TAB in Windows OS or CMD+TAB in MacOS or as a 2x2 grid when 4 workspaces are being shown. In early systems, thumbnails were referenced to by numbers. They are now often presented as a thumbnail of the actual workspace with its applications positioned as in the workspace itself.

In existing mobile environments, the main screen can sometimes be extended to display additional static information such as extra application icons. These application launcher spaces can be represented as a line of dots in the main menu where one dot is highlighted, indicating the workspace in view. Some phones also propose solutions close to the concept of multiple workspaces in desktop computing such as the Nokia N900 where multiple applications can be running in additional virtual space. Yet, there is no point of reference to what the user is currently viewing with respect to the virtual space. In all instances presented above, there is a spatial relationship between the workspaces: one can be represented next to the other, above or below.

6.1.1.4 Conceptual Conclusions

As discussed in the theoretical background Yee presented Peephole [230], an interaction technique for a spatially-aware display that provides a window on a larger virtual workspace. They mention that this window could be used to display several applications on the same workspace where users could draw connections between applications. In mSpaces, we take this concept further by mapping the location of each distinct virtual workspace to a physical location, relative to the user. The user can access each workspace by orienting the mobile device in the direction of the workspace, hereby receiving permanent visual feedback to which workspace they are looking at. This technique utilizes the intrinsic properties of a context-aware mobile device, the users micro spatial memory through kinaesthetic cues that will ostensibly alleviate some of the users mental workload.

Furthermore, as analyzed in the related work, Cao and Balakrishnan [37] explore using a handheld projector to access multiple items on a single virtual space. We take a step further by accessing multiple virtual spaces at different physical locations with constant visual feedback. In m- and pSpaces, each virtual space is linked to a location relative to the user, who then points at the physical location to display the associated virtual space. In order to determine how virtual workspaces can be displayed in the mobile environment, we conducted a user study comparing various implementations.

6.1.2 User Study 1: Mobile Virtual Workspaces

This study aims to find out the type of interface is suited to implement virtual workspaces in a fully mobile environment. While there are many ways to implement mobile virtual workspaces, we compared the current use of mobile phones to two probable implementations.

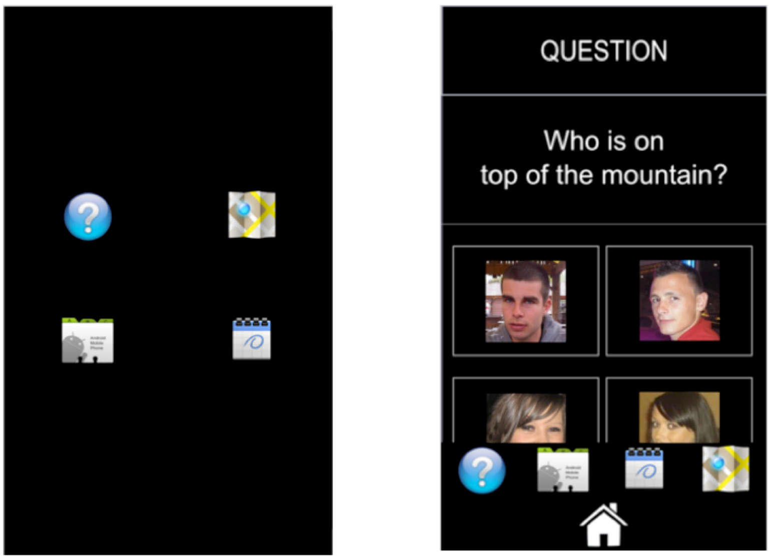


Figure 6.2: Simple Workspace Interfaces - No virtual workspace(left), menu of icons-Workspace switcher (right): a bar of icons representing the workspaces appears at the bottom of the screen.

6.1.2.1 Virtual Workspaces

In order to keep the current display paradigm used in mobile phones, each workspace contains one application only. The mobile phones screen would quickly become cluttered if more than one application was displayed at a time on such a small display. Ringel [173] indicates that the average number of virtual workspaces used in desktop computing is four; we therefore implemented the prototypes with four virtual workspaces and four running applications

6.1.2.2 Apparatus

The three conditions for the experiment are: no virtual workspace, which corresponds to the current use of mobile phones; workspace switcher, which provides a representation of the workspaces, and mSpaces, which distributes the workspaces across space. All conditions have been implemented on the same mobile phone, a Samsung Galaxy S running Android 2.3 OS.

No Virtual Workspace This condition reflects the current usage of mobile phones. To switch applications, the user returns to the main menu by performing a short click on the home button at the bottom of the screen. There, they click on the icon corresponding to the application they want to open. This operation must be performed every time the user wants to switch application. This technique is a typical interaction technique for browsing through applications in mobile phones. The menu displayed consists of a 2x2 grid of icons, each one representing an application (Figure 6.2 left).

Workspace Switcher This condition simulates the current metaphors for switching workspaces in the desktop environment. The workspace switcher is a graphical representation of the available workspaces, consisting of a bar of icons that appears at the bottom of the screen superimposing and partially hiding the current visible workspace (Figure 6.2 right). Each icon represents a workspace as in the no virtual workspaces condition menu. The user performs a long click (500 ms.) on the home button to access the workspace switcher as they would typically do for switching context on mobile phones when the functionality is available.

mSpaces The third condition, mSpaces, is a prototype that allows the user to choose which virtual workspace they want to display by moving to the physical space. The technique is in the same idiom as the interaction techniques presented in Imaginary Interfaces [80], Virtual Shelves [119] and Peephole [230]. In mSpaces however, we exploit the users spatial memory through kinaesthetic cues that are attached to the workspace switches. The user accesses the distinct workspaces with permanent visual feedback; without having to press any button; just by moving the device to a new physical location. mSpaces is a spatially- aware device, in which 6 degrees of freedom (DOF) tracking is realized using a NaturalPoint OptiTrack motion- capture system through IR-reflecting markers attached to the prototype (Figure 6.3). Workspace are separated from one another on the horizontal axis by a 30 degree angle as is advised in the literature [119]. In the experiment, the workspaces are positioned on a single vertical level but the prototype would in addition support having workspaces at multiple heights.

The position of the workspaces was fixed for the experiments and participants had a chance to become familiar with the system prior to the experiment.

6.1.2.3 Tasks

To evaluate how virtual workspaces need to be designed for mobile devices, we asked participants to answer questions for which they need to look up information using familiar mobile applications. We used spatial search tasks where participants needed to access multiple workspaces to find the right answer. Since the tasks are spatial, the applications do not require any user interaction, aside from navigating between workspaces and using the touch screen to answer the question. The applications visible to the users are pre-designed screenshots of applications that may contain some clues to answer the trial question. The user can retrieve the clues by perusing the screenshots using the workspace switching technique. This task is representative of a task where a user consults their agenda to give a date or location to a person on the phone or next to them. This task is not designed for on-the-move scenario but instead a scenario where

Figure 6.3: mSpaces - Spatially-aware mobile device. Different workspaces appear depending on the devices position.

the user would stop for a short moment to consult some data on their phone as this is current practice.

For the four workspaces available, four applications were used: a question (Figure 6.4a), contact list (Figure 6.4b), calendar (Figure 6.4c) and map (Figure 6.4d). These applications correspond to everyday tasks commonly undertaken on mobile phones. Four types of tasks were presented to generate different sets of workspace switches. type1 respectively involves looking at the contact list and the map; type2 at the map and the calendar; type3 at the contact list and the agenda; while type4 involves all workspaces and is therefore harder than other types.

The aim of each task is to answer the trial question. The participant doesnt know the type of the task; and for each trial, all four workspaces are available even if they do not all provide clues to answering the question. Once the answer is found, the participant gets back to the initial question workspace to validate their choice by touching an answer out of four choices on the touch screen (Figure 6.4a).

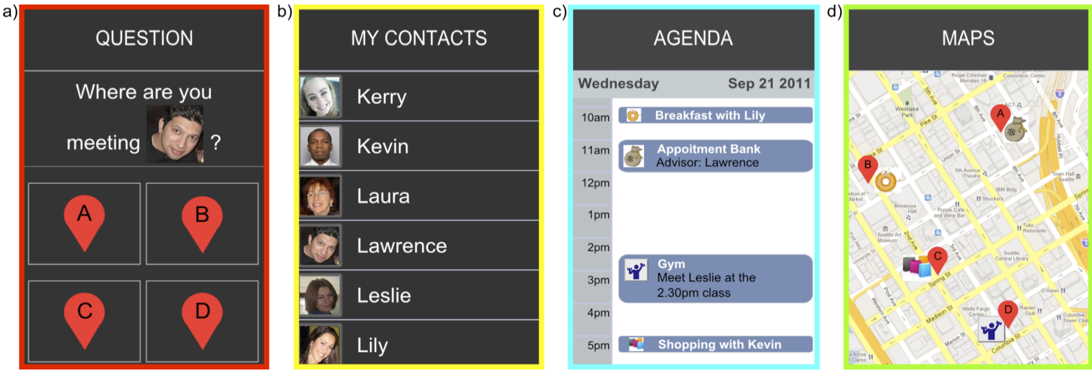


Figure 6.4: Screenshot of a Trial - To answer Question (a), the participant needs to match the picture with the persons name on the Contact list (b) and use the Agenda (c) to identify what activity they will be doing with this person. The next step is to use the Map (d) to locate the activity. Once the participant has navigated through all workspaces, they can answer the Question.

6.1.2.4 Procedure

12 volunteers (4 female) aged between 23 and 44 (avg . 29.5) were recruited from our research institution. All were right-handed, familiar with smartphones and touch-screen technologies. We used a within-subjects experimental design where each participant had to answer all questions and the type of virtual workspace was counterbalanced across participants. The task was explained individually to each participant who could try out each condition with a randomly chosen task in their own time. When they felt ready, they pressed the Start button to start the experiment and again before each trial. At the end of the experiment, users filled out a NASA TLX survey.

The independent variables were: The type of virtual workspaces: No virtual workspace (Nv), workspace switcher (Ws), mSpaces (mSp). The type of task: type 1 to 4 (t1 to t4). There were 6 questions of 4 different types per condition (i.e. type of workspace), which corresponds to 24 questions for each condition and 72 trials overall per participant. The experiment had: 12 participants x 3 virtual workspaces x 4 tasks types x 6 trials resulting in 864 data points.

6.1.2.5 Measures

The experimental software recorded trial completion time (MT) and error rate (ER) as dependent variables. MT is the total time taken to complete the task and is defined as the time taken for the user to perform a trial. The counter begins when the user presses Start and stops when the user clicks on one of the response buttons. If the user did not select the right answer an error was registered and the user was allowed to progress to the next trial. In an exit questionnaire we asked participants to complete the NASA TLX questionnaire. This allows assessing on a 7-point Likert scale subjective information for mental, physical and temporal demand; performance; effort and frustration. In addition to the NASA TLX we asked participants to rank the techniques and comment on their personal preferences.

6.1.2.6 Results

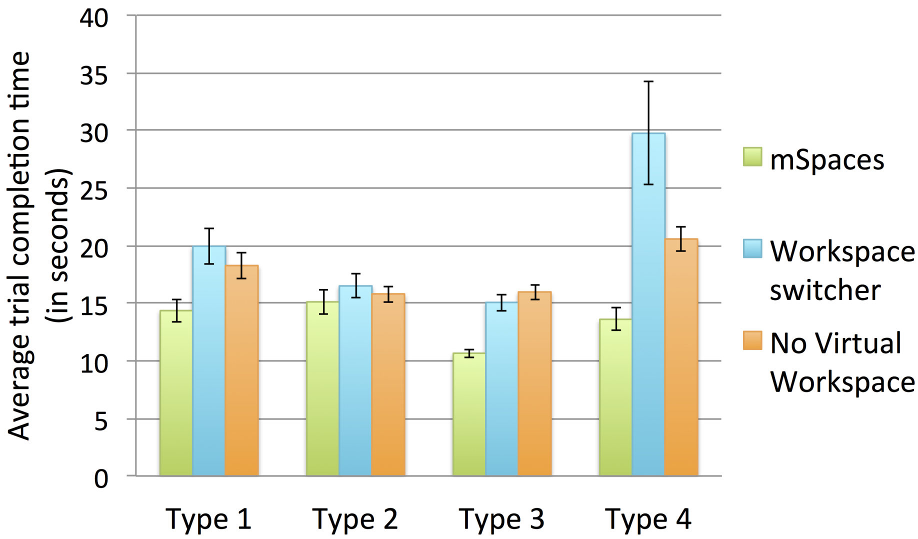


Figure 6.5: mSpaces User Study Trial Completion Time - Mean trial completion time for each technique and task. The horizontal axis shows the four task types.

We used the univariate ANOVA with Tamhane post-hoc pairwise (unequal variance) comparisons for our analyses.

Factor	$\mathbf{X^2(12)}$	**P**	**mSp**	**Nv**	**Ws**
Mental Demand	6.645	<0.05	1.54	2.08	2.38
Performance	7.0	<0.05	1.67	1.92	2.42
Effort	11.862	<0.05	1.42	2.08	2.50

Table 6.1: Results of the mSpaces NASA TLX questionnaire

Error Rate There were 59 errors out of 864 trials. With 8 incorrect trials, participants made fewest errors with mSp followed by 21 incorrect trials for Nv and 30 for Ws. All 59 trials with incorrect responses were removed from further analysis.

Completion Time (MT) The average trial completion time over all tasks and techniques was 16.6s with standard deviation of 4.1s. There was a significant effect of technique on trial completion time ($F2,22 = 10.85$, $p<0.01$); mSpaces was significantly faster than the other two techniques followed by No virtual workspace (Nv) and Workspace switcher (Ws). We found no significant difference between Nv and Ws. There was a significant effect of task type on MT ($F3,33 = 19.6$, $p<0.01$). Figure 6.5 shows the mean MT with standard error-bars for each technique and task type.

Subjective Evaluation The NASA TLX questions were analysed separately using non-parametric tests (k-related samples with Freidman Test Type) (Table 6.1). We found a significant difference for the following pairs: Mental Demand: (Ws,mSp), Performance: (mSp,Ws) and Effort: (mSp,Ws) & (mSp,Nv). All other combinations did not reveal significant differences. Low performance value shows that users felt they performed well. Users felt mSpaces required the least mental demand and effort. In overall ranking of techniques, 8 out of 12 participants preferred mSpaces to the other two conditions (Figure 6.6).

6.1.2.7 Discussion

Virtual Workspaces in Mobile Devices Participants felt that the traditional use of a phone, as in the no virtual workspace condition, although it was already common and well understood, was annoying to always start in the menu. Our experiment shows

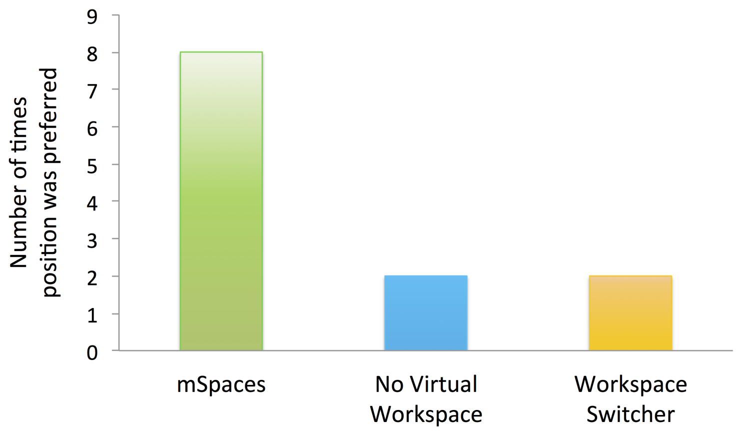


Figure 6.6: mSpaces User Study Subjective Preference - Subjective preferences of the different techniques.

that the use of virtual workspaces to complete tasks requiring serial switching through different applications can be significantly faster and less prone to errors than the traditional use of mobile phones. This is the case when comparing mSpaces to the no virtual workspace condition.

In the workspace switcher condition, the results were very similar to the ones of the no virtual workspace condition and slightly better for type3 tasks and worse for type4. This provides evidence that, although virtual workspaces can foster significant improvement over current use of mobile phones, they need to be carefully designed to fully realize their potential.

Micro Spatial Memory to Position Virtual Workspaces The results show that mSpaces improves decision-making accuracy. Additionally, the NASA TLX questionnaire shows that participants felt that they were less frustrated and required less effort to use mSpaces. This implies that virtual workspaces can therefore be managed on mobile phones using spatially-aware techniques. With mSpaces, people use their micro spatial memory through kinaesthetic cues to intuitively understand the positions of the various workspaces. In addition to being more efficient than the other conditions, mSpaces was preferred by 67% of the participants who enjoyed the opportunity to build a spatial knowledge of the location of apps in space and were able ”to arrange [their] apps around [them]". They found mSpaces faster, easy to use, quick and advanced, very intuitive and one mentioned that they could browse everything pretty fast and easily look something up again.

Our implementation of mSpaces is built for displaying virtual workspaces, nonetheless the strong results and very positive feedback obtained from the study lead us to believe that the use of their micro - or macro spatial memory to interact with mobile technology is very promising in spite of being under-exploited. It could be used not only to navigate through workspaces (mSpaces), menus [80] and shortcuts [119], but also to help the user when continuing a task after an interruption similar to [27].

Application Location Memory Aids The results show that using micro spatial memory only; participants can locate the different workspaces and navigate between them. Participants noted that after a short learning phase it was easy and comfortable to switch between apps and navigational help on the display would be useful. This suggests that mSpaces might requires some memory aid. This will also be useful as a reminder when users havent used mSpaces in some time. One way to provide a memory aid would be to display a map with the phones current position compared to the overall position of all workspaces. We therefore implemented a workspace viewer and conducted a user study to investigate if this functionality would improve usability.

Home Button Curse In this experiment, we noticed that imposing an extra click to request the workspace switch is time consuming and frustrating. Participants commented that the long press was irritating and that the no virtual workspace condition required a ridiculous amount of clicking. Users should indeed be able to access a different workspace without having to first return to a main menu. One way would be to use a button dedicated to workspace navigation, which could display all workspaces available and allow navigation between them. This notion of a dedicated workspace button is described in more detail in the design considerations section. mSpaces obviously does not suffer from the Home button curse as the device is simply moved in order to switch to another workspace, neatly avoiding the issue. The current use of mobile phones with a home button to return to a main menu may then not be optimal, especially for tasks requiring several workspaces or application switches.

One-Handed Interaction Technique We did not instruct participants on how the phone should be held. They could hold it as they felt comfortable and all held it naturally as they would with their own mobile phone. Nonetheless we noticed that for mSpaces, 75% of the participants held the device in their dominant hand only and interacted with the dominant thumb, leaving their second hand completely free. Only two held the phone in their dominant hand while interacting with the other hand. Finally one person used both hands after the first three trials as they [feared] to drop the phone when moving it too fast. For the two other conditions: workspace switcher and no virtual workspace, 7 out of 12 participants used two hands to hold the prototype. This is despite both techniques being implemented on the same prototype as mSpaces.

Number of Workspaces For a task with a higher number of switches such as type4 tasks, there was a significant difference in time completion across the techniques. mSpaces was more efficient and less error prone than the other two conditions. The number of workspaces may well therefore influence which technique is most appropriate. With the number of applications being concurrently used on mobile phones growing, mSpaces seems better suited than other techniques that would clutter the phones screen with extra icons, switches or scroll bars. Yet, mSpaces scalability will need to be determined in future work. According to Ringel [173], some users prefer partitioning information on their screens by using external displays rather than virtual workspaces. We are going beyond this statement by proposing to improve the use of mSpaces by adding an external display (pico-projector) on the mobile phone and fitting it with our mSpaces approach.

6.1.3 User Study 2: Extending mSpaces

We propose to test if we can further improve mSpaces by including a mobile projected display as well as a workspace viewer. We propose to implement pSpaces, a projected version of mSpaces. pSpaces has the added advantage of displaying the workspaces externally from the device via the projection beam, which may improve the speed and accuracy of users when performing a task. In this study, we also compare both solutions to a hybrid version: m+pSpaces where the main workspace is displayed on the phones screen while other workspaces are being projected one at a time depending on the position of the projection as in pSpaces. For both pSpaces and m+pSpaces prototypes, the user points the projector at the physical space to display the virtual workspace corresponding to the location.

6.1.3.1 Apparatus

All prototypes used a Samsung Galaxy S and an attached Microvision ShowWX+ pico-projector, of size 14mm and weight 122g. To guarantee the comparability of results the mSpaces prototype used in this experiment uses the same hardware as pSpaces and m+pSpaces with the projector turned off.

Figure 6.7: pSpaces Workspace Viewer - Spatial representation of the virtual workspaces on the top part of the phones screen. The phones position is represented by a white dot.

pSpaces We implement mSpaces on a pico-projector connected to a phone where the user can point at the virtual workspace to display it (which can be seen in Figure 6.1). The pico-projector is fixed to the phone and the participant moves the projector-phone to different locations to display different content. The motion-capture system used is the same as for the mSpaces prototype described in the prior user study. The computer determines which workspace to display depending on the prototypes position.

All workspaces are accessible and displayed via the projection. In pSpaces, a spatial representation of the virtual workspaces, workspace viewer (Figure 6.7), is displayed on the phones screen. Workspaces are represented by a thumbnail and not an icon, as in current mobile phones displays; allowing users to benefit from both their visual and spatial memory [117]. The workspace viewer provides constant information on users location in the environment compared to other workspaces. By this it represents a *Figural space* representation of the real environment. This answers some of the concerns addressed by participants in user study 1 where some felt that they couldnt remember the exact physical location of workspaces: "you first had to discover and "save" [remember] the positions of the apps". In order for participants to complete the task and to not introduce a new interaction technique, participants used the touch screen to answer the question as for mSpaces and by this employing a *Input on the projection device* technique. Participants needed to return to the question workspace for the answer buttons to appear on the screen below the workspace viewer.

m+pSpaces This condition is a hybrid version of mSpaces and pSpaces where the main workspace in use is displayed on the phones screen while the other workspaces are accessible by projecting towards their physical locations (as for pSpaces). As the only input needed from the user is on the question workspace, the latter is defined as main workspace, constantly displayed on the phones screen. Since the screen will be occupied by the main workspace, the workspace viewer will not be displayed or made available. We discuss in the design considerations section some interaction techniques that can be used to define what workspace to display on the phones screen.

6.1.3.2 Procedures

The tasks used for this experiment are the same as for the previous experiment. We recruited 12 new participants (3 female) for this study aged between 23 and 45 (avg. 31.3), all but one right-handed and all but two smartphone owners. While all participants were familiar with touch screens, only two had used a pico-projector prior to the study. We used a within-subjects experimental design where all participants had to answer all questions and prototypes were counterbalanced across participants. The tasks were explained individually to each participant who could try out the prototypes with a randomly chosen task in their own time. When they felt ready, they pressed a Start button to begin the experiment and again to initiate each trial. After the experiment, users filled out a NASA TLX survey.

The independent variables were: The virtual workspaces prototype: mSpaces (m), m+pSpaces (m+p) and pSpaces (p). The type of task: type 1 to 4 (t1 to t4). There were 6 questions of 4 types per prototype corresponding to 24 questions for each prototype and 72 trials overall per participant. In summary the experimental design was: 12 participants x 3 virtual workspace prototypes x 4 types of tasks x 6 trials = 864 data points.

6.1.3.3 Measures

The experimental software recorded trial completion time (MT) and error rate (ER). In addition, the number of switches between workspaces (SW) was recorded as dependent variable. SW corresponds to the number of times (switches) the user stops on a workspace during a trial. We measure this data to ensure that visual separation effectsin the m+p and pSpaces conditions do not hinder the results. We count a switch each time the user spends at least 300 consecutive ms on a workspace. There is no maximum number of switches as users can change workspaces as many times as they want until they find the answer to the question. Since m+pSpaces has one workspace displayed on the screen, SW is recorded as the actual number of switches per task minus the minimum number of switches required to perform this type of task with a given prototype. We gathered the same qualitative data as for user study 1.

6.1.3.4 Results

We used the univariate ANOVA with Tamhane post-hoc pairwise (unequal variance) comparisons for our analyses.

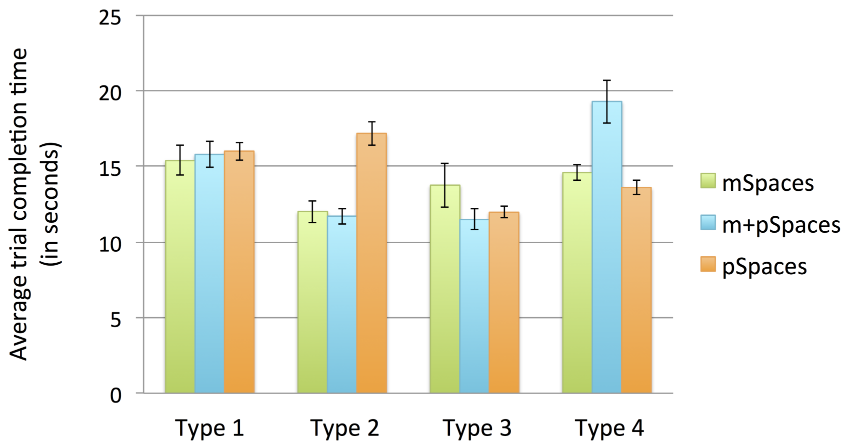


Figure 6.8: m+pSpaces User Study Trial Completion Time - Mean trial completion time.

Error Rate There were a total of 43 errors out of 864 trials. With 9 incorrect trials, participants made fewest errors with the pSpaces technique followed by 16 incorrect trials for mSpaces and 18 for m+pSpaces. The 43 trials with incorrect responses were removed from further analysis.

Completion Time (MT) The average trial completion time overall was 14.4s with standard deviation of 2.6s. There was a significant effect of type of task on trial completion time ($F_{3,33}$=17.09, $p<0.001$). Yet, we found no significant effect of workspace prototype (compare Figure 6.8).

Switches Between Workspaces (SW) The average number of switches overall was 4.2 with standard deviation of 0.2. There was a significant effect of prototype used on number of switches ($F_{2,22}$=5.83, $p<0.05$). mSpaces (m) and m+pSpaces (m+p) resulted in significantly less switches than pSpaces (p) (Figure 6.9). There was also a significant effect of type of task on trial number of switches ($F_{3,33}$ = 67.9, $p<0.001$)

Subjective Evaluation The NASA TLX questions were analysed separately using non-parametric tests (k-related samples with Freidman Test Type). We found no sta-

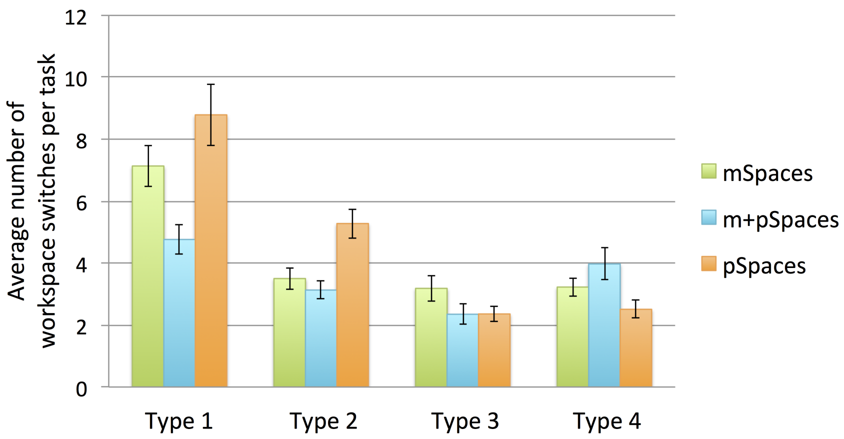


Figure 6.9: m+pSpaces User Study Workspace Switches - Switches between workspaces for each technique and task.

tistical difference between the different prototypes on any of the NASA TLX factors. In terms of preferences, 8 out of 12 participants preferred m+pSpaces, 2 preferred mSpaces and 1 preferred pSpaces (compare Figure6.10).

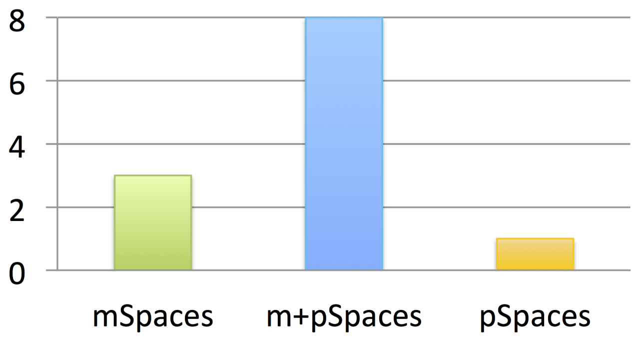


Figure 6.10: m+pSpaces User Study Subjective User Preference - Results of the subjective user preference.

6.1.3.5 Discussion

Spatially Aware Virtual Workspaces In this experiment, we did not find any significant difference in the overall task completion time over the three spatially-aware virtual workspaces prototypes. In the first study, for the same tasks, we found that using spatially aware workspaces was significantly faster than current usages of mobile devices. We conclude that when exploiting micro spatial memory greatly improves current usages, the way in which the spatially aware system is designed does not seem to influence the systems pace. This further reinforces our first user study finding that virtual workspaces need to be designed with considerations to users micro spatial memory.

Application Location Memory Aids During the first study where mSpaces is compared to current usage of mobile phones, users have reported needing time to learn the position of the workspaces and mentioned that navigational help on the display would be useful. We decided to provide users with a workspace viewer in the pSpaces condition since all workspaces were projected and the phones screen could then be used for displaying the workspace viewer. The idea was to provide users with a constant reference to their position in the environment. This technique also relates well to current existing techniques for switching workspaces in the desktop environment. To our surprise, when we asked users if they found the workspace viewer helpful and if they used it, 10 out of 12 participants who answered, unanimously replied that they were not using it nor finding it useful. The reasons they provided were that it was quite easy to spot the projections and easy to remember the positions. They mentioned they "concentrated on the projection", "knew the arrangement already from the task [they] did before" and that the other screens were sufficient.

This suggests that with pSpaces, users are able to remember the position of the workspaces without needing any workspace viewer. Projection onto the external space has rich spatial memory cues e.g. landmarks that, when combined with the kinesthetic cues of moving the device, help users remember the location of the workspace. Additionally, no participant mentioned struggling with finding the position of the workspaces across all conditions, contrary to the first study. Furthermore, in the second study, contrary to the first, all conditions expect users to remember the physical locations of the workspaces so it is possible that the nature of the task condition people to remember the workspaces locations better than in the first study. This leads us to think that when there is no reference to other types of interaction; people feel comfortable and lose their apprehension towards using such interfaces.

Workspaces Switches The number of workspace switches is significantly higher for pSpaces than for both mSpaces and m+pSpaces (Figure 6.9) while there is no significant difference in the average time needed to perform the trials; which shows that users switched between workspaces in a faster way using pSpaces. At the same time, the NASA TLX shows that participants did not find pSpaces more mentally or physically demanding or even more frustrating than the other two prototypes. Since we found that participants had no issue finding the position of the workspaces, we can conclude that participants chose, whether consciously or not, to switch more often betwen workspaces in the pSpaces condition. This is also very likely to be the reason why users made considerably less errors in performing the task with pSpaces. This higher number of switches seems to indicate that pSpaces provide an easy avenue to externalize users thoughts as in Chapter 5.1.

One Handed vs. Two-Handed Interaction In this experiment, only two participants held the prototypes in their dominant hand; while all others always held the prototypes with both hands. We hoped that the projector would be small and light enough to not affect the interaction technique. Unfortunately, the device was bigger and heavier than in the first study and that affected the interaction. Since this mSpaces prototype was held in both hands instead of one for the first study prototype, we believe that there is a potential for pSpaces and m+pSpaces to also be one handed techniques provided a smaller embodiment of pico-projection technology inside phones.

Projected Virtual Workspaces In terms of performance, pSpaces appears to be the best technique as users answered more accurately for the same completion time and swap workspaces more often, probably as a way to externalize their thoughts. Nonetheless, m+pSpaces was preferred by 75% of the participants and for performances similar to mSpaces. Participants preferred mSpaces "as it [is] useful to have the task visible all the time while working on it" and "as it was interesting to have one workspace always in sight". Some participants also liked "having the screens in [a] big size on the wall and at the same time to have the question at hand" and finally one mentioned that it "somehow "divided" the task and space".

In summary, whether due to performances (pSpaces) or user preferences (m+pSpaces), external projection improves the capabilities of spatially aware virtual workspaces.

6.1.4 Design Considerations

Our experiments show that virtual workspaces have the potential to improve the usability of mobile environments. In this section we would like to propose some design considerations for m- and pSpaces.

Creating Virtual Workspaces and Managing Applications There are different aspects to take into consideration when designing virtual workspaces. The workspace needs to be created and an application needs to be allocated to this particular workspace. The virtual workspace also has some attributes such as size and position that need to be defined. We present this set of design considerations below.

Creating and Positioning New Workspaces There are two strategies for creating workspaces in the space around the user. The first consists of creating an empty workspace and moving applications inside it in separate actions. The second consists in directly positioning an application and creating the workspace at the same time. The latter strategy is the most efficient when there is only one application per workspace. We advise no more than one application per workspace to keep the current mobile device interaction paradigm, and due to the small amount of screen real-estate available for phone handsets.

We propose to introduce a specific button – software or hardware – to trigger and control the workspaces management. This will preserve the one-handedness of the interaction technique while keeping it intuitive for the user. The interaction can be hold and release based where the user holds the button, moves the device to a physical location and releases the button to complete the operation. This technique allows a direct allocation of the application in view to the newly created workspace.

Another solution is to implement a drag-and-drop approach similar to what Boring et al. [28] proposed. For pSpaces, an application launcher could be displayed on the mobile screen for the selected application to open up on the actual external projection at the current pointing direction of the device. The workspace position can then be directly controlled by pointing or moving the device to the physical location corresponding to where the workspace will be residing.

Workspaces Dimension While in mSpaces the workspace size is limited to the size of the mobile phones screen, as discussed in the previous section, in pSpaces the corners and size of the workspaces could directly be defined by manipulating the projection area in the environment through an appropriate gesture via the device itself. We could envisage having a virtual workspace larger than the size of the phones screen in mSpaces, similarly to the concept of Peephole displays [230], but this would require a set-up stage via the touch screen or some movement recognition technique and incur additional interaction controls to scroll through the virtual workspace. This will be discussed later on in this Chapter as well.

Moving Applications Across Workspaces When moving applications across workspaces we need to differentiate between m- and pSpaces. In mSpaces we propose to use a similar approach to the way applications are already organized on a smartphone application launcher space. In pSpaces we can take advantage of the mobile device screen to grab an application and drop it on an arbitrary workspace after having pointed at it. We propose a special grab gesture (such as quickly tilting towards the user, as in Borings Tilt interface [29]) to copy the actual projected workspace to the screen in m+pSpaces so the user can provide touch input to an application.

Repositioning Workspaces The thumbnails representing workspaces or applications can be rearranged on the workspaces viewer on the mobile devices screen to rearrange the actual physical location of workspaces. The spatial alignment of the thumbnails and their sizes on the screen will directly translate into the spatial alignment and actual sizes of the workspaces surrounding the user in an appropriate manner such as by inserting at least 30 degree separation angles between workspaces. The rearrangement of thumbnails on the screen will then lead to the corresponding spatial realignment of the workspaces themselves. In the case of m+pSpaces, the same technique is used for ad-hoc selection and modification of which workspace is to be displayed on the screen and respectively the projection.

Finding Virtual Workspaces We show that users get shortly accustomed to their current workspaces spatial configuration and do not need memory aid. Yet, this will be useful when people move workspaces around or create new ones and disturb the established spatial arrangement. Similarly, it will benefit users who have high numbers of workspaces or havent used the system in a long time. An overview of the arrangement of the workspaces on the device screen is needed as well as a finding function which allows rapid access to a workspace. Finding a particular workspace could be done by displaying arrows on the screen pointing to the direction of the workspaces location. Haptic feedback could be used to indicate the position of a workspace, which would allow users to simply wave the device until they receive the haptic feedback (similar to Sweep-Shake [175]).

Favourite Configuration Storing favourite configurations would allow for the configuration of different arrangements based on context as Böhmer [26] proposes for icons. In fact, when the phone is used in different contexts e.g. personal or profesional or when used by more than one person, different favorite configurations may exist. For multi-users, it is equivalent to starting ones own session on a shared computer.

Applications and Tasks Awareness To be used for more effective task management, we can provide awareness that a task in another virtual workspace requires a users attention by adding some visual feedback. We could display a coloured bar on one side of the display (reflective of the position of the other workspace) and for example changing its colour [140] to indicate the status of the task. Another possibility is to use a halo technique [14] or an off-screen visualization pointing triangle [67]; and haptic feedback could also be used. For example, a light vibration would indicate an alert on a workspace which is currently not visible associated to a stronger vibration when the user hovers over the workspace in question as we describe in the creating virtual workspaces section above.

Deleting a Virtual Workspace A gesture, such as drawing a cross while pointing at the workspace could be used to delete it. Alternatively, a specific button could indicate deletion of the workspace. This may require a confirmation click or movement. If the deletion is not linked to the deletion of the application on the workspace, then the confirmation is not compulsary as the effects of a mistake will only be minimal.

6.2 Dynamic Peephole Displays

While the first part of this chapter covered a novel technique to switch applications on a mobile projected display, the second part investigates an interface design of a spatial interaction technique for such displays. To not only allow fast switches between applications but also to allow interaction with such displays.

As discussed in the related work, one disadvantage of mobile projected displays is that *Direct Interaction on the projection* is only possible when the user is very close to the display resulting in very small displays. Due to physical constraints the user is forced to be a certain distance away from the projection surface to create a large projection. For multi-user scenarios the possibility exists that one person is holding the device while others interact with the projection [53]. But in single-user scenarios one is facing a distributed user interface where parts of the interface remain on the display of the device while other parts are moved to the projection. This will cause attention shifts between the projection and the device, which inherits the controls for manipulation and interaction as discussed in Chapter 5. Even though the distance and with that the distribution may be rather small it is still a problem that restrains projector phones from exploiting their full capabilities.

But not only the interface is distributed also the information shown can be spread over the two displays. For example most users would prefer to not project private information and rather only see them on the display of the projector phone where they can be kept private. Depending on the relative orientation of the projected display and the devices display to each other, the distributed information may not be visible at one glance and the users are forced to shift the attention. As presented in Chapter 5.1 visual separation is not a big problem in those cases. But the results of chapter should be kept in mind when designing such distributed interfaces.

To tackle these problems of projector phones we will discuss an expanding interface design that employs a dynamic peephole metaphor to explore large-scale information. Peephole Displays are closely related to the category of spatially aware displays. As mentioned in the theoretical background, dynamic peephole interaction is a technique where a spatially aware display is moved and through this the viewport showing a part of the virtual information space is changed. In this way it is possible to make the whole information space visible which otherwise would not fit within the small display. This technique creates a positional mapping between the virtual information space that the user wants to explore and the real world. Doing so they allow to exploit the users micro spatial memory for navigation in the virtual information space. This technique has in several different use cases already been proven to be superior to static peepholes [39, 92, 230].

6.2.1 Dynamic Peephole Interaction for Projector Phone

Due to the distance that the user has to bring between him and the projection surface to create a projection from suitable size he might not able to reach projection with one of his extremities. In a multi-user scenario this possible as another user, which is not holding the projector, can reach the projected display which was for example shown in [53]. To enable interaction with the projected display in a single-user scenario new user interfaces and interaction techniques are needed to handle the whole information

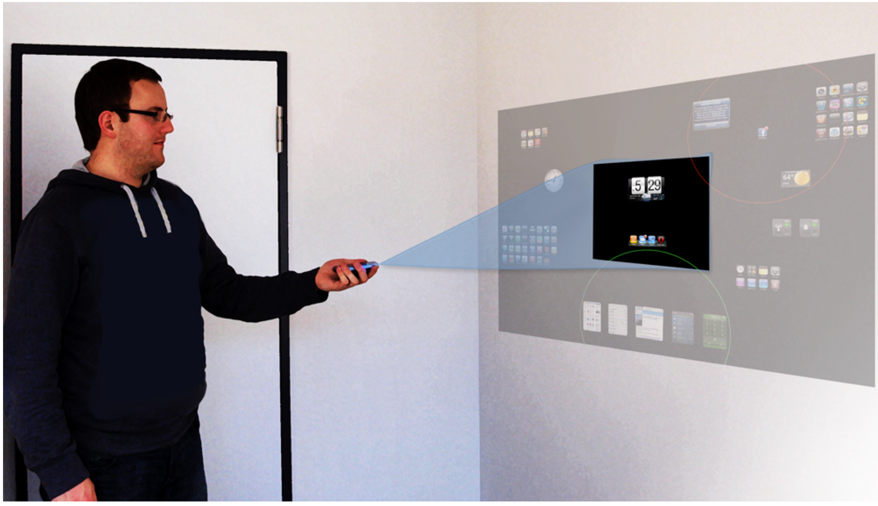


Figure 6.11: Dynamic Peephole Interaction for Projector Phone - Design of a dynamic peephole interface for projector phones. Only a part of the information space is visible (in this figure the opaque part). Notifications on different parts of the information space are indicated using the Halo technique [14].

space of the device. With the user interface being distributed between the device and the projection the user has to shift his attention from the content of the projection to the content on the devices display or the input constraints on the device. The principal goal of our presented design, is to reduce the amount of attention switches to a minimum. To achieve this, the focus of interaction will be the projected display. This is achieved by moving most of the interface elements to the projection and making the elements that are needed on the device as intuitive to interact with as possible. The devices display is only included for information display if necessary or if the user wish to use it. Especially for private information this is an advantage. Since the total information space is often even larger than the projection and to assure that everything is readable and visible from distance only an excerpt of the whole is projected. To navigate in the projected virtual information space we propose to make use of the dynamic peephole display metaphor, so that only a part of the whole information space is shown (compare Figure 6.11).

Similar to SpotLight [170, 171] the change of the viewport of the projected information space is done by device movements. But compared to spotlight we do not rely on device movements and rather focus on the devices orientation. By changing the device orientation the projection on a wall would move and could cover the whole wall as well. The drawback would be a distorted image but therefore solutions already exist [56]. Most modern mobile phones are equipped with sensors like accelerometer and compass and Apples iPhone 5 or Googles Nexus 4 additionally host a gyroscope. These sensors can be used to exactly determine the devices orientation and to measure even the slightest change. By using the devices orientation rather than movement it is also possible for the user to retain a comfortable position and still explore the whole information space. Sometimes the projection surface maybe limited and this technique may not be feasible therefore an easy switching method to a static peephole metaphor should be integrated as well. This static peephole could then be shifted using the touchscreen of the device (or an isometric joystick). This input should then be restricted to a designated area on the touch-screen or whole touchscreen should be used. When using the whole screen as soon as the user starts to move the touching finger the input should not be processed by the underlying UI-element but be interpreted as a desired movement of the static peephole.

When the information space is larger than the projection the center of the information space is right in front of the user in the starting position. But in many cases the information of interest is not in the center and the user wants to change the center to get a better viewing angle. Therefore, when using a dynamic peephole metaphor, we argue for adding a clutching method. This technique has been already successfully employed by Rapp et al. [170, 171]. By pressing the clutch button on the devices screen the movement detection is stopped and the user can move the last shown viewport to the desired new position on the projection space and when he releases the button the center of the information space is shifted accordingly.

The selection of a specific object on the projected display is problematically as direct interaction is not possible. An obvious solution would be indirect input in a mouse-cursor-like way controlled by the touchscreen. This seems to be cumbersome because it would require precise input on the touchscreen, to steer the cursor towards the desired goal, while at the same time controlling the dynamic peephole and focusing on the projection. Therefore we propose the cursor to be static in the center of the projection. By that it would move over the information space accordingly with the peephole metaphor. This would make everything easily accessible. For more complicated and detailed operations we propose to use the devices screen. One example is selecting an item from a combo box or a list, the user select the combo box on the projection and then the different selectable possibilities are shown on the devices screen where one can select a specific item. This would for example allow the interaction with private data that then would only be shown on the devices screen.

Objects that are not shown in the projection - which are off-screen - but desire the attention of the user e.g. an incoming text message, are indicated with a Halo [14]. Furthermore private information such as text messages or emails should always be firstly shown off-screen in the projection and additionally a small preview should be shown on the screen of the device. With that the user has to decide actively if he wants to explore the private information on the projection or the screen of the device. Furthermore this could be a general solution for private information that needs the users attention. A tactile feedback could indicate that the user now has to focus on the devices screen. With SpotLight [170] the usage of a zoomable interface was already presented in a mobile projection setup. While in SpotLight it was controlled using a scrollwheel we argue for interpreting physical movements towards or further away from the projection surface as a zoom-in and -out. Technically this could be achieved e.g. by using the optical flow of the integrated camera of the mobile device. Physical movement would have the disadvantage of the projection screen getting smaller or bigger depending on the distance. On the other side this could facilitate the interaction with private data in such a way, that private information is only visible when the user is close enough to the projection surface to shield the projection for uninvolved passersby.

The purpose of all the aforementioned design decisions is to minimize the number of context switches needed to operate the device. Therefore we also propose to use haptics to indicate which of the two screens should be in the center of attention. With current haptic possibilities experiences indicating a direction are possible on todays mobile devices. Using such a vibro-tactile feedback that e.g. goes from the direction of the projection towards the user, an incoming notification for which the user has to switch his focus to the devices display can be communicated. Such indications can be used to actively steer the users attention and therefore can lead to fewer context-switches [43, 82].

6.2.2 Prototype

Figure 6.12: Dynamic Peephole Prototype - This image was created by blending two images of the prototype taken from the same viewpoint. The accuracy is very high even though a small can be seen in the blending which is due to position inaccuracy.

To show the validity of the presented design and how it could simplify everyday tasks like finding a hotel on a map we created a map application that includes nearly all design ideas presented in the section above (compare Figure 6.12). This prototype is created to inform for future evaluation of the design in small focus groups in controlled user studies. Since Hang et al. already showed that map interaction could benefit from a projection [82] we have chosen the task of map interaction to explore the possibilities of our design. While Hang et al. relayed on panning using buttons for their study our concept allows easy interaction with the map using the presented dynamic peephole design concept.

6.2.2.1 Map Interaction

To explore the whole map the user simply changes the orientation of the device. By this, the application than reveals the whole map, part by part using the peephole metaphor (compare Figure 6.12). In line with a normal map application, the map can be enhanced with different layers containing Points of Interest (POI) like georeferenced Wikipedia articles, Hotels or the Google Latitude position of friends. The user can explore the detailed content of the POIs by moving the crosshair on the projection over the points and click on the select button. The information connected to the POI is then either shown on the projection or the devices screen when it is private data. For example for a Foursquare position of a friend the projection only shows the point but the name of the friend is only revealed on the devices screen. To select more than one point the user can simply press down the select button and drag the crosshair just like a mouse cursor. Layers can be added or removed to the map using either a drop-down menu or check-boxes. They are located on the devices screen to maintain easy and fast access (compare Figure 6.13). Off-screen POIs are indicated by Halo‘s [14]. With that most of the discussed design concepts were realized inside one example application. Of course a complete phone interface with all the capabilities of modern mobile phones would be much more complicated to realize.

6.2.2.2 Implementation

Since up to now no suitable projector phone is commercially available that is able to show different contents on the screen of the device and the projection we had to create our own prototype. On account of this we use a Microvision ShowWX laser projector connected to a fourth generation iPod Touch as can be seen in Figure 6.13. The two devices together weigh 220g and with that are easy to handle. The projector provides only 15 lumens but since its laser technology it provides the advantage to be always in focus. The iPod touch was chosen since it is one of the few devices that allows to present different content on the display and the TV-Out and additionally contains a gyroscope. We used the gyroscope and the accelerometer in the device to determine its orientation highly precise and with that be able to control the viewport of the projected peephole. The projected image is not corrected before it is projected, so that depending on the angle between the device and the projection surface, it can get distorted. This

Figure 6.13: Dynamic Peephole Prototype Device - The prototype consists of an 4th generation iPod Touch and a Microvision ShowWX Laser Projector.

problem could be overcome using the method developed by Dao et al. [56], which was created for a similar use case. The application for the iPod touch is based on the iOS UIMapKit, which already include many basic map functionalities.

6.2.3 Evaluation

To evaluate the created system we conducted two user studies. At first we consulted a group of experts to get general feedback about our system. In the second study we assessed the general performance of this spatial interaction technique.

6.2.3.1 Qualitative Expert User Study

To get first feedback on our interface design we conducted an informal expert study with 12 experts (10 male) on mobile interaction in unstructured interviews. The participants were selected from our University and all had at least 2 years of expertise in developing novel mobile solutions and applications. All participants had the possibility to use the prototype after a short introduction into the operation of the device. The participants all found the application overall easy to use and most were surprised how well the peephole metaphor worked with the device tracking the movements itself.

Furthermore they all agreed that the map exploration with this was a lot easier than on the small screen and that they were able to use their micro spatial memory to relocate POIs that they explored before as it has also been confirmed by Kaufmann and Ahlström [107].

The biggest problem that eight of the participants criticized was the way of performing the selection. The following comment summarizes the problem (male, 28) "when I want to select something, I move the cursor over the POI and then I have to look at the devices screen to search for the select button and when I press it, I already accidentally moved the crosshair a little bit and selected something different". This problem was observed for multiple users. Two participants stated that a double tap or something similar on every part of the devices display should replace the select button. One participant also stated that he wish to go even one step further and replace all the UI elements through gestures. For our next evaluation we decided to remove the select button and adapted the idea of a double tap on every part of the display. If a UI element is placed under the double tap position it will simply be ignored. This was well accepted beyond all participants and the selection was rated much easier. Most of the participants (9 out of 12) disagreed with removing all UI elements stating that they are easier to use and that there would be no need to remember gestures that in the end could be different for every application.

6.2.3.2 Quantitative User Study

The informal study was mainly meant to get an inside on how to design interfaces for a dynamic peephole display, using a mobile projector. Even though the experts consulted in our informal study were all positive about responsiveness and performance of the developed system we still have no insight on its actual performance. Therefore we conducted a second study with the aim to investigate the performance of the peephole pointing using a mobile projection interface.

While it has been shown that peephole pointing is a valid and natural interaction technique for AR applications on mobile devices [179], it hasn't been shown that in terms of task performance it actually has any advantages. Interestingly Kaufman and Ahlström didn't compare the performance of their system against other interaction techniques [106, 107]. To assess the task performance time we conducted a user study comparing the dynamic peephole pointing with a touch-based cursor control.

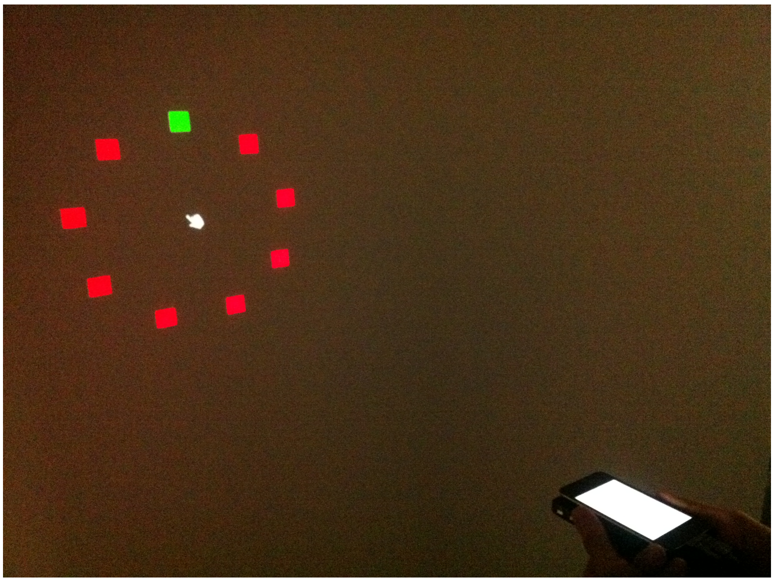


Figure 6.14: Dynamic Peephole User Study - The cyclical multi-direction pointing task that the users had to fullfill.

Task The user study adopted the cyclical multi-direction pointing task paradigm of ISO 9241-9 [98] in which the cursor was steered either using the before mentioned peephole interface or using the touch-screen of the iPod Touch. In the cyclical selection paradigm the subject is required to move the cursor to sequentially numbered targets which are equally spaced around the circumference of a circle. The task starts when the user moves the cursor to the topmost target and ends when the sequence is completed (the topmost target is selected again). The next item to select is always close to opposite clockwise around the circle. Selection was done by moving the cursor on top of the target. Selection was done by double tapping on the screen as a result of the qualitative user study.

For our setting we choose 9 targets which were altered in distance. We chose three different distance, the small circle had a diameter of 360 pixel (three quarters the projection size), the medium circle 720 pixel (one and a half times the projection size) and the large circle was 1440 pixel (three times the projection size). These sizes were deliberately chosen to emulate different size of content that the users would explore on their mobile devices, such as small images (small circle size), to large maps (large circle size).

For the touch-based technique the cursor is controlled using relative positioning as on a touch-pad. This decision was made since in pilot tests we found that users cannot estimate well where the finger initially touches the touch pad, such that absolute positioning was not deemed suitable. The mapping of the touch input to the velocity of the cursor followed a linear acceleration. The setup can be seen in Figure 6.14.

Participants & Procedure The study was conducted with 18 participants, 10 female, 8 male, ages 18-41 (average 26.5). The subjects were mostly undergraduates with varying degrees of technical background. 17 subjects were right-handed and one was left-handed. The study took including questionnaire approximately 50 minutes and the subjects were paid an incentive of 10 Euro for their participation.

Interaction techniques as well as circle sizes were counterbalanced between participants to ensure that no learning effects could occur. After a short introduction in which all participants had the chance to test the two interaction techniques, participants started either with the dynamic peephole interaction or the touch interaction and then conducted all three circle sizes. Participants were required to stand on a prior marked position in front of a white wall so that a constant size of the projection was ensured. Between every task the participants had the chance to pause. The cursor was in the beginning of each task placed in the centre of the circle. For each technique and circle size we conducted 10 trials. Overall this results in 1080 data points (18 users $\times$ 2 interaction techniques $\times$ 3 target sizes $\times$ 10 trials).

As for measures we took the task completion time starting from the moment the users pressed the start button and ending when the last target was selected. An error occurred when a user tried to select a target while the cursor is not placed on the current target.

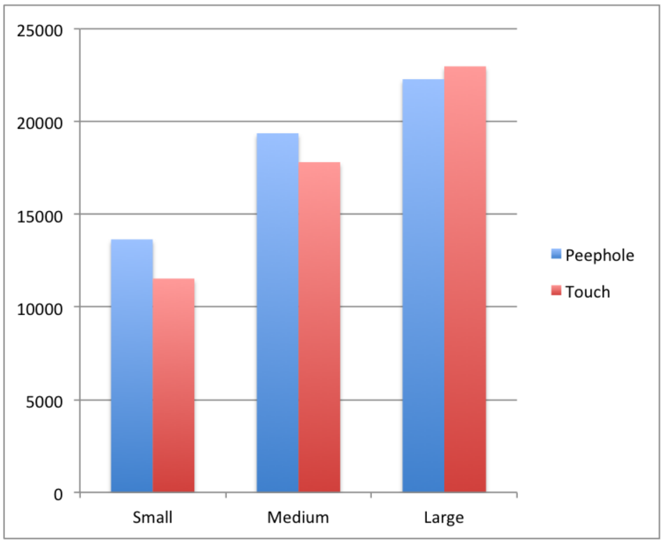


Figure 6.15: Dynamic Peephole User Study Results - Average task completion time of the study in ms.

Reuslts For the analysis of the task completion time we used a two way factorial ANOVA with interaction technique and circle size as factors. Both factors showed a significant effect ($p < 0.05$). For the small sized circles the touch based interaction (avg. 11.53s) was significantly faster ($p < 0.01$) compared to the peephole technique (avg. 13.64s). For the medium sized circles the touch based interaction (avg. 17.81s) again was significantly faster ($p < 0.05$) compared to the peephole technique (avg. 19.36s). For the large scale circle no significant difference could be found even though on average the peephole interaction (avg. 22.01s) was faster than the touch based interaction (avg. 22.96s). In terms of errors, overall participants tried to select wrongly for the peephole technique 48 times and for the touch based technique 53 times. The difference is negligible as both are around 10%.

6.2.3.3 Discussion

Surprisingly the touch-based interaction outperformed the peephole interaction. This might also explain why results on such studies have not been reported before (as mentioned above). A possible explanation for the results could be simply the familiarity of the users with such touch pad style interactions. All participants owned a laptop that had a touchpad, so all participants where familiar with this kind of interaction. Therefore they felt much more comfortable directly from the first trial. Furthermore the peephole interaction required to move the projection which might have led to confusion as users are not used to having a fully mobile display in current desktop settings. Nevertheless for the large circle size the participants on average were faster. This result could have been even more explicit for even larger diameters. This indicates an advantage when exploring large scale content such as maps or websites on a projector phone which also is supported by the findings of Hang et al. [82]. Since Kaufmann and Ahlström [107] showed that using a dynamic peephole interface for a mobile projected display can significantly enhance the users micro spatial memory compared to a static peephole, the dynamic peephole is a valid interaction technique when interacting with large scale content.

6.3 Summary

In this chapter we worked towards answering the research question of **What are suitable spatial interaction techniques for mobile projected displays**. For this we followed a two-fold approach, on the one side we investigated a technique that allowed switching of virtual workspaces and applications. On the other side we investigated the design and effectiveness of a dynamic peephole interface for mobile projected displays.

For our investigation of mobile virtual workspacese started with existing techniques for workspace management in the desktop environment with the view to translate them to mobile settings. We show that extra spatial awareness, that is possible in the mobile context, vastly enhances users' performance. We have presented mSpaces, a spatially aware prototype for virtual workspaces, which allows workspace switching by moving the phone to various physical locations, exploiting the users' micro spatial memory. To determine if a mobile projected display further enhances users' performance, we designed the m+p and pSpaces prototypes. In a final user study we compared the three techniques to show that projected display in fact enhances users' abilities to switch between workspaces. We concluded with design considerations to create, manipulate and manage virtual workspaces in the spatially-aware mobile environment.

Additionally we discussed a design for a distributed peephole interface for projector phones. This design concept is a first try to facilitate the integration of projector phones in every day-life task by minimizing the drawbacks that evolve from the distributed interface of a projector phone. Through the combination of sophisticated user interface elements of prior research we developed a design that has the capability to tackle common problems on such devices. We presented our exemplarily implementation of this prototype and the corresponding evaluation. In our comparison study surprisingly the touch-based interaction outperformed the peephole interaction. As the difference in task completion time was rather small and might be outweighed by the advantage of the micro spatial memory generation presented by Kaufmann and Ahlström [107]. Therefore we can conclude that the dynamic peephole can be a valuable interaction technique especially for large-scale content. Furthermore the idea of the dynamic peephole interface can also be extended to create a virtual projection sphere around the user, which would allow for CAVE like interactions, as it has been developed by Koskenranta et al. [113].

7

Windows on the World Revisited

In this chapter we want to introduce a system that incorporates the findings that have been made in this thesis into a holistic mobile projection system. The goal of this system is to provide evidence in how far an added projector can bring an advantage over current mobile devices. As discussed in the introduction the added value through the projection in current projector phones is rather small. In this chapter we describe a proximity aware mobile projector depth-camera system that allows seamless interaction independent of the distance between the user and the projection surface. This is achieved by switching interaction techniques depending on the distance between the user and the projected display. Furthermore, it allows to easily access different applications by making use of the user's micro spatial memory through kinesthetic cues (compare Chapter 2 and 6). The system allows the user to distribute his application windows in the world surrounding him.

The chapter is based on the following previously published papers: [55, 121, 126, 129].

7.1 Idea

In 1993, Feiner et al. introduced an AR approach that allows interaction with 2D application windows in the real world using a head-mounted display (HMD) [69]. This concept is a promising approach when applied to projector phones. Instead of just showing one window or application in the projection - which is the current practice

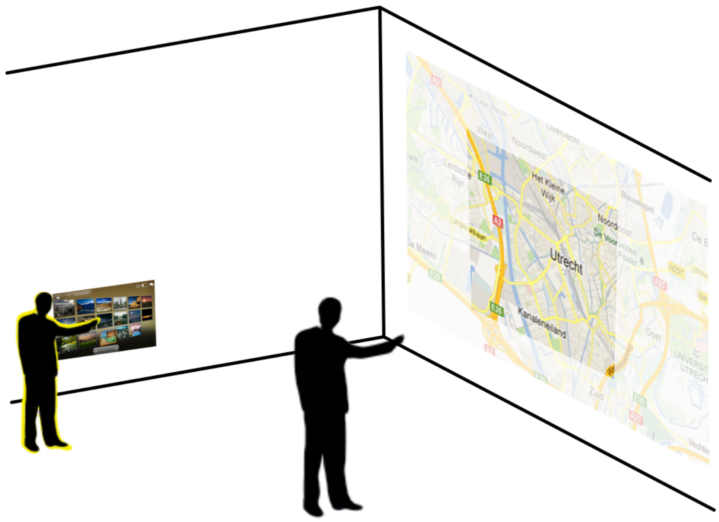


Figure 7.1: Distribution of applications around the user - A user interacting with two different applications at different positions with different interaction techniques. When being closer to the wall the user can touch the projection (yellow) and when being further away the user can use a dynamic peephole projection.

for projector phones - the user could arrange his applications in the real world around him. He could easily access the different applications by pointing the device at the position where the application was placed beforehand (compare Figure 7.1). Although a spatially aware mobile device would be needed for this approach, it would enable making use of the user's micro spatial memory through kinaesthetic cues. This would ostensibly alleviate some of the user's mental workload and it could increase performance when conducting tasks that need multiple applications.

That such a real world distribution of the applications increases task performance has been shown in chapter 6.1 . However, there we only analysed information perception in static applications and did not investigate how interaction will influence this behaviour. Or, to be more specific, which interaction techniques are actually suitable for such kind of novel device. With the applications distributed in the environment, they are easily recallable and accessible but since they might be located on different surfaces and different distances the interaction with them can become problematic. To the best of our knowledge, this problem has not yet been solved.

Even though a multi-user scenario might be the most beneficial use case for mobile projection, we are convinced that a single user can also avail from a large projected display. For example, in a situation where the user wants to access large-scale information while on the go or present something to a bigger audience, he still might want to interact with the projected content just by himself. Nevertheless the interaction scenarios described in this chapter allow for single- as well as multi-user purposes.

The concept of our proximity-aware system allows seamless interaction with mobile projected displays independent of the user's distance relative to them. The system is spatially aware and allows distributing the user's applications in the environment, and with that, the user can easily access the applications by pointing the device towards the direction they are located. Our system is based on prior findings of this thesis as well as other related research and integrates different techniques to allow a holistic experience that underpins the capabilities of interaction with mobile projected displays.

7.2 Proximity Based Interaction for mobile Projection

As discussed before, the main influencing factors when interacting with hand-held mobile projection devices is that size and position of the projected display are always coupled to the device and its motions. Even though there are approaches for decoupling the projection [42], the distance from the device to the projection surface would be a crucial aspect even in these cases. Depending on the distance, it is expected that people would be willing to show different content - e.g. if they are able to shield the projection they could project personal information [133] (compare Chapter 2). Furthermore, different distances would allow and even require the usage of different interaction techniques. For example, if the user is close enough to reach the projection with one of his extremities, direct interaction would be possible. This is of course not possible when the projection is further away, and with that, out of the user's reach.

Even though a variety of different interaction techniques for mobile projected displays exist today, they are mostly tailored to a specific use case as we have seen in the related work. To allow seamless interaction across all distances, we propose to adapt the interaction technique by choosing different techniques for different distances, and with that, make mobile projected displays proximity-aware (compare Figure 7.2). This idea is closely related to the approach of Vogel and Balakrishnan [206]. When moving the projected display closer to the user or further away from him, the interaction technique will be changed automatically to the most suited one.

Figure 7.2: Proximity-aware change of projected information - While in the public space (left) the projection would only show impersonal content such as maps, the privacy of information that would be available in the private space (right) would increase to information such as email.

7.2.1 Concept

The work of Feiner et al. [69] allowed to position 2D windows in the real world and accessing them by looking at the point where they have been placed. To realize this, they used a HMD, which resulted in the drawback that only people equipped with such a device could access the application (compare Chapter 2.1.2). For our system, we use a mobile projector instead of a HMD. It projects the different application windows onto nearby surfaces. This approach has the ability to make the applications accessible for multiple users, while only needing a small amount of instrumentation. To place an application, the user simply points the projector phone at a specific position in the real world and decides then, which application he wants to project at this point. This is a result of the design guidelines presented in Chapter 6.1. When he points away with the device, the application will be hidden. Pointing back at this position would result in projecting again the application, or parts of it, depending on the amount captured in the field of projection. This window definition procedure can be repeated for the desired amount of applications. The spatial arrangement around the user allows him to make use of his micro and macro spatial memory and to easily access different applications. This is a result of Chapter 6.1.

Two means of defining the positions of the applications around the users are possible: (1) absolute - and (2) relative positioning. An absolute tracking would have the ability to incorporate real world features into the application. Those features could help to reduce mental demand since they could act as mental aids. Here we also envision to exploit semantic meaning of real world objects. For example a paper-based calendar could be recognized and then the calendar application of the projector phone could be projected directly next to it or even as an AR overlay on top it. Through this, the system exploits the user's macro spatial memory as well. This might even increase task performance time. The real world features and objects would help by becoming landmarks, allowing to retrieve the position of applications even faster [142, 167].

Nevertheless, real mobility can only be preserved when using a relative tracking as in most cases no absolute tracking (e.g. a IR-based marker tracking) will be available. This would mean that the applications would be positioned relative to the user e.g. to the right in an angle of $90\,^{\circ}$. Every time the user points in this direction, the application would appear and it would not be spatially bounded to the environment. Relative positioning still makes use of the users micro spatial memory - kinaesthetic senses [168] as we employed it in chapter 6.1. In a scenario where the user is not moving, absolute- and relative positioning would be equal.

When using mobile projection, privacy is always an issue. To compensate for this using proxemics might be a helpful tool. As discussed in Chapter 2.3, even though the proxemic regions are not universal, they could be adapted to the region the user is living in. This could establish privacy aware interfaces for projector phones. Depending on the space the projection is located in, the system decides whether or not a specific content should be projected. Applications might not be made accessible, since they were placed in closer proximity where they, for example, have been shielded by the users body (compare Figure 7.2). Of course, it should be optional for the user to decide whether or not he wants applications to be made inaccessible. Furthermore, the integration of semantic privacy annotations for applications could be one possibility [128]. Even though it enables easy decision making which content should be projected or not, since most data is not annotated, this might be not easily realizable. Furthermore, proxemics can be used to change the interaction technique and with that give the ability to seamlessly interact with different content in different ways.

The system described in this chapter follows an approach that allows privacy based on interaction. Whenever the user wants to hide the current content, he simply points the device in to a different direction, which would then hide the current application.

7.2.2 Interaction Techniques

As discussed in the related work (Chapter 3), a variety of possible interaction techniques for mobile projection devices exist, but most of them have their limitations. The techniques that we selected for our system originates from the aforementioned related work and our own investigations in Chapter 6.

Direct manipulation of the projection is a promising approach independent whether using a foot [42] or a hand [84]. However, it is impossible if the projection is out of the reach of the extremities in a single-use case. In a multi-user scenario one person can hold the device while the other users can touch the large-scale projection. Nevertheless, it is not socially acceptable to touch every surface (e.g. wallpapers). Therefore different interaction techniques are needed as well. Another suitable technique is the spotlight metaphor utilizing a dynamic peephole. As discussed in Chapter 6 the dynamic peephole interaction has some very promising abilities especially to explore large-scale content. An example of a user interacting with a peephole interface can be seen in Figure 6.11. For interaction with the shown information a cross hair in the middle of the projection is projected on top of the application. By double tapping on the screen the user can select the object beneath the crosshair.

As we have shown in Chapter 6, the dynamic peephole interaction is not the fastest interaction technique in terms of task completion time, if the content does not exceed a certain size. Furthermore, it is not applicable when the user is so far away from the projection surface that the projected display is already taking most of the space that is suited for projecting on it. Therefore, we added the cursor control with the phone acting as a touchpad as discussed in Chapter 6. This allows the projection to remain stable, while the user could easily interact with it.

We choose these three interaction techniques as they are suitable for a variety of application and privacy scenarios. Furthermore, all three of them have been shown to be from certain convenience for interaction with mobile projected displays. Switching between the different interaction techniques can be either done automatically according to the proximity of the user to the projected display, or the user might assign a distinct interaction technique to a certain application (e.g. the dynamic peephole interaction to the map application).

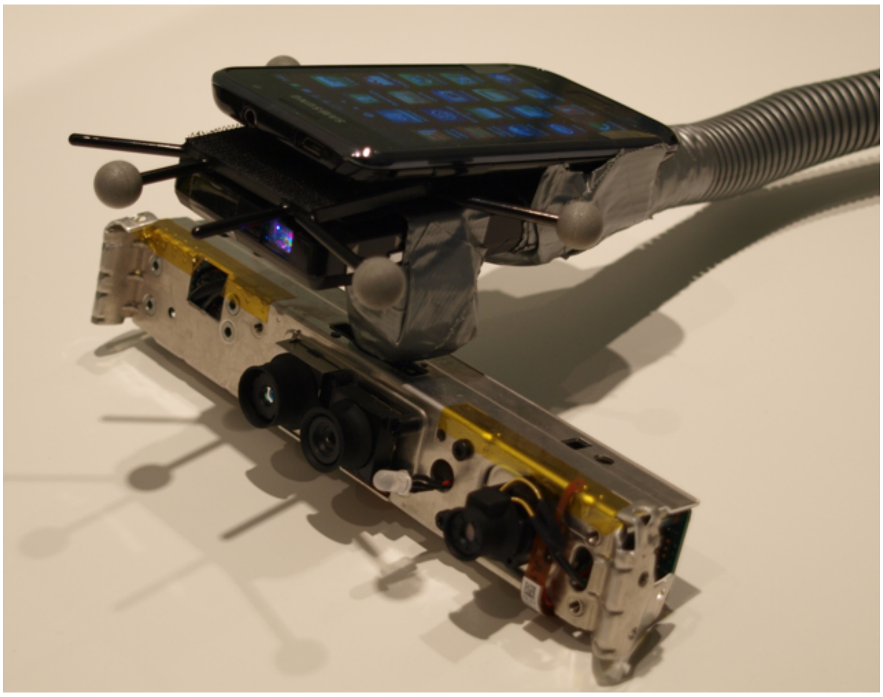


Figure 7.3: Hardware Prototype - Prototype consisting of a stripped down Kinect, a Microvision SHOWWX+ and a Samsung Galaxy S

7.3 Implementation

Albeit the availability of projector phones, these devices are very limited in their current form. Besides the fact that they are only able to mirror the image of the devices display there exist only models with an integrated LED-based projector. This is reasonable taking the lower battery consumption into account, but those projectors suffer from the need for manual focus, which would be problematic when projecting at different distances. Therefore, we created our own prototype using a laser projector (which has the advantage of always being in focus) that although not being fully mobile still allows for exploration of the presented system. We implemented the three interaction techniques mentioned above and the system switches automatically between these three. This is done based on the proximity and orientation of the device towards the projection. A short vibration of the device indicates the change of the interaction technique. If the distance is less than 90cm (arm reachable) the user is able to interact with the projection by touching it. Afterwards, either the spotlight metaphor or interacting through the touchpad of the device is available.

7.3.1 Hardware

On the hardware side, the prototype consists of a Microsoft Kinect for Windows, a Microvision ShowWX+ laser projector and a Samsung Galaxy S. The Samsung Galaxy S was chosen because it has a TV-Out which is connected to the laser projector. To reduce weight, the housing and pedestal of the Kinect were removed (compare Figure 7.3). All three devices were mounted onto each other in such a way that it is possible to hold them in one hand. Even though it needs a cable connection, with an overall weight of 376g it is still light enough to handle it with ease. The Kinect is connected to a MacBook Pro running Windows 7 by cable while the communication between the phone and the notebook is handled via WiFi. The image content of the projection was provided by the Samsung Galaxy S. The device's orientation is tracked in the environment using an external NaturalPoint OptiTrack optical tracking system. For the alignment of the projected display and the device's display we followed the findings of chapter 5.1. The display are aligned in a 70° angle. Even though in chapter 5.1 we used a more radical 60° angle, the 70° angle still ensures that the mobile devices display and the projection are in the same field of view, very similar to our SurfacePhone prototype in Chapter 5.2.

One might assume that due to physical limitations the integration of a Kinect-like depth sensor might not be possible, since it needs a certain distance between the infrared projector and the infrared detector. But this distance is just 9cm, for comparison a Samsung Galaxy Nexus is 12.5cm long, so that an integration would be possible. Of course the infrared projector and detector would need to undergo a certain miniaturization since they are approximately 2cm in height, but still it could be made possible with technological advances. The most problematic issue will be the power supply for such a device, but for example the Asus Xtion Sensor, which is comparable to the Kinect is powered only by USB. Partially the goal of a working mobile Kinect has also already been realized [159].

To achieve the spatial awareness of the prototype we used a Natural Point OptiTrack System, which required the prototype to have infrared reflective markers attached (compare Figure 7.3). The system tracks only the yaw, pitch and roll of the prototype in our environment. Currently the use of the OptiTrack system restricts the mobility of the prototype but it could be replaced by a sensor fusion of gyroscope and accelerometer as it has been done in Chapter 6.2. Since the Samsung Galaxy S that we use (because of its TV-Out capabilities) has no integrated gyroscope, sensor fusion was not possible.

7.3.2 Software

The prototype is implemented in C# using the Microsoft Kinect SDK. We implemented three simple Android applications for the phone - an image browser, a map application and a light weight email client that would only allow to read emails. They are all usable with all three interaction techniques. Those applications were chosen since they highlight the capabilities of a mobile projection device for single and multi-user scenarios. Furthermore they are frequently used in every-day life.

7.3.2.1 Application Positioning

The first step in the interaction with our proximity aware prototype is the positioning of the applications around the user. The user can select an application by pressing the

Menu button on the Samsung phone, and then point the prototype in the direction the application should be positioned. By double tapping on the phones screen the position of the application is saved. Therefore, the orientation of the prototype supplied by the OptiTrack system and the depth towards the projection surface measured by the Kinect are used to determine the position and size of projected display. A 3D model of the environment was not used to ensure that a replacement of the OptiTrack through e.g. an inertial measurement unit will be easily possible. Besides positioning also resizing of the application is possible by performing a pinch gesture on the device.

Switching between applications is easily possible by simply pointing the device into the direction, the switching is then done automatically. This is done in line with the way we evaluated in Chapter 6. If the user points into a direction where no application is located the projection turns dark. This is done so that it is possible to hide private information fast.

Figure 7.4: Processing pipeline for touch detection - Steps of the hand detection algorithm based on the depth image. From left to right: Canny Edge detection, Determine Contours after Dilation and Erosion, fill Contours, determine convex Hull.

7.3.2.2 Touch detection

The detection of touches on the projection is achieved using image processing algorithms based only on the depth data of the Kinect (compare Figure 7.5). For the processing of the depth data EmguCV [64], an OpenCV port for C#, was used. The detection of the touches is split into two steps, one is the detection of the hand and the other is the decision whether or not it is touching. The hand detection is done using the following operations (compare Figure 7.4):

1. Canny-edge detection to determine the hand

2. Dilation and Erosion (10×10 pixel box)

3. Determine contours (findContours with more then 1000pixel)

4. Fill contours

The touch detection is then done as follows:

1. Determine the convex hull and its center

2. Determine the edge of the finger

3. Determine the distance to the background (Ray from center to edge of hull extended by 20%)

4. Threshold of distance

All thresholds were experimentally determined through extended evaluation of the system. Even though this is a rather naive approach it is working at 60FPS (which is the maximum frame rate of the Kinect) on a common Notebook. For details about accuracy and mobile applicability compare the technical evaluation. Since simply every contour with more then 1000 pixel is taken into account, multi-touch is natively detected by the algorithm. Additionally, pointing with larger objects such as bottles is possible. The positions of the touches are then send to the Android phone through the WiFi connection using UDP packets guaranteeing a minimal lag.

7.3.2.3 Dynamic Peephole and Touchpad

For the dynamic peephole pointing, the OptiTrack System was used. By determining the orientation towards the projection surface, only the part of the application covered in the projection would be shown. The distance between the user and the projection surface - measured by the Kinect - is used to change the scaling of the application so that the shown part of the application would stay constant. In the middle of the projection a red crosshair is shown (compare Figure 2.8) and a double tap on the phone screen selects the object beneath it. This is implemented using the double tap listener from the Android API.

Figure 7.5: Touch Interaction - A user interacting with the projection by direct touch.

The Touchpad interaction is realized by a cursor that is controlled by the touch-screen of the Samsung Galaxy S. The mapping of the touch input to the velocity of the cursor in the projection is following a linear acceleration. A standard cursor arrow, as it is common in today's operating systems, was chosen to effectively distinguish the cursor control from the peephole metaphor. While for the peephole metaphor a red crosshair is laid over the application, the cursor symbol in the touchpad interaction obviously shows the changed affordance when switching between different interaction techniques.

7.4 Evaluation

To evaluate our system, we decided to conduct two studies to outline the performance of the system. First we conducted a technical evaluation to show the accuracy of the touch registration on the projection as well as a performance analysis of out touch recognition algorithm on a smartphone. Secondly we conducted a user experience evaluation. We did not focus on task performance as the system is in an early stage of development and we could have only tested a rather made up task. Therefore, we decided to only focus on qualitative measures.

7.4.1 Technical Evaluation

To evaluate the accuracy and the performance of our touch-detection algorithm we conducted a technical evaluation. The registration of touch events is guaranteed as long as the projection surface is flat or has only smooth curves and the Canny edge algorithm is able to extract the hand contour out of the depth image. When it failed, it always resulted from a flat angle between the hand and the surface while the finger was actually touching the surface. Hence it is not recommended to touch the surface directly, but it is enough to operate on some millimetres distance to it. This also fulfils social norms, e.g. not to touch the wallpaper.

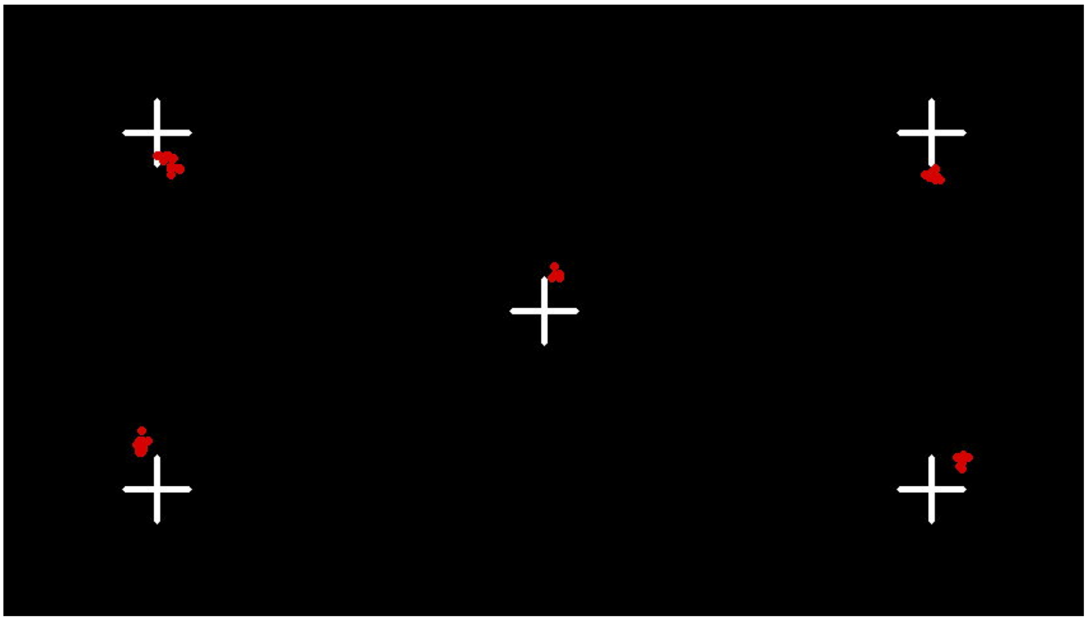


Figure 7.6: Technical evaluation of the touch detection - The technical evaluation application displayed a crosshair at five different positions, that the users then had to touch. The red dots indicate the registered touches in the evaluation.

We developed a distinct application for this evaluation shown in Figure 7.6. It displayed white cross-hair targets in each of the projection corners and one in the center. The pico projector resolution was 848×480 pixels and each cross-hair had a width and height of 50 pixels. The projection distance for this evaluation was set to 100 cm and the width of the projection image was 80 cm. Therefore a 1 cm line in the projection

Marker	Point Position	Avg. detected touch	SD [pixel]		Dev. [cm]	
			x	y	x	y
Upper left	(120,100)	(130.8,105.9)	7.3	7.7	1.5	0.9
Upper right	(728,100)	(729.2,117.7)	6.1	11.4	0.6	2.0
Down left	(120,380)	(101.0,358.6)	5.5	13.1	1.9	2.4
Down right	(728,380)	(750.5,371.9)	10.7	7.3	2.3	1.0
Center	(424,240)	(421.9,241.8)	5.2	4.4	0.5	0.4

Table 7.1: Results of the evaluation of touch registration

image consisted of approximately 11 pixels and a cross-hair had a size of about 4.7 cm.

The projector was aligned perpendicular to a plane projection surface and rested on a table. The participants were asked to click the five cross-hairs in arbitrary order, but such that in the end, each of them was clicked 20 times. In the meanwhile, the application recorded the 100 detected positions but did not display them during the evaluation to avoid side effects such as the users trying to compensate the registration inaccuracies. We acquired 10 students (4 female) from our university, aged 20 to 24 years (avg. age 22 years). This resulted in 1000 gathered touch-points.

Our findings are depicted in the table 7.2. The average SD over all markers is 7.0 pixels for x and 8.8 pixels for y. Hence, 95% of the detected touches have a distance of 14.0 pixels or less to the average in x direction and 17.6 pixels in y direction. To make sure that the user interface can be touched accurate enough, we suggest a button size of at least 38.0×37.0 pixels. We designed our applications accordingly. Of course the accuracy of such a system can be enhanced through a personal touch model as it has been employed in [211]. But still the accuracy is good enough, as long as the user interface is designed accordingly.

Since the fast adoption into mobile devices was one of the main drivers behind this system, we evaluated how the above described algorithm performs on a current smartphone. We implemented the same algorithm for OpenCV4Android [161] using the Android NDK. Since currently no driver is available that allows to use the Kinect with an Android device, we saved 3000 frames of Kinect depth data. This data is then read in native code and the frames are processed using the algorithm. On a Samsung Galaxy S3 the processing rate of our algorithm performs at ∼24fps (without any kind of visualization, just the touch processing). This illustrates the portability of the algorithm.

7.4.2 Qualitative Evaluation

As current prototypes only focus on very specific application scenarios (compare Chapter 3) and no holistic mobile projection system exist so far, we decide to not compare our system in a quantitative user study, but rather focus on qualitative feedback.

7.4.2.1 Method

To assess the participant's experience and perception of our system, we employed an adapted version of the Microsoft Production Reaction Cards [18]. Even though this method was originally designed to evaluate the emotional response and desirability of a software design, it has also been applied to product design [11]. Instead of presenting separate cards we presented the participants a sheet of paper with 26 antonym word pairs. The participants where requested to select five out of these 52 words that they found to describe their experience most accurately. Additionally we asked the users to complete a NASA TLX on a 5-point Likert scale for each of the interaction techniques.

7.4.2.2 Study Set-up

13 volunteers (5 female) aged between 21 and 35 ($\mu = 26.5$) were recruited from our research institution. All were right-handed and familiar with smartphones and touch-screen technologies. 7 participants were aware of the current development and had seen a mobile projector before, but only 3 had ever used one. One male participant had to be removed afterwards because of a technical problem during the experiment leaving 12 valid data sets.

The users were given the device and after a short explanation phase they were asked to use all three implemented applications (Map application, gallery application and email client) each with all three interaction techniques (dynamic peephole, touchpad and direct touch on the projection). They were free to choose the projection surface as well as the amount of time they spend with each technique on each application.

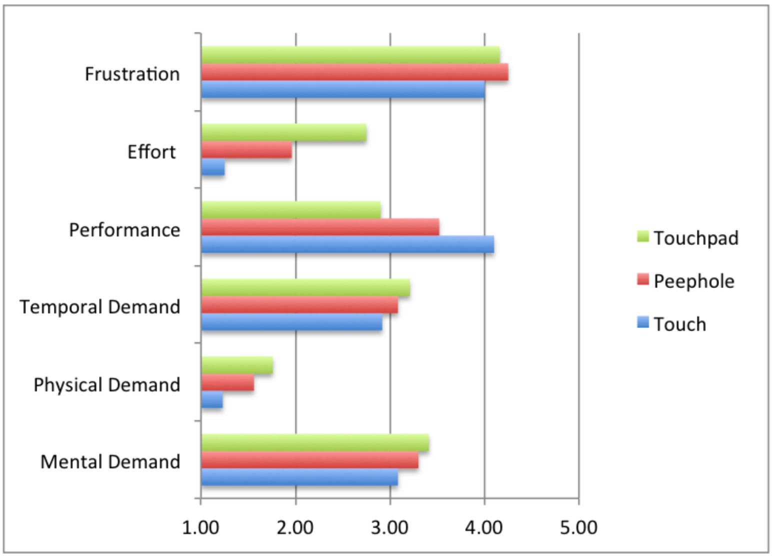


Figure 7.7: Results of the NASA TLX Questionnaire - Results of the questionnaire on a 5-point Likert scale (1 - Low to 5 - High). Statistically significant results were found for Effort and Performance ($p < 0.05$).

7.4.2.3 Results

Overall the Microsoft Reaction Cards show a clear evidence that the system was well accepted and highly appreciated by the participants. Only three negative terms and eighteen neutral were selected by all participants, compared to the 39 positive terms. The top six selected terms can be found in table 7.2.

The NASA TLX questions were analysed separately using non-parametric tests (k-related samples with Friedman Test Type). We found a significant difference ($p < 0.05$) for the following pairs: Performance: (direct touch > touchpad and direct touch > dynamic peephole) and Effort: (direct touch < touchpad and dynamic peephole < touchpad). The results can also be seen in Figure 7.7.

7.4.3 Discussion

Overall the direct touch was rated best, which is not surprising, taking into account that most people are familiar with this technique from e.g. tablet devices. But still

Top selected terms
Creative (10/12)
Exciting (10/12)
Convenient (8/12)
Effective (7/12)
Innovative (7/12)
Attractive (5/12)

Table 7.2: Results of the Microsoft Reaction Cards
The six most common terms selected by the paricipants to describe the system (n = 12)

this technique cannot be applied if a larger projection is desired e.g. in a presentation scenario. Even though all participants were familiar with touchscreen interaction, the touchpad interaction could not convince the participants. The dynamic peephole interaction was rated to be a very natural form of exploring large-scale information. This implies that for such applications, e.g. the implemented map application or a web browser, this technique is very useful. The results in terms of user experience emphasizes that the participants liked the system and found it convenient to use. When asked about the automatic switching, all participants except one stated that they liked the idea, especially when moving away from the projected display and with that out of reach. Three participants stated that they were a little bit confused when approaching the display, as they wanted to keep using the before used technique.

7.5 Summary

In this chapter we address the question of how we can leverage the full capabilities of mobile projected displays by incorporating spatial features. We presented a system that incorporates findings of this thesis in order to create a holistic experience for spatial interaction with mobile projected displays. We base the different interaction techniques on the proximity of the user towards the projected displays. Furthermore, the system allows to incorporate objects in the environment to exploit the micro - and macro spatial memory of the users. The system enables spatial interaction with mobile projected display in a more holistic way than ever before.

8

Conclusion & Future Work

In this last chapter we will summarize the contributions of this work. We will also line out possible directions for future work that is based on the results from this thesis but that go beyond its scope. Finally we will close this thesis with final remarks.

8.1 Summary

The goal of this thesis was to exploit the capabilities of mobile projected displays to a higher degree as before by incorporating spatial interaction techniques and let applications for such devices become spatially aware. Therefore, the research followed three different directions besides a categorization of existing work. In Chapter 2 we presented the theoretical background and we developed the concept of micro- and macro spatial memory that combines findings from the field of spatial cognition with physiological and working memory theories. This does not only enable us to judge spatial interaction techniques in a better way, but also allow fellow researchers to be more precise when investigating about spatial memory in the field of HCI. To build up a basis for more specific categorizations of projected displays, we developed a notion of such, based on their mobility in Chapter 3. Additionally, this chapter also presents existing interaction techniques for mobile projected displays derived from the related work.

Following this foundation we investigated the three different directions: Applications exploiting spatial features (Chapter 4); Analysis of spatial alignment of mobile projected displays (Chapter 5); and Interaction techniques exploiting spatial memory techniques (Chapter 6). In Chapter 7 we combined these three directions again and presented a system that allows for a novel, more holistic way of interacting with mobile projected displays.

For the investigation of spatially-aware applications (Chapter 4) we decided to tackle the field of AR as it represents the closest combination of the real world and the mobile projected display. We presented different interface designs for mobile projection-based AR applications taking hardware issues of future projector phones into account. The hardware design guidelines allow manufacturers to make educated decisions of how to design their future products to enable such spatially aware applications. Additionally we presented a system that eases guitar learning based on a mobile projected display.

Our analysis of the alignment of mobile physical - and projected displays revealed that current alignments of projector phones such as the Samsung Galaxy Beam [76] are not ideal to support the user. Our design guidelines suggest to align them in one field of view. That this can be effective was also shown with the SurfacePhone prototype, which was rated as very enjoyable to work with in our qualitative user studies.

As for spatial interaction techniques for mobile projected displays we presented two different approaches. On the one side the m+p-spaces, a technique that allows to switch between different workspaces and on the other side an in-depth investigation of the dynamic peephole metaphor applied to a mobile projected display. We were able to show where these technique have advantages compared to other techniques and where their limitations lie.

Lastly, we presented a system that combined the findings from the related work as well as from this thesis. The switch between different interaction techniques is based on the proximity of the user towards the projected displays. Furthermore, the system allows to incorporate objects in the environment to exploit the micro - and macro spatial memory of the users. The system enables spatial interaction with mobile projected display in a more holistic way than ever before.

8.1.1 Contributions

Here we want to outline the achieved contributions of this thesis with respect to their role as theoretical, technical or design contributions:

- **Theoretical contributions:** In Chapter 2 we presented existing approaches and knowledge from the field of spatial cognition. By classifying them into micro- and macro spatial memory (Compare Chapter 2) techniques we allow characterization of interaction techniques more fine-grained. Furthermore, we presented in Chapter 3, four classes of types of projected displays based on their mobility. Additionally, we developed a classification for possible interaction techniques with such. These six classes allow fellow researchers to classify their techniques as well as provide opportunities for future research. Additionally we developed a novel analysis technique for investigating the effects of spatial separation of mobile displays (as described in the user study in Chapter 5). By adopting mobile eye-tracking analysis we were capable of gaining a deeper insight in the behaviour of users when interacting with mobile displays. Furthermore, we drew first insight of user behaviour in such MMDEs.

- **Technical contributions:** As for technical contributions we presented two new methods for tracking direct touch interaction with a mobile projected display. The first one, presented in Chapter 7, is a depth camera based approach that utilizes an algorithm that is simple enough to achieve a high performance on current smartphones while still being highly accurate. The second approach that is applied for the SurfacePhone, presented in Chapter 5, is based on simple commodity hardware of current devices only relying on the camera and the accelerometer of an iPhone 5. Furthermore we present a sensor-fusion system for creating a dynamic peephole interface for mobile projected displays in Chapter 6. Additionally, we showed advanced technical solutions to create AR applications for mobile projected displays. These technologies allow a seamless fusion of the projected display and the environment around the users. With the SurfacePhone we also make an open source contribution to the community. The whole instructions including sources, building manual and the 3d printing file of the case can be found online. From this we hope to spark further investigations of the SurfacePhone concept.

- **Design contributions:** In terms of the design of interaction with projected displays we contributed on three different areas. Firstly, we investigated suited hardware designs that enable AR applications. These include the design of a guitar learning system, including different visualization and interaction techniques. Secondly, our investigations of different alignments of mobile displays contribute to a set of design guidelines for such combinations of displays. Our analysis show that current alignments do not cope with the users demands. Finally, we presented the design of suited interaction techniques for mobile projected displays that exploit the users spatial memory. Especially here the design of mobile virtual workspaces has been outlined in Chapter 6.1.

8.2 Future Work

Even though the findings and contributions we provide in this thesis open new opportunities but they also leave open challenges that can be addressed in future work. Our categorizations shows that on the one side only little research has been conducted in the field of *Self-actuated Projectors*. In terms of interaction techniques the areas of *Manipulation of the projection surface* and *Bi-Manual manipulation of Surface and Projector* still provide space for further research. Especially multi-user interaction has so far been only sporadically investigated, even though it presents a compelling use-case for mobile projectors. Additionally technical limitations prohibited in-the-wild investigations and corresponding audience reactions.

8.2.1 Interaction Techniques using Surface Manipulation

As discussed in Chapter 3 so far only three systems exist that employ an interaction technique that rely on *Manipulation of the projection surface* or *Bi-Manual manipulation of Surface and Projector*. As such techniques have a high degree of freedom in terms of dimensions that can be manipulated, they provide profound expressiveness for interaction. This does not have to be limited flexible surfaces, even rigid objects can still be rotated and moved closer or further away from the projection. One specific example for this could be an addition to the Map Torchlight [132, 188] prototype presented in Chapter 4. Instead of moving the projection unit over the map, the map could be moved in front of the projection unit. This could for example allow to recreate interaction techniques that have been applied to classic paper maps for centuries [180]. This not necessarily has to be an augmentation of an already existing map, but also could facilitate natural interaction with maps projected on a blank projection surface. In general such interaction techniques are very well suited for AR applications as they enable a very natural way of exploring the AR overlay by manipulating the object where it is created.

8.2.2 Multi-User Spatial Interaction

By today, only very few researchers investigated real multi-user interaction with mobile projected displays. The majority of this thesis only investigated single user settings, even though with the SurfacePhone we presented a multi-user, multi projection prototype investigation (compare Chapter 5.2). When multiple users interact with one or multiple mobile projected displays, spatial features of the environment will be of increased importance. Additionally, so far it has not been investigated how multiple users are supposed to interact with only one mobile projected display. Especially when multiple users are located at different distances relative to the projected display, it is unclear how they are supposed to manipulate it. If no direct interaction techniques are applied, not all users might be aware of the actions of the other users. Therefore we see a huge potential for spatial interaction techniques in such settings.

8.2.3 In-The-Wild Studies and Audience Reactions

So far technical limitations prohibited an *in-the-wild* [44, 176] research approach. As current pico projectors provide only limited brightness, they are not suitable for bright or sunlight settings. This so far has prohibited *in-the-wild* evaluations. Since pico projection technology is expected to drastically increase in terms of brightness these future mobile projectors will be suited for evaluation in the real world instead of relying on laboratory settings only. Such *in-the-wild* studies will also allow to measure the audience experience. As only little research has been conducted in this area [54, 74], sophisticated insight is so far missing. Compared to the displays on our current smartphones, such mobile projected displays are visible to passersby as well. This of course limits content that is suitable for projection and furthermore opens options for mischief. For example uninvolved passersby could become the victim of involuntary projections or get blinded. So far no social or governmental rules exist that prohibit such behaviour.

This might be especially dangerous if for example a car driver is blinded. It is unclear whether social rules will form, as they to some extent, have been established for loud-speakers in current smartphones.

Furthermore the rather small distribution of projector phones have made *Appstore-based* [88] research approaches impossible. Even though the Samsung Galaxy Beam [76] has been released, limited application and developer support have preluded a huge penetration of the market. Recently first information have been leaked that Samsung is preparing a successor of the Galaxy Beam that most certainly will have an increased projector brightness as well as better developer support. It will be interesting to see what kind of third party applications will emerge once such devices are more widely distributed.

8.2.4 Self-actuated Projectors

The field of *Self-actuated Projectors* so far has only been investigated scientifically by Scheible et al. [184] with their Display Drone. Those devices have the advantage that problems such as hand tremor or a perspectively distorted display can be neglected or easily solved. With the decrease in price of multi-copters such as the Parrot AR.Drone the exploration of this field will be of increased interest for researchers in the near future. Personal assistants that fly around the user could be made possible. They could be enabled to project personal displays in the vicinity of the user. This of course requires a high amount of spatial awareness of these devices to deliver content not only at the right time but also at the right place. Additionally no interaction techniques so far have been applied to such *Self-actuated Projectors*. Especially *Direct Interaction on the projection* will be very efficient. As the user does not need to carry the projection unit himself, the size of the projected display is not affected by the users position relative to the projected display. This will enable interaction with large-size mobile projected displays.

8.3 Closing Remarks

In this thesis we investigated different application and interaction scenarios for mobile projected displays that rely on spatial features to enable a better user experience. The holistic view of this thesis incorporating not only technical but also design guidelines for such mobile projected displays enables to exploit their possibility to a higher degree as before. Still a variety of open question exist as outlined above. Especially interesting will be the technological evolution of projection technology and whether the fast increase of recent years will continue. But even if pico projectors fail to attract attention, the developed concepts of this thesis still can be applied to different technologies. Besides steerable environmental projections, flexible displays might act as a replacement to the projection in certain cases. Still, mobile projection provides features that are unmatched by such displays. The availability of large-scale displays independent from the environmental factors as outlined in the example of the Chilean miners in the introduction, exemplifies the capabilities of such devices [65]. We should keep an open mind about where this development takes us next.

Acknowledgements

In the last years I had the pleasure to meet and work with a lot of insanely great people that supported me writing this thesis. Therefore I want to use this possibility to thank some of them in particular and I apologize to the ones I forgot.

First of all, I am deeply grateful for the all the freedom and support that **Antonio Krüger** granted me through all those years. Especially I would thank him for all the comments he provided on this thesis and related publications. Furthermore thanks to **Jonna Häkkilä** for not only providing helpful advice, support during the writing process but also for reviewing my thesis.

Furthermore I had the pleasure to collaborate with a lot of cool and bright people. I would like to thank therefore: **Jessica Cauchard**, **Sriram Subramanian** and the rest of the BIG group for having me there and letting me participate in their awesome work. Thanks as well to **Rod Murray-Smith** and his IDI group for a wonderful time in Glasgow. And of course a big thanks for the work conducted with **Christian Winkler** and **Enrico Rukzio.**

Maybe one of the luckiest things that happen to me was the inspiring and creative environment at DFKI. First of all thanks, to my office mates **Florian Daiber** and **Sven Gehring**! Furthermore I want to thank my past and current colleagues from IRL / DFKI (in alphabetical order) **Matthias Böhmer**, **Christina Di Valentin**, **Benedict Fehringer** (you are better at Fifa), **Ralf Jung**, **Gerrit Kahl**, **Frederic Kerber**, **Christian Lander**,

Petter Nurmi, **Denise Paradowski**, **Johannes Schöning**, **Lübomira Spassova**, and **Marco Speicher**. Especially I would also like to thank **Frederik Wiehr** and **Felix Kosmalla** for proof reading the thesis. I had a great time working with all of you!

Most certainly my deepest thank-you goes to my whole family. I want especially thank my parents **Antonius** and **Christa** for supporting my dreams through all these years. Thank you for always understanding and supporting me.

List of Figures

List of Tables

References

[1] W. Albert and T. Tullis. *Measuring the user experience: collecting, analyzing, and presenting usability metrics.* Newnes, 2013. 12

[2] J. Ängeslevä, I. Oakley, S. Hughes, and S. OModhrain. Body mnemonics portable device interaction design concept. In *Proceedings of UIST*, volume 3, pages 2–5, 2003. 28, 29

[3] D. R. Ankrum. Workplace ergonomics supplement - visual ergonomics in the office-use these guidelines for monitor placement and fighting to improve productivity and promote injury-free work. *Occupational Health and Safety*, 68(7):64–70, 1999. 99

[4] P. M. Associates. Fourth annual pico projector manufacturer survey, 2012. (Online; last accessed 15th of May 2014). 3

[5] R. T. Azuma et al. A survey of augmented reality. *Presence*, 6(4):355–385, 1997. 16

[6] S. Bachl, M. Tomitsch, K. Kappel, and T. Grechenig. The effects of personal displays and transfer techniques on collaboration strategies in multi-touch based multi-display environments. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-computer Interaction - Volume Part III*, INTERACT'11, pages 373–390, Berlin, Heidelberg, 2011. Springer-Verlag. 123

[7] A. D. Baddeley and G. J. Hitch. Developments in the concept of working memory. *Neuropsychology*, 8(4):485, 1994. 27

[8] G. Bailly, A. Oulasvirta, T. Kötzing, and S. Hoppe. Menuoptimizer: Interactive optimization of menu systems. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, pages 331–342, New York, NY, USA, 2013. ACM. 25

[9] T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, pages 121–130, New York, NY, USA, 2010. ACM. 31, 32

[10] J. A. Bargas-Avila and K. Hornbæk. Old wine in new bottles or novel challenges: A critical analysis of empirical studies of user experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2689–2698, New York, NY, USA, 2011. ACM. 12

[11] C. M. Barnum and L. A. Palmer. More than a feeling: Understanding the desirability factor in user experience. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '10, pages 4703–4716, New York, NY, USA, 2010. ACM. 191

[12] H. C. Bastian. The muscular sense; its nature and cortical localisation. *Brain*, 10(1):1–89, 1887. 28

[13] P. Baudisch, E. Cutrell, K. Hinckley, and R. Gruen. Mouse ether: accelerating the acquisition of targets across multi-monitor displays. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '04, pages 1379–1382, New York, NY, USA, 2004. ACM. 99

[14] P. Baudisch and R. Rosenholtz. Halo: a technique for visualizing off-screen objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 481–488, New York, NY, USA, 2003. ACM. 162, 164, 166, 168

[15] beaMover. beamover 75, 2014. (Online; last accessed 15th of May 2014)). 39

[16] P. Beardsley, J. Van Baar, R. Raskar, and C. Forlines. Interaction using a handheld projector. *IEEE Computer Graphics and Applications*, 25(1):39–43, 2005. 45, 58

[17] F. S. Bellezza. Mnemonic methods to enhance storage and retrieval. *Memory*, pages 345–380, 1996. 27

[18] J. Benedek and T. Miner. Microsoft product reaction cards. In *http://www.microsoft.com/usability/ uepostings/productreactioncards.doc*, 2002. (Online; last accessed 15th of February 2014). 191

[19] H. Benko and S. Feiner. Multi-monitor mouse. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1208–1211, New York, NY, USA, 2005. ACM. 109

[20] H. Benko and S. Feiner. Pointer warping in heterogeneous multi-monitor environments. In *Proceedings of Graphics Interface 2007*, GI '07, pages 111–117, New York, NY, USA, 2007. ACM. 99

[21] H. Benko, A. D. Wilson, and R. Balakrishnan. Sphere: Multi-touch interactions on a spherical display. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, UIST '08, pages 77–86, New York, NY, USA, 2008. ACM. 41

[22] X. Bi, S.-H. Bae, and R. Balakrishnan. Effects of interior bezels of tiled-monitor large displays on visual search, tunnel steering, and target selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 65–74, New York, NY, USA, 2010. ACM. 96, 100, 103

[23] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and magic lenses: the see-through interface. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 73–80, New York, NY, USA, 1993. ACM. 66

[24] O. Bimber and R. Raskar. Modern approaches to augmented reality. In *ACM SIGGRAPH 2006 Courses*, SIGGRAPH '06, New York, NY, USA, 2006. ACM. 16, 17, 18, 19

[25] G. Blasko, S. Feiner, and F. Coriand. Exploring interaction with a simulated wrist-worn projection display. In *ISWC '05: Proceedings of the Ninth IEEE International Symposium on Wearable Computers*, pages 2–9, Washington, DC, USA, 2005. IEEE Computer Society. 57

[26] M. Böhmer and G. Bauer. Exploiting the icon arrangement on mobile devices as information source for context-awareness. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, MobileHCI '10, pages 195–198, New York, NY, USA, 2010. ACM. 161

[27] M. Böhmer, C. Lander, S. Gehring, D. P. Brumby, and A. Krüger. Interrupted by a phone call: Exploring designs for lowering the impact of call notifications for smartphone users. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 3045–3054, New York, NY, USA, 2014. ACM. 150

[28] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch projector: mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2287–2296, New York, NY, USA, 2010. ACM. 5, 160

[29] S. Boring, M. Jurmu, and A. Butz. Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, OZCHI '09, pages 161–168, New York, NY, USA, 2009. ACM. 5, 160

[30] C. Both and J. Diamond. Augmented reality application about roman forum and satricum, 2013. (Online; last accessed 15th of June 2013). 85

[31] R. Bradbury. *Zen in the Art of Writing*. Bantam Books, 1992. i

[32] Brother. Wd-100g, 2012. (Online; last accessed 15th of May 2014). 18

[33] A.-M. Burns and M. M. Wanderley. Visual methods for the retrieval of guitarist fingering. In *Proceedings of the 2006 conference on New interfaces for musical expression*, NIME '06, pages 196–199, Paris, France, France, 2006. IRCAM — Centre Pompidou. 85

[34] A. Butz and A. Krüger. *Mensch-Maschine-Interaktion*. De Gruyter Oldenbourg, 2014. 33

[35] A. Butz, M. Schneider, and M. Spassova. Searchlight - a lightweight search function for pervasive environments. In *Pervasive*, pages 351–356, 2004. 39, 40

[36] O. Cakmakci, F. Bérard, and J. Coutaz. An augmented reality based learning assistant for electric bass guitar. In *Proc. of the 10th International Conference on Human-Computer Interaction, Crete, Greece*. Citeseer, 2003. 85, 87

[37] X. Cao and R. Balakrishnan. Interacting with dynamically defined information spaces using a handheld projector and a pen. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, UIST '06, pages 225–234, New York, NY, USA, 2006. ACM. 53, 54, 55, 142

[38] X. Cao, C. Forlines, and R. Balakrishnan. Multi-user interaction using handheld projectors. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, UIST '07, pages 43–52, New York, NY, USA, 2007. ACM. 46, 57

[39] X. Cao, J. J. Li, and R. Balakrishnan. Peephole pointing: modeling acquisition of dynamically revealed targets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1699–1708, New York, NY, USA, 2008. ACM. 56, 163

[40] S. K. Card and A. Henderson, Jr. A multiple, virtual-workspace interface to support user task switching. In *Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface*, CHI '87, pages 53–59, New York, NY, USA, 1987. ACM. 138, 139

[41] J. Cauchard, M. Löchtefeld, M. Fraser, A. Krüger, and S. Subramanian. M+pspaces: Virtual workspaces in the spatially-aware mobile environment. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pages 171–180, New York, NY, USA, 2012. ACM. vi, 58, 137

[42] J. R. Cauchard, M. Fraser, T. Han, and S. Subramanian. Steerable projection: exploring alignment in interactive mobile displays. *Personal Ubiquitous Comput.*, 16(1):27–37, Jan. 2012. 31, 54, 73, 114, 115, 179, 182

[43] J. R. Cauchard, M. Löchtefeld, P. Irani, J. Schoening, A. Krüger, M. Fraser, and S. Subramanian. Visual separation in mobile multi-display environments. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, pages 451–460, New York, NY, USA, 2011. ACM. vi, 95, 167

[44] A. Chamberlain, A. Crabtree, T. Rodden, M. Jones, and Y. Rogers. Research in the wild: Understanding 'in the wild' approaches to design and development. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 795–796, New York, NY, USA, 2012. ACM. 12, 199

[45] C.-H. Chen and Y.-H. Chien. Effects of rsvp display design on visual performance in accomplishing dual tasks with small screens. *International Journal of Design*, 1(1):27–35, 2007. 104

[46] N. Chen, F. Guimbretiere, M. Dixon, C. Lewis, and M. Agrawala. Navigation techniques for dual-display e-book readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1779–1788, New York, NY, USA, 2008. ACM. 101

[47] X. A. Chen, N. Marquardt, A. Tang, S. Boring, and S. Greenberg. Extending a mobile device's interaction space through body-centric interaction. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pages 151–160, New York, NY, USA, 2012. ACM. 28

[48] K. Cheng and P. Graham. Spatial memory: Properties and organization. In *Handbook of Spatial Cognition*, pages 137–155. American Psychological Association, 2013. 26

[49] A. D. Cheok, K. H. Goh, W. Liu, F. Farbiz, S. W. Fong, S. L. Teo, Y. Li, and X. Yang. Human pacman: A mobile, wide-area entertainment system based on physical, social, and ubiquitous computing. *Personal Ubiquitous Comput.*, 8(2):71–81, May 2004. 18

[50] S. Clinch. Smartphones and pervasive public displays. *Pervasive Computing, IEEE*, 12(1):92–95, Jan 2013. 3

[51] K. Corporation. Kyocera echo, 2013. (Online; last accessed 15th of June 2013). 2

[52] S. Corporation. Pj780ve high definition flash-camcorder, 2013. (Online; last accessed 15th of June 2013). 2

[53] L. G. Cowan and K. A. Li. Shadowpuppets: supporting collocated interaction with mobile projector phones using hand shadows. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2707–2716, New York, NY, USA, 2011. ACM. 7, 63, 115, 162, 163

[54] L. G. Cowan, N. Weibel, W. G. Griswold, L. R. Pina, and J. D. Hollan. Projector phone use: Practices and social implications. *Personal Ubiquitous Comput.*, 16(1):53–63, Jan. 2012. 199

[55] R. Dachselt, J. Häkkilä, M. Jones, M. Löchtefeld, M. Rohs, and E. Rukzio. Pico projectors: firefly or bright future? *interactions*, 19(2):24–29, 2012. vii, 177

[56] V. N. Dao, K. Hosoi, and M. Sugimoto. A semi-automatic realtime calibration technique for a handheld projector. In *VRST '07: Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, pages 43–46, New York, NY, USA, 2007. ACM. 44, 67, 165, 169

[57] R. Davis. Sketch understanding in design: Overview of work at the MIT AI lab. In *Sketch Understanding, Papers from the 2002 AAAI Spring Symposium*, pages 24–31, Stanford, California, March 25-27 2002. AAAI Press. 79

[58] G. Debord. *Guide psychogéographique de Paris.* Édité par le Bauhaus imaginiste, 1957. 24, 25

[59] S. Devices. Smartq u7h, 2013. (Online; last accessed 12th of August 2013). 2

[60] C. Dickie, J. Hart, R. Vertegaal, and A. Eiser. Lookpoint: an evaluation of eye input for hands-free switching of input devices between multiple computers. In *Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, OZCHI '06, pages 119–126, New York, NY, USA, 2006. ACM. 109

[61] L. Dittmar. Static and dynamic snapshots for goal localization in insects. *Commun Integr Biol*, 4:17–20, 2011. 26

[62] P. Dourish and S. D. Mainwaring. Ubicomp's colonial impulse. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 133–142, New York, NY, USA, 2012. ACM. 8

[63] S. Elrod, R. Bruce, R. Gold, D. Goldberg, F. Halasz, W. Janssen, D. Lee, K. McCall, E. Pedersen, K. Pier, J. Tang, and B. Welch. Liveboard: A large interactive display supporting group meetings, presentations, and remote collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, pages 599–607, New York, NY, USA, 1992. ACM. 4, 5

[64] Emgu. Emgu-cv, 2013. (Online; last accessed 15th of May 2014)). 186

[65] Engadget. Samsung galaxy beam projector smartphone helped to light up the lives of the chilean miners, 2010. (Online; last accessed 11th of May 2014). 4, 7, 201

[66] Engadget. Urrobot's robii hides a multitouch projector in his circus tent, 2011. (Online; last accessed 15th of June 2013). 3

[67] B. Ens, D. Ahlström, A. Cockburn, and P. Irani. Characterizing user performance with assisted direct off-screen pointing. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, pages 485–494, New York, NY, USA, 2011. ACM. 162

[68] P. Eslambolchilar and R. Murray-Smith. Tilt-based automatic zooming and scaling in mobile devices - a state-space implementation. In S. Brewster and M. Dunlop, editors, *Mobile Human-Computer Interaction - MobileHCI 2004*, volume 3160 of *Lecture Notes in Computer Science*, pages 120–131. Springer Berlin Heidelberg, 2004. 2

[69] S. Feiner, B. MacIntyre, M. Haupt, and E. Solomon. Windows on the world: 2d windows for 3d augmented reality. In *Proceedings of the 6th annual ACM symposium on User interface software and technology*, UIST '93, pages 145–155, New York, NY, USA, 1993. ACM. 177, 180

[70] G. W. Fitzmaurice. Situated information spaces and spatially aware palmtop computers. *Commun. ACM*, 36(7):39–49, July 1993. 33, 34, 140

[71] C. Forlines, C. Shen, D. Wigdor, and R. Balakrishnan. Exploring the effects of group size and display configuration on visual search. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, CSCW '06, pages 11–20, New York, NY, USA, 2006. ACM. 100, 104, 109

[72] T. Ganz. *Die Welt im Kasten: von der Camera obscura zur Audiovision*. Verlag Neue Zürcher Zeitung, 1994. 7

[73] A. Gomes, A. Nesbitt, and R. Vertegaal. Morephone: a study of actuated shape deformations for flexible thin-film smartphone notifications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 583–592, New York, NY, USA, 2013. ACM. 2

[74] A. Greaves, P. Akerman, E. Rukzio, K. Cheverst, and J. Hakkila. Exploring user reaction to personal projection when used in shared public places: a formative study. *CAM3SN Workshop at MobileHCI*, 2009. 199

[75] S. Greenberg, M. Boyle, and J. LaBerge. Pdas and shared public displays: Making personal information public, and public information personal. *Personal Technologies*, 3(1-2):54–64, 1999. 97

[76] S. Group. Samsung galaxy beam, 2012. (Online; last accessed 11th of August 2013). 2, 3, 4, 8, 55, 69, 115, 196, 200

[77] S. Group. Samsung galaxy s4, 2013. (Online; last accessed 15th of May 2014)). 1

[78] J. Grudin. Partitioning digital worlds: focal and peripheral awareness in multiple monitor use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 458–465, New York, NY, USA, 2001. ACM. 113

[79] Y. Guiard, M. Beaudouin-Lafon, J. Bastin, D. Pasveer, and S. Zhai. View Size and Pointing Difficulty in Multi-scale Navigation. In *Proc. AVI '04*, pages 117–124, New York, NY, USA, 2004. ACM. 33

[80] S. Gustafson, D. Bierwirth, and P. Baudisch. Imaginary interfaces: spatial interaction with empty hands and without visual feedback. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, UIST '10, pages 3–12, New York, NY, USA, 2010. ACM. 144, 150

[81] E. Hall. *The hidden dimension*. Doubleday Anchor Books. Doubleday, 1966. 31

[82] A. Hang, E. Rukzio, and A. Greaves. Projector phone: a study of using mobile phones with integrated projector for interaction with maps. In *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 207–216, New York, NY, USA, 2008. ACM. 45, 109, 167, 174

[83] T. L. Hankins, R. J. Silverman, and W. Hackmann. *Instruments and the Imagination*. Cambridge Univ Press, 1995. 5, 6

[84] C. Harrison, H. Benko, and A. D. Wilson. Omnitouch: wearable multitouch interaction everywhere. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, pages 441–450, New York, NY, USA, 2011. ACM. 47, 48, 53, 182

[85] C. Harrison, D. Tan, and D. Morris. Skinput: Appropriating the body as an input surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 453–462, New York, NY, USA, 2010. ACM. 48, 52, 53

[86] B. Hecht, M. Rohs, J. Schöning, and A. Krüger. Wikeye using magic lenses to explore georeferenced wikipedia content. In *Proceedings of the 3rd International Workshop on Pervasive Mobile Interaction Devices (PERMID*, 2007. 19

[87] D. A. Henderson, Jr. and S. Card. Rooms: the use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface. *ACM Trans. Graph.*, 5(3):211–243, July 1986. 139

[88] N. Henze and M. Pielot. App stores: external validity for mobile hci. *interactions*, 20(2):33–38, 2013. 12, 200

[89] K. Hinckley, M. Dixon, R. Sarin, F. Guimbretiere, and R. Balakrishnan. Codex: a dual screen tablet computer. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1933–1942, New York, NY, USA, 2009. ACM. 98, 100, 102, 116, 117

[90] K. Hinckley, G. Ramos, F. Guimbretiere, P. Baudisch, and M. Smith. Stitching: pen gestures that span multiple displays. In *Proceedings of the working conference on Advanced visual interfaces*, AVI '04, pages 23–31, New York, NY, USA, 2004. ACM. 101, 103

[91] K. Hosoi, V. N. Dao, A. Mori, and M. Sugimoto. Cogame: Manipulation using a handheld projector. In *ACM SIGGRAPH 2007 Emerging Technologies*, SIGGRAPH '07, New York, NY, USA, 2007. ACM. 51

[92] W. Hürst and T. Bilyalov. Dynamic versus static peephole navigation of vr panoramas on handheld devices. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*, MUM '10, pages 25:1–25:8, New York, NY, USA, 2010. ACM. 56, 163

[93] F. IDG. Augmented reality application about roman forum and satricum, 2009. (Online; aufgerufen 4-Mai-2009). 20

[94] A. Inc. Macbook pro retina 2013, 2013. (Online; last accessed 15th of August 2013). 1

[95] F. Inc. Fretlight, 2013. (Online; last accessed 15th of June 2013). 86

[96] G. Inc. Google glass, 2013. (Online; last accessed 15th of May 2014). 17, 18

[97] Y. Inc. Yamaha ez-eg, 2013. (Online; last accessed 15th of June 2013). 85

[98] ISO. *ISO 9241-9 Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices.* ISO copyright office, 2000. 171

[99] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon. Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera. In *Proceedings*

of the 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11, pages 559–568, New York, NY, USA, 2011. ACM. 46

[100] R. M. Johnson, J. van der Linden, and Y. Rogers. Musicjacket: the efficacy of real-time vibrotactile feedback for learning to play the violin. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '10, pages 3475–3480, New York, NY, USA, 2010. ACM. 86

[101] B. R. Jones, H. Benko, E. Ofek, and A. D. Wilson. Illumiroom: Peripheral projected illusions for interactive experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 869–878, New York, NY, USA, 2013. ACM. 118

[102] S. K. Kane, D. Avrahami, J. O. Wobbrock, B. Harrison, A. D. Rea, M. Philipose, and A. LaMarca. Bonfire: A nomadic system for hybrid laptop-tabletop interaction. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, pages 129–138, New York, NY, USA, 2009. ACM. 116

[103] C. K. Kantarjiev, A. Demers, R. Frederick, R. T. Krivacic, and M. Weiser. Experiences with x in a wireless environment. In *Mobile & Location-Independent Computing Symposium on Mobile & Location-Independent Computing Symposium*, MLCS, pages 11–11, Berkeley, CA, USA, 1993. USENIX Association. 4, 5

[104] T. Karitsuka and K. Sato. A wearable mixed reality with an on-board projector. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pages 321–, Washington, DC, USA, 2003. IEEE Computer Society. 47, 52

[105] H. Kato. Artoolkit. *http://www. hitl. washington. edu/artoolkit/*, 1999. 20

[106] B. Kaufmann and D. Ahlström. Revisiting peephole pointing: A study of target acquisition with a handheld projector. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pages 211–220, New York, NY, USA, 2012. ACM. 57, 171

[107] B. Kaufmann and D. Ahlström. Studying spatial memory and map navigation performance on projector phones with peephole interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3173–3176, New York, NY, USA, 2013. ACM. 57, 170, 171, 174, 175

[108] Keecker. Projection robot, 2014. (Online; last accessed 15th of May 2014). 49

[109] C. Kerdvibulvech and H. Saito. Vision-based guitarist fingering tracking using a bayesian classifier and particle filters. In *Proceedings of the 2nd Pacific Rim conference on Advances in image and video technology*, PSIVT'07, pages 625–638, Berlin, Heidelberg, 2007. Springer-Verlag. 85

[110] M. Khalilbeigi, R. Lissermann, W. Kleine, and J. Steimle. Foldme: interacting with double-sided foldable displays. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, TEI '12, pages 33–40, New York, NY, USA, 2012. ACM. 2

[111] A. Kircher. *Ars Magna Lucis et umbrae*. Amsterdam, 1646. 5

[112] D. Kirsh and P. Maglio. On distinguishing epistemic from pragmatic action. *Cognitive science*, 18(4):513–549, 1994. 112

[113] O. Koskenranta, A. Colley, and J. Häkkilä. Portable cave using a mobile projector. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, UbiComp '13 Adjunct, pages 39–42, New York, NY, USA, 2013. ACM. 175

[114] D. M. Krum, E. A. Suma, and M. Bolas. Augmented reality using personal projection and retroreflection. *Personal and Ubiquitous Computing*, 16(1):17–26, 2012. 59

[115] B. Lahey, A. Girouard, W. Burleson, and R. Vertegaal. Paperphone: understanding the use of bend gestures in mobile devices with flexible electronic paper displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1303–1312, New York, NY, USA, 2011. ACM. 2

[116] J. R. Levin, M. E. Levin, L. D. Glasman, and M. B. Nordwall. Mnemonic vocabulary instruction: Additional effectiveness evidence. *Contemporary Educational Psychology*, 17(2):156–174, 1992. 27, 28

[117] J. P. Lewis, R. Rosenholtz, N. Fong, and U. Neumann. Visualids: automatic distinctive icons for desktop interfaces. *ACM Trans. Graph.*, 23(3):416–423, Aug. 2004. 153

[118] Lewis, James R. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. Technical 54.786, Human Factors Group. 129

[119] F. C. Y. Li, D. Dearman, and K. N. Truong. Virtual shelves: interactions with orientation aware devices. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, UIST '09, pages 125–128, New York, NY, USA, 2009. ACM. 29, 144, 150

[120] H. Liu, X. Xie, W.-Y. Ma, and H.-J. Zhang. Automatic browsing of large pictures on mobile devices. In *Proceedings of the eleventh ACM international conference on Multimedia*, MULTIMEDIA '03, pages 148–155, New York, NY, USA, 2003. ACM. 2

[121] M. Löchtefeld and M. Braun. Towards a mobile everywhere display based on a depth-camera projector unit. In *MobiVis - Mobile Vision and HCI Workshop at MobileHCI'12*, 2012. vii, 177

[122] M. Löchtefeld, S. Gehring, R. Jung, and A. Krüger. guitar: Supporting guitar learning through mobile projection. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 1447–1452, New York, NY, USA, 2011. ACM. vi, 42, 65

[123] M. Löchtefeld, S. Gehring, R. Jung, and A. Krüger. Using mobile projection to support guitar learning. In L. Dickmann, G. Volkmann, R. Malaka, S. Boll, A. Krüger, and P. Olivier, editors, *Smart Graphics*, volume 6815 of *Lecture Notes in Computer Science*, pages 103–114. Springer Berlin Heidelberg, 2011. vi, 42, 65

[124] M. Löchtefeld, S. Gehring, and A. Krüger. Distributed dynamic peephole interfaces for mobile map navigation on projector phones. In *Workshop on Distributed User Interfaces at CHI 11*. 2011. vii, 137

[125] M. Löchtefeld, S. Gehring, and A. Krüger. Distributed user interfaces for projector phones. In *Distributed User Interfaces*, pages 113–123. Springer London, 2011. vii, 137

[126] M. Löchtefeld, S. Gehring, and A. Krüger. Proxemic interaction for projector phones. In *Workshop on Mobile and Personal Projection MP2 at CHI 2011*, 2011. vii, 177

[127] M. Löchtefeld, S. Gehring, J. Schöning, and A. Krüger. Pinwi: Pedestrian indoor navigation without infrastructure. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, NordiCHI '10, pages 731–734, New York, NY, USA, 2010. ACM. 75

[128] M. Löchtefeld, S. Gehring, J. Schöning, and A. Krüger. Shelftorchlight: augmenting a shelf using a camera projector unit. In *Proc. 1st Workshop on Personal Projection via Mobile and Wearable Pico Projectors*, pages 20–23. Citeseer, 2010. v, 59, 65, 182

[129] M. Löchtefeld and A. Krüger. Evaluation of an in-call interaction technique for projector phones. 2013. vii, 55, 177

[130] M. Löchtefeld, A. Krüger, and M. Rohs. Mobile projection interfaces for augmented reality applications. In *Handbook of Augmented Reality*, pages 177–197. Springer New York, 2011. vi, 65

[131] M. Löchtefeld, J. Schöning, M. Rohs, and A. Krüger. Littleprojectedplanet: An augmented reality game for camera projector phones. *Proceedings of the Workshop on Mobile Interaction With the Real World (MIRW) at MobileHCI'09*, 2009. v, 59, 65

[132] M. Löchtefeld, J. Schöning, M. Rohs, and A. Krüger. Marauders light: replacing the wand with a mobile camera projector unit. In *Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia*, MUM '09, pages 19:1–19:4, New York, NY, USA, 2009. ACM. v, 59, 65, 198

[133] M. Löchtefeld, J. Schöning, M. Rohs, and A. Krüger. Projector phones: a new class of interfaces for augmented reality. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 2(3):1–14, 2010. v, 31, 59, 65, 179

[134] K. Lyons, T. Pering, B. Rosario, S. Sud, and R. Want. Multi-display composition: Supporting display sharing for collocated mobile devices. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part I*, INTERACT '09, pages 758–771, Berlin, Heidelberg, 2009. Springer-Verlag. 98

[135] S. Malacria, E. Lecolinet, and Y. Guiard. Clutch-free panning and integrated pan-zoom control on touch-sensitive surfaces: the cyclostar approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2615–2624, New York, NY, USA, 2010. ACM. 2

[136] R. L. Mandryk, M. E. Rodgers, and K. M. Inkpen. Sticky widgets: pseudo-haptic widget enhancements for multi-monitor displays. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1621–1624, New York, NY, USA, 2005. ACM. 96, 99

[137] D. Martinez-Plasencia, E. Joyce, and S. Subramanian. Mistable: Reach-through personal screens for tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, New York, NY, USA, 2014. ACM. 8

[138] T. Matsumoto and S. Hashimoto. Pileus internet umbrella: Tangible mobile interface of a lovely umbrella. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, TEI '09, pages 41–42, New York, NY, USA, 2009. ACM. 42

[139] T. Matsumoto, D. Horiguchi, S. Nakashima, and N. Okude. Z-agon: mobile multi-display browser cube. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '06, pages 351–356, New York, NY, USA, 2006. ACM. 100

[140] T. Matthews, M. Czerwinski, G. Robertson, and D. Tan. Clipping lists and change borders: improving multitasking efficiency with peripheral information design. In

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06, pages 989–998, New York, NY, USA, 2006. ACM. 162

[141] D. C. McFarlane and S. M. Wilder. Interactive dirt: Increasing mobile work performance with a wearable projector-camera system. In *Proceedings of the 11th International Conference on Ubiquitous Computing*, Ubicomp '09, pages 205–214, New York, NY, USA, 2009. ACM. 54

[142] T. McNamara, J. Sluzenski, and B. Rump. 2.11 - human spatial memory and navigation. In E. in Chief:John H. Byrne, editor, *Learning and Memory: A Comprehensive Reference*, pages 157 – 178. Academic Press, Oxford, 2008. 181

[143] T. P. McNamara. Spatial memory: Place learning, piloting, and route knowledge. In *Handbook of Spatial Cognition*, pages 173–191. American Psychological Association, 2013. 23, 25

[144] D. Merrill, J. Kalanithi, and P. Maes. Siftables: towards sensor network user interfaces. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, TEI '07, pages 75–78, New York, NY, USA, 2007. ACM. 98

[145] P. Milgram, D. Drascic, J. Grodski, A. Restogi, S. Zhai, and C. Zhou. Merging real and virtual worlds. In *Proceedings of IMAGINA*, volume 95, pages 218–230, 1995. 16

[146] P. Mistry, P. Maes, and L. Chang. Wuw - wear ur world: a wearable gestural interface. In *CHI EA '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 4111–4116, New York, NY, USA, 2009. ACM. 48, 62, 71

[147] D. Modjeska. Navigation in electronic worlds: Research review. Technical report, Computer Systems Research Group, University of Toronto, 1997. 77

[148] B. Moggridge and G. C. Smith. *Designing interactions*, volume 17. MIT press Cambridge, 2007. 12

[149] D. Molyneaux, S. Izadi, D. Kim, O. Hilliges, S. Hodges, X. Cao, A. Butler, and H. Gellersen. Interactive environment-aware handheld projectors for pervasive computing spaces. In J. Kay, P. Lukowicz, H. Tokuda, P. Olivier, and A. Krger, editors, *Pervasive Computing*, volume 7319 of *Lecture Notes in Computer Science*, pages 197–215. Springer Berlin Heidelberg, 2012. 40, 44, 45, 46, 52, 53, 62, 63

[150] D. R. Montello. Scale and multiple psychologies of space. In *Spatial information theory a theoretical basis for gis*, pages 312–321. Springer, 1993. 20, 21, 23, 26, 32

[151] D. R. Montello, D. Waller, M. Hegarty, and A. E. Richardson. Spatial memory of real environments, virtual environments, and maps. *Human spatial memory: Remembering where*, pages 251–285, 2004. 21

[152] Y. Motokawa and H. Saito. Support system for guitar playing using augmented reality display. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '06, pages 243–244, Washington, DC, USA, 2006. IEEE Computer Society. 85, 87

[153] H. Müller, J. Gove, and J. Webb. Understanding tablet use: a multi-method exploration. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*, MobileHCI '12, pages 1–10, New York, NY, USA, 2012. ACM. 1

[154] J. Müller, F. Alt, and D. Michelis. *Pervasive Advertising*. Springer, 2011. 3

[155] M. A. Nacenta, S. Sallam, B. Champoux, S. Subramanian, and C. Gutwin. Perspective cursor: perspective-based interaction for multi-display environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 289–298, New York, NY, USA, 2006. ACM. 99

[156] nchu eucl. "third paradigm" computing — ubiquitous computing, 2011. (Online; last accessed 11th of May 2014). 5

[157] K. Ng. Interactive feedbacks with visualisation and sonification for technology-enhanced learning for music performance. In *Proceedings of the 26th annual ACM international conference on Design of communication*, SIGDOC '08, pages 281–282, New York, NY, USA, 2008. ACM. 86

[158] T. Ni, A. K. Karlson, and D. Wigdor. Anatonme: Facilitating doctor-patient communication using a projection-based handheld device. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 3333–3342, New York, NY, USA, 2011. ACM. 58, 59

[159] U. of Bristol. Mobile kinect, 2013. (Online; last accessed 15th of May 2014)). 185

[160] J. O'keefe and L. Nadel. *The hippocampus as a cognitive map*, volume 3. Clarendon Press Oxford, 1978. 28

[161] OpenCV. Opencv4android, 2013. (Online; last accessed 15th of May 2014)). 190

[162] S. Ota, Y. Takegawa, T. Terada, and M. Tsukamoto. A method for wearable projector selection that considers the viewability of projected images. *Comput. Entertain.*, 8(3):17:1–17:16, Dec. 2010. 48

[163] C. Pinhanez, R. Kjeldsen, A. Levas, G. Pingali, M. Podlaseck, and N. Sukaviriya. Applications of steerable projector-camera systems. In *In Proceedings of the IEEE International Workshop on Projector-Camera Systems at ICCV 2003*. IEEE Computer Society Press, 2003. 40

[164] C. S. Pinhanez. The everywhere displays projector: A device to create ubiquitous graphical interfaces. In *Proceedings of the 3rd international conference on Ubiquitous Computing*, UbiComp '01, pages 315–331, London, UK, UK, 2001. Springer-Verlag. 39, 40, 44

[165] T. Pouli and S. Subramanian. Mobile projectors versus mobile displays: An assessment of task performance. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '12, pages 103–110, New York, NY, USA, 2012. ACM. 55

[166] J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey. *Human-computer interaction*. Addison-Wesley Longman Ltd., 1994. 12

[167] C. C. Presson and D. R. Montello. Points of reference in spatial cognition: Stalking the elusive landmark. *British Journal of Developmental Psychology*, 6(4):378–381, 1988. 23, 24, 181

[168] U. Proske and S. C. Gandevia. The kinaesthetic senses. *The Journal of physiology*, 587(17):4139–4146, 2009. 27, 29, 181

[169] R. Rädle, H.-C. Jetter, J. Müller, and H. Reiterer. Bigger is not always better: Display size, performance, and task load during peephole map navigation. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 4127–4136, New York, NY, USA, 2014. ACM. 34, 35

[170] S. Rapp. Spotlight navigation: a pioneering user interface for mobile projection. *Ubiprojection, Workshop on Personal Projection in conjuction with Pervasive 2010*, 2010. 45, 56, 164, 165, 166

[171] S. Rapp, G. Michelitsch, M. Osen, J. Williams, M. Barbisch, R. Bohan, Z. Valsan, and M. Emele. Spotlight navigation: interaction with a handheld projection device. *advances in pervasive computing: a collection of contributions presented at PERVASIVE*, pages 397–400, 2004. 45, 56, 164, 165

[172] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. ilamps: geometrically aware and self-configuring projectors. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Courses*, page 7, New York, NY, USA, 2006. ACM. 42, 43, 45

[173] M. Ringel. When one isn't enough: an analysis of virtual desktop usage strategies and their implications for design. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, pages 762–763, New York, NY, USA, 2003. ACM. 139, 140, 143, 151

[174] G. Robertson, M. Czerwinski, K. Larson, D. C. Robbins, D. Thiel, and M. van Dantzich. Data mountain: using spatial memory for document management. In *Proceedings of the 11th annual ACM symposium on User interface software and technology*, UIST '98, pages 153–162, New York, NY, USA, 1998. ACM. 25

[175] S. Robinson, P. Eslambolchilar, and M. Jones. Sweep-shake: finding digital resources in physical environments. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '09, pages 12:1–12:10, New York, NY, USA, 2009. ACM. 161

[176] Y. Rogers. Hci theory: classical, modern, and contemporary. *Synthesis Lectures on Human-Centered Informatics*, 5(2):1–129, 2012. 12, 199

[177] M. Rohs. Visual code widgets for marker-based interaction. In *Distributed Computing Systems Workshops, 2005. 25th IEEE International Conference on*, pages 506–513. IEEE, 2005. 78

[178] M. Rohs, J. Schöning, A. Krüger, and B. Hecht. Towards real-time markerless tracking of magic lenses on paper maps. *Adjunct Proceedings of the 5th Intl. Conference on Pervasive Computing (Pervasive), Late Breaking Results*, pages 69–72, 2007. 76

[179] M. Rohs, J. Schöning, M. Raubal, G. Essl, and A. Krüger. Map navigation with mobile devices: virtual versus physical movement with and without visual context. In *Proceedings of the 9th international conference on Multimodal interfaces*, ICMI '07, pages 146–153, New York, NY, USA, 2007. ACM. 56, 75, 170

[180] R. E. Roth. *Interacting with Maps: The science and practice of cartographic interaction.* PhD thesis, The Pennsylvania State University, 2011. 199

[181] A. Roudaut, A. Karnik, M. Löchtefeld, and S. Subramanian. Morphees: toward high "shape resolution" in self-actuated flexible mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 593–602, New York, NY, USA, 2013. ACM. 2, 60

[182] E. Rukzio, P. Holleis, and H. Gellersen. Personal projectors for pervasive computing. *IEEE Pervasive Computing*, 11(2):30–37, Apr. 2012. 51

[183] J. Scheible, A. Hoth, J. Saal, and H. Su. Displaydrone: A flying robot based interactive display. In *Proceedings of the 2Nd ACM International Symposium on Pervasive Displays*, PerDis '13, pages 49–54, New York, NY, USA, 2013. ACM. 50

[184] J. Scheible, A. Hoth, J. Saal, and H. Su. Displaydrone: A flying robot based interactive display. In *Proceedings of the 2Nd ACM International Symposium on Pervasive Displays*, PerDis '13, pages 49–54, New York, NY, USA, 2013. ACM. 200

[185] J. Scheible and T. Ojala. Mobispray: Mobile phone as virtual spray can for painting big anytime anywhere on anything. In *ACM SIGGRAPH 2009 Art Gallery*, SIGGRAPH '09, pages 5:1–5:10, New York, NY, USA, 2009. ACM. 43

[186] D. Schmidt, D. Molyneaux, and X. Cao. Picontrol: Using a handheld projector for direct control of physical devices through visible light. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, pages 379–388, New York, NY, USA, 2012. ACM. 51

[187] S. Schneegass, F. Alt, J. Scheible, S. Albrecht, and H. Su. Midair displays: Exploring the concept of free-floating public displays. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '14, New York, NY, USA, 2014. ACM. 50

[188] J. Schöning, M. Rohs, S. Kratz, M. Löchtefeld, and A. Krüger. Map torchlight: a mobile augmented reality camera projector unit. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, pages 3841–3846, New York, NY, USA, 2009. ACM. v, 59, 65, 198

[189] R. Shilkrot, S. Hunter, and P. Maes. Pocomo: Projected collaboration using mobile devices. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, pages 333–336, New York, NY, USA, 2011. ACM. 58

[190] B. Shirmohammadi and C. J. Taylor. Self-localizing smart camera networks. *ACM Trans. Sen. Netw.*, 8(2):11:1–11:24, Mar. 2012. 122

[191] B. Shneiderman. Human-computer interaction. chapter Direct manipulation: A step beyond programming languages, pages 461–467. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987. 33

[192] C. Stahl, J. Baus, B. Brandherm, M. Schmitz, and T. Schwartz. Navigational- and shopping assistance on the basis of user interactions in intelligent environments. In *Proceedings of the IEE International Workshop on Intelligent Environments (IE)*, University of Essex, Colchester, UK, 2005. 22

[193] J. Steimle, A. Jordt, and P. Maes. Flexpad: Highly flexible bending interactions for projected handheld displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 237–246, New York, NY, USA, 2013. ACM. 60

[194] S. Strachan, R. Murray-Smith, and S. O'Modhrain. Bodyspace: Inferring body pose for natural control of a music player. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '07, pages 2001–2006, New York, NY, USA, 2007. ACM. 28

[195] I. C. Study. Growing the digital signage economy beyond 7 billion dollar annually, 2011. (Online; last accessed 11th of May 2014). 3

[196] R. E. Su and B. P. Bailey. Put them where? towards guidelines for positioning large displays in interactive workspaces. In *Proceedings of the 2005 IFIP TC13 international conference on Human-Computer Interaction*, INTERACT'05, pages 337–349, Berlin, Heidelberg, 2005. Springer-Verlag. 96, 99, 101, 109, 114

[197] D. S. Tan and M. Czerwinski. Effects of visual separation and physical discontinuities when distributing information across multiple displays. In *Proc. Interact03*, pages 252–255, 2003. 96, 99, 100, 104, 109

[198] D. S. Tan, R. Pausch, J. K. Stefanucci, and D. R. Proffitt. Kinesthetic cues aid spatial memory. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '02, pages 806–807, New York, NY, USA, 2002. ACM. 30

[199] A. P. Tarun, P. Wang, A. Girouard, P. Strohmeier, D. Reilly, and R. Vertegaal. Papertab: an electronic paper computer with multiple large flexible electrophoretic displays. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 3131–3134, New York, NY, USA, 2013. ACM. 2

[200] A. Technologies. P300 pico projector, 2014. (Online; last accessed 15th of May 2014). 10

[201] H. Tobita, S. Maruyama, and T. Kuzi. Floating avatar: Telepresence system using blimps for communication and entertainment. In *CHI '11 Extended Abstracts on*

Human Factors in Computing Systems, CHI EA '11, pages 541–550, New York, NY, USA, 2011. ACM. 50

[202] J. Trumbo. The spatial environment in multimedia design: Physical, conceptual, perceptual, and behavioral aspects of design space. *Design Issues*, pages 19–28, 1997. 23, 26

[203] UrRobot. Robii, 2012. (Online; last accessed 15th of May 2014)). 43, 44

[204] A. Virolainen, P. Akerman, and J. Häkkilä. Burn-to-share: Content sharing with mobile projectors. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*, MUM '10, pages 18:1–18:4, New York, NY, USA, 2010. ACM. 45

[205] A. Virolainen, M. Paldanius, A. Lehtiö, and J. Häkkilä. Projector-based multi-touch screen for situated interaction with a mobile phone. In *Mobile and Personal Projection workshop at CHI*, volume 11, 2011. 41

[206] D. Vogel and R. Balakrishnan. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, UIST '04, pages 137–146, New York, NY, USA, 2004. ACM. 32, 180

[207] K. Vredenburg, J.-Y. Mao, P. W. Smith, and T. Carey. A survey of user-centered design practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, pages 471–478, New York, NY, USA, 2002. ACM. 12

[208] R. Want, B. Schilit, N. Adams, R. Gold, K. Petersen, D. Goldberg, J. Ellis, and M. Weiser. The parctab ubiquitous computing experiment. In T. Imielinski and H. Korth, editors, *Mobile Computing*, volume 353 of *The Kluwer International Series in Engineering and Computer Science*, pages 45–101. Springer US, 1996. 4, 5

[209] K. Warren, J. Lo, V. Vadgama, and A. Girouard. Bending the rules: bend gesture classification for flexible displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 607–610, New York, NY, USA, 2013. ACM. 2

[210] S. Wars. Episode iv-a new hope. *Dir. George Lucas. Perf. Harrison Ford, Carrie Fisher, Mark Hamill, Alec Guinness. Twentieth Century Fox*, 1997, 1977. 8, 49

[211] D. Weir, S. Rogers, R. Murray-Smith, and M. Löchtefeld. A user-specific machine learning approach for improving touch accuracy on mobile devices. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, UIST '12, pages 465–476, New York, NY, USA, 2012. ACM. 2, 134, 190

[212] M. Weiser. The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun. Rev.*, 3(3):3–11, July 1999. 4, 22

[213] J. Wesiwood, H. Hoffman, D. Stredney, and S. Weghorst. The virtual retinal display: a new technology for virtual reality and augmented vision in medicine. *Medicine Meets Virtual Reality: Art, Science, Technology: Healthcare (r) evolution*, 50:252, 1998. 18

[214] K. D. Willis. A pre-history of handheld projector-based interaction. *Personal and Ubiquitous Computing*, 16(1):5–15, 2012. 6, 7, 45, 56, 58

[215] K. D. Willis, I. Poupyrev, S. E. Hudson, and M. Mahler. Sidebyside: ad-hoc multi-user interaction with handheld projectors. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, pages 431–440, New York, NY, USA, 2011. ACM. 46, 47, 55, 58

[216] K. D. Willis, I. Poupyrev, and T. Shiratori. Motionbeam: A metaphor for character interaction with handheld projectors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1031–1040, New York, NY, USA, 2011. ACM. 45, 58

[217] K. D. D. Willis, T. Shiratori, and M. Mahler. Hideout: Mobile projector interaction with tangible objects and surfaces. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '13, pages 331–338, New York, NY, USA, 2013. ACM. 61

[218] A. D. Wilson. Playanywhere: a compact interactive tabletop projection-vision system. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, UIST '05, pages 83–92, New York, NY, USA, 2005. ACM. 43, 44, 53, 134

[219] A. D. Wilson and H. Benko. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 273–282, New York, NY, USA, 2010. ACM. 123

[220] C. Winkler, M. Löchtefeld, D. Dobbelstein, A. Krüger, and E. Rukzio. Surfacephone: A mobile projection device for single- and multiuser everywhere tabletop interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, New York, NY, USA, 2014. ACM. vi, 44, 54, 95

[221] C. Winkler, K. Pfeuffer, and E. Rukzio. Investigating mid-air pointing interaction for projector phones. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '12, pages 85–94, New York, NY, USA, 2012. ACM. 62, 63

[222] C. Winkler, C. Reinartz, D. Nowacka, and E. Rukzio. Interactive phone call: synchronous remote collaboration and projected interactive surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '11, pages 61–70, New York, NY, USA, 2011. ACM. 53, 55

[223] C. Winkler, J. Seifer, D. Dobbelstein, and E. Rukzio. Pervasive information through constant personal projection: The ambient mobile pervasive display (amp-d). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, New York, NY, USA, 2014. ACM. 48, 54

[224] C. Winkler, J. Seifert, C. Reinartz, P. Krahmer, and E. Rukzio. Penbook: Bringing pen+paper interaction to a tablet device to facilitate paper-based workflows in the hospital domain. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '13, pages 283–286, New York, NY, USA, 2013. ACM. 42, 53, 54

[225] J. O. Wobbrock, A. D. Wilson, and Y. Li. Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, UIST '07, pages 159–168, New York, NY, USA, 2007. ACM. 91, 133

[226] X. Xiao and H. Ishii. Mirrorfugue: communicating hand gesture in remote piano collaboration. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, TEI '11, pages 13–20, New York, NY, USA, 2011. ACM. 85

[227] X.-D. Yang, E. Mak, D. McCallum, P. Irani, X. Cao, and S. Izadi. Lensmouse: augmenting the mouse with an interactive touch display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2431–2440, New York, NY, USA, 2010. ACM. 100

[228] F. A. Yates. *The art of memory*, volume 64. Random House, 1992. 27

[229] Z. Ye and H. Khalid. Cobra: Flexible displays for mobilegaming scenarios. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '10, pages 4363–4368, New York, NY, USA, 2010. ACM. 60

[230] K.-P. Yee. Peephole displays: pen interaction on spatially aware handheld computers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 1–8, New York, NY, USA, 2003. ACM. 34, 56, 141, 144, 160, 163

[231] T. Yoshida, Y. Hirobe, H. Nii, N. Kawakami, and S. Tachi. Twinkle: Interacting with physical surfaces using handheld projector. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 87–90. IEEE, 2010. 59