Nomadic ProjectionWithin Reach: Overcoming Deficiencies in Nomadic Information Management through Mobile Projected Interfaces

Dissertation zur Erlangung des Doktorgrades

Dr. rer. nat.

der Fakultät für Ingenieurwissenschaften, Informatik

und Psychologie der Universität Ulm

von

Christian Winkler

aus Lemgo

Institut für Medieninformatik

2016

Acting dean:

Prof. Dr. Tina Seufert

Referees:

Prof. Dr. Enrico Rukzio, Ulm University
Prof. Dr. Michael Weber, Ulm University
Prof. Dr. Albrecht Schmidt, University of Stuttgart

Day of Defense:

January 22, 2016

Christian Winkler:

Nomadic Projection Within Reach

Doctoral dissertation

This document was typeset in LaTeX using the typographical look-and-feel classicthesis developed by André Miede
available at http://code.google.com/p/classicthesis/.

For Tine

ABSTRACT

The achievements in pervasive and nomadic computing, today, allow us to take and use our mobile devices such as smartphones and tablets everywhere we go. However, the mobility came at a cost: the small screen space of mobile devices leads to a deficient Human-Computer-Interaction (HCI) across several aspects, ranging from missing overview and awareness about the own digital information flow, to lacking support for multi-tasking and (privacy-respectful) collaboration. Any more complex tasks quickly become cumbersome to perform on mobile devices, which created the demand for constantly increasing display sizes in recent years. However, physical screen sizes are constrained by the required mobility (smartphones have to fit pockets, the hand, and the ear), such as display resolutions are constrained by the capabilities of the human eye—and both have reached their limits as far as smartphones and tablets are regarded.

Conversely, mobile projection promises large projected displays to be created anywhere, anytime, from very tiny physical form factors by means of battery-powered pico-projectors that can be integrated, for instance, to smartphones. But while this is generally true, the low luminance of these projectors that falls short magnitudes below the requirements dictated by ambient light, precluded them from being used in nomadic on-the-go scenarios so far and instead relegated them to a niche existence. Further on, while successfully applied to some application domains, for instance gaming, interaction concepts for nomadic information management with these projected interfaces are yet to be presented.

To this matter, this thesis first provides an analysis of deficiencies in nomadic information management that might be addressable by mobile projection. It then continues by presenting a new framework called Nomadic Projection Within Reach that aims at allowing mobile projection to fulfill its full promise on improving nomadic interaction, today. By drastically decreasing the projection distance, the framework—compared to traditional usage rather counter-intuitively—promotes the utility of a small but bright projected display instead of a larger darker one. The additional nearby display leads to a touchable mobile multi-display environment (MMDE), bringing nomadic projection *within reach*, both physically as well as figuratively. All in all, these changes provide completely new opportunities for overcoming the current deficiencies in nomadic information management using mobile projection. Further on, for use cases that require to cover a larger distance, an advancement of the framework to Nomadic Projection Within *Extended* Reach

is proposed which crosses boundaries between within-reach and out-of-reach projection and interaction.

The framework is established and evaluated through five case studies that systematically investigate its application to the aforementioned deficiencies. Through technical and conceptual explorations, innovations, and evaluations as well as qualitative and empirical user research, they demonstrate the framework's ability to alleviate all of these deficiencies through tailored interaction concepts that are enabled by touch-interaction and the MMDE. The proposed framework is complemented by a set of 12 design guidelines, which are derived from the case studies. These assist in deciding whether the framework is applicable to a (new) type of device and if so, provide guidance in decisions regarding integration and placement of the projector, positioning of the projection, techniques for transferring content between displays, to name a few.

Finally, as mobile and nomadic computing are quickly advancing fields, the process of developing the proposed framework has witnessed the (re-)emergence of several new nomadic device and display categories like, for instance, smart watches and glasses. These have very distinct advantages and disadvantages compared to mobile projection and prompt the question about the future role of mobile projection within this new ecosystem of (wearable) nomadic devices and displays. A thorough prospect on future work at the end of this thesis, pursues this question by providing an overview of strengths and weaknesses of devices within this ecosystem. It demonstrates, how by combining these devices the strengths of one device can be used to surmount the weaknesses of another, thereby also highlighting a possible unique role of mobile projection in the future. Further, two different approaches for these device combinations are proposed, as well as two further case studies that already started investigating these approaches, hopefully spurring new research agendas in this exciting new area of nomadic information management.

ZUSAMMENFASSUNG

Dank der Entwicklungen in den Bereichen „Pervasive-“ und „Nomadic Computing“ können wir heutzutage unsere mobilen Endgeräte wie Smartphones und Tablets überall mit hinnehmen und dort auch uneingeschränkt benutzen. Diese Mobilität ging allerdings deutlich zu Lasten der Interaktionsmöglichkeiten mit den Geräten. So erlauben die kleinen Bildschirmgrößen z.B. weder einen guten Überblick über die eigene digitale Informationswelt noch Multi-Tasking oder (die Privatsphäre schützende) Kollaboration. Jegliche komplexere Interaktionen auf den mobilen Geräten werden schnell unübersichtlich, was dazu geführt hat, dass die Displaygrößen in den letzten Jahren ständig gestiegen sind. Aber dieses Wachstum ist klar limitiert durch die erforderliche Mobilität der Geräte, ebenso wie deren Auflösung durch die Fähigkeiten des menschlichen Auges limitiert ist – und beide Grenzen sind zumindest bei Smartphones und Tablets bereits überschritten.

Im Gegensatz dazu verspricht mobile Projektion die Möglichkeit, große Bildschirme nahezu überall und zu jeder Zeit zu erzeugen – so werden sie zumindest vermarktet – und das auch noch aus sehr kleinen physischen Formfaktoren, welche sich z.B. in Smartphones integrieren lassen. Allerdings gibt es ein Problem, welches darin besteht, dass die Lichtstärke dieser Projektoren noch um ein Vielfaches hinter dem zurückliegt, was für eine uneingeschränkte, mobile Interaktion unter nahezu beliebigen Umgebungslichtbedingungen notwendig wäre, wodurch sie bislang ein Nischendasein fristen. Darüber hinaus sind zwar bereits Interaktionskonzepte mit mobilen Projektionen in einigen Anwendungsbereichen gezeigt worden, unter anderem z.B. im Spielebereich, Interaktionskonzepte zum (gemeinsamen) Handhaben von Informationen unterwegs sind aber bislang noch unerforscht.

Hierzu präsentiert die vorliegende Arbeit nun ein neues Rahmenwerk mit dem Titel „Nomadic Projection Within Reach “, zu Deutsch „Projektion unterwegs in greifbarer Nähe“. Mit diesem rückt mobile Projektion nicht nur im wörtlichen, sondern auch im übertragenen Sinne in „greifbare Nähe“ und kann für das Informationsmanagement unterwegs erfolgreich eingesetzt werden. Die drastische Verkürzung der Projektionsdistanz führt nämlich zwar zu sehr viel kleineren, aber dafür auch deutlich helleren Projektionsflächen, die, wenn auch ungewöhnlich für Projektion, als zusätzliche Displays großen, dunklen, projizierten Displays vorzuziehen sind. Die daraus resultierende mobile Mehrbildschirm-Umgebung und die Möglichkeit auf der kurzen Projektionsdistanz über Fingerberührung zu interagieren eröffnen ganz neue Möglichkeiten für die Handhabe von Informationen unterwegs. Darüber hinaus wird auch noch ein erweitertes Rahmenwerk „Nomadic Projection Within *Extended* Reach“, zu Deutsch etwa „Projektion

unterwegs in bedingt greifbarer Nähe“, vorgestellt, welches die grenzübergreifende Interaktion zwischen naher und weit entfernter Projektion und Interaktion dort erlaubt, wo nahe Interaktion alleine nicht ausreichend ist.

Das Framework wurde anhand von fünf Fallstudien entwickelt und evaluiert, welche sich systematisch mit den zu Beginn angesprochenen Nachteilen aktueller mobiler Endgeräte befassen. Durch technische sowie konzeptionelle Untersuchungen, Innovationen und Studien sowie qualitativer und empirischer Benutzerforschung zeigen und belegen diese Studien die Fähigkeit des Rahmenwerks, durch neue Interaktionskonzepte die gegenwärtigen Nachteile zu beseitigen. Das Framework wird darüber hinaus von einem Satz von 12 Designregeln komplementiert, welche aus den Fallstudien abgeleitet wurden und Entwickler neuer Geräte bei der Entscheidung, ob und wie das Rahmenwerk angewendet werden soll, unterstützen können. Dies beinhaltet u.a. Fragen zur Integration und Platzierung des Projektors im Gerät, zur Position und Ausrichtung der Projektion um das Gerät herum sowie Techniken zum Transfer von Inhalten zwischen den Geräten.

Zuletzt, da es sich beim „Nomadic Computing“ um ein schnell wachsendes Forschungsfeld handelt, haben während der Entwicklung dieses Frameworks einige andere Geräte- und Displaytechnologien (wieder) an Bedeutung hinzugewonnen, z.B. intelligente Uhren oder Brillen. Diese haben im Vergleich zu mobiler Projektion sehr unterschiedliche Vor- als auch Nachteile. Das wirft die Frage auf, welche Rolle mobile Projektion in der Zukunft für dieses Ökosystem aus mobilen Geräten spielen wird. Dieser Frage geht das Ausblickskapitel dieser Arbeit nach, indem es einen Überblick über die Stärken und Schwächen der Geräte dieses Ökosystems bietet und anhand von Beispielen aufzeigt, wie in der Kombination die Stärken des einen Gerätes genutzt werden können, um die Schwächen eines anderen zu kompensieren (so wie es zuvor auch schon für mobile Projektion und Smartphones gezeigt werden konnte). Darin zeigt sich dann ebenfalls die einzigartige Nützlichkeit mobiler Projektion in diesem Ökosystem. Darüber hinaus werden zwei generelle Ansätze zur Kombination von Geräten innerhalb dieses Ökosystems vorgestellt sowie zwei erste Fallstudien, welche diese Ansätze untersuchen und eine Grundlage für zukünftige Forschungsarbeiten, zum Informationsmanagement unterwegs, darstellen können und diese hoffentlich motivieren.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

[W1] Gugenheimer, J., **Winkler, C.**, Wolf, D., Rukzio, E., "Interaction with Adaptive and Ubiquitous User Interfaces." In: *Companion Technology - A Paradigm Shift in Human-Technology Interaction*. Red. by J. Carbonell, M. Pinkal, H. Uszkoreit, M. M. Veloso, W. Wahlster, and M. J. Wooldridge. Cognitive Technologies. Springer, 2016, to appear (cit. on p. 185).

[W2] **Winkler, C.**, Gugenheimer, J., De Luca, A., Haas, G., Speidel, P., Dobbelstein, D., Rukzio, E., "Glass Unlock: Enhancing Security of Smartphone Unlocking Through Leveraging a Private Near-eye Display." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. ACM, 2015, pp. 1407–1410 (cit. on pp. 228, 231 sq.).

[W3] Gugenheimer, J., Knierim, P., **Winkler, C.**, Seifert, J., Rukzio, E., "UbiBeam: Exploring the Interaction Space for Home Deployed Projector-Camera Systems." In: *Human-Computer Interaction – INTERACT 2015*. Vol. 9298. Lecture Notes in Computer Science. Springer International Publishing, 2015, pp. 350–366 (cit. on p. 35).

[W4] **Winkler, C.**, Seifert, J., Dobbelstein, D., Rukzio, E., "Pervasive Information Through Constant Personal Projection: The Ambient Mobile Pervasive Display (AMP-D)." In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14, **Honorable Mention Award**. ACM, 2014, pp. 4117–4126 (cit. on p. 160).

[W5] **Winkler, C.**, Löchtefeld, M., Dobbelstein, D., Krüger, A., Rukzio, E., "SurfacePhone: A Mobile Projection Device for Single- and Multiuser Everywhere Tabletop Interaction." In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 3513–3522 (cit. on p. 101).

[W6] **Winkler, C.**, "Peripheral Interaction On-The-Go." In: *Workshop on Peripheral Interaction: Shaping the Research and Design Space at CHI 2014*. CHI '14. 2014 (cit. on pp. 183, 212, 219).

[W7] **Winkler, C.**, Rukzio, E., "Projizierte tischbasierte Benutzungsschnittstellen." In: *Informatik-Spektrum* 37.5, Springer (2014), pp. 413–417.

[W8] **Winkler, C.**, Seifert, J., Reinartz, C., Krahmer, P., Rukzio, E., "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, **Best Note Award**. ACM, 2013, pp. 283–286 (cit. on p. 87).

[W9] **Winkler, C.**, Pfeuffer, K., Rukzio, E., "Investigating mid-air pointing interaction for projector phones." In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*. ITS '12. ACM, 2012, pp. 85–94 (cit. on pp. 57, 84).

[W10] **Winkler, C.**, Hutflesz, P., Holzmann, C., Rukzio, E., "Wall Play: A Novel Wall/Floor Interaction Concept for Mobile Projected Gaming." In: *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services Companion*. MobileHCI '12. ACM, 2012, pp. 119–124 (cit. on pp. 36, 39 sq., 52, 54, 109).

[W11] **Winkler, C.**, Reinartz, C., Nowacka, D., Rukzio, E., "Interactive phone call: synchronous remote collaboration and projected interactive surfaces." In: *Proceedings of the 2011 ACM international conference on Interactive tabletops and surfaces*. ITS '11. ACM, 2011, pp. 61–70 (cit. on p. 132).

[W12] **Winkler, C.**, Broscheit, M., Rukzio, E., "NaviBeam: Indoor Assistance and Navigation for Shopping Malls through Projector Phones." In: *MP2: Workshop on Mobile and Personal Projection*. CHI 2011. 2011 (cit. on pp. 170, 182).

Further co-authored publications that are not directly related to the thesis' topic are:

[W13] Schaub, F., Könings, B., Lang, P., Wiedersheim, B.**, Winkler, C.**, Weber, M., "PriCal: Context-adaptive Privacy in Ambient Calendar Displays." In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '14. ACM, 2014, pp. 499–510.

[W14] Seifert, J., Boring, S.**, Winkler, C.**, Schaub, F., Schwab, F., Herrdum, S., Maier, F., Mayer, D., Rukzio, E., "Hover Pad: Interacting with Autonomous and Self-actuated Displays in Space." In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. UIST '14. ACM, 2014, pp. 139–147 (cit. on pp. 20, 216).

[W15] Derthick, K., Scott, J., Villar, N.**, Winkler, C.,** "Exploring smartphone-based web user interfaces for appliances." In: *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*. MobileHCI '13. ACM, 2013, pp. 227–236.

[W16] Valderrama Bahamondez, E. d. C., **Winkler, C.**, Schmidt, A., "Utilizing multimedia capabilities of mobile phones to support teaching in schools in rural panama." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. ACM, 2011, pp. 935–944.

[W17] Shirazi, A., **Winkler, C.**, Schmidt, A., "SENSE-SATION: An extensible platform for integration of phones into the Web." In: *Internet of Things (IOT), 2010*. Internet of Things (IOT), 2010. 2010, pp. 1–8.

[W18] Shirazi, A. S., **Winkler, C.**, Schmidt, A., "Flashlight Interaction: A Study on Mobile Phone Interaction Techniques with Large Displays." In: *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '09. ACM, 2009, 93:1–93:2.

To accomplish great things,
we must not only act, but also dream,
not only plan, but also believe.

— Anatole France

ACKNOWLEDGMENTS

Working on this thesis has been the most exciting and challenging part of my life and I'm deeply grateful to many for their professional and emotional support. First and foremost, I want to thank **Enrico Rukzio** for supervising my work and the creative work environment he provided throughout the years. In particular, his commitment and trust in my research, supporting even unconventional research ideas with money and thought, was key to my academic success. Besides, he always was a great source of advice, focus, and inspiration and a role model of good leadership that I could not have better wished for and led to a friendship that I really appreciate. Since moving with the group to Ulm, the good work environment was all the same credit to the great personality of **Michael Weber**, head of the institute, whom I further thank for consenting to examine this thesis based on his expertise in the field of pervasive computing. The good work environment last but not least was enabled by the DFG whom I thank for their financial support as part of the Emmy Noether research funding on "Mobile Interaction with Pervasive User Interfaces".

Furthermore, I'm deeply thankful to **Albrecht Schmidt**, to whom I owe finding my interest in research, in Ubicomp and HCI, and the decision to pursue a PhD. Over the years, he kept regularly checking on me and supporting my case through recommendations for internships and applications and I'm glad to have won him as another examiner of my thesis.

I'm happily looking back to awesome travels around the globe (especially Tokyo, Vancouver, and Copenhagen) with **Alireza Sahami Shirazi** and legendary tennis matches (my pleasure to let you win) with **Florian Alt**, two of my first colleagues who became good friends over time. Throughout the years, my office roommate **Julian Seifert** has been a great colleague and friend. Having started together and having been the only two PhD candidates in the research group for some time, I happily look back (and not without some chuckling) to our endless discussions about the secret "how to get accepted at CHI"-recipe (I'm not sure whether we really found it but at least we came a lot closer $\ddot\smile$). Besides, I was regularly impressed by his profound knowledge across various domains and picking his mind for comments and solutions was always a great benefit. I'm further thankful for his unselfish help in reviewing drafts and last-minute help in getting projects finished.

Christian Reinartz, **Ken Pfeuffer**, and **David Dobbelstein** have been wonderful students and later colleagues to start with during my first PhD years in Essen. Thank you for staying overnights with me and your contributions to successful publications together. All the more, the steadily growing research group in Ulm became an awesome place to be—not only because of great developing group habits like the regular table soccer and squash matches but because of the great people I got to know. First of all, I'm deeply thankful to my dear colleagues **Jan Gugenheimer** and **David Dobbelstein** for reviewing this thesis and providing more good comments than I could have possibly included. As much as I liked to impart my experience, I certainly have not learned any less from them and am grateful for a great time of discussions and their friendship. All the same I'm thankful for the friendship of **Bastian Könings** and **Björn Wiedersheim** and for having been always quick to help with anything. But also the other members of the Institute of Media Informatics deserve my gratitude for having been great colleagues: **Florian Schaub**, **Frank Honold**, **Felix Schüssel**, **Florian Geiselhart**, **Katrin Plaumann**, **Philipp Hock**, **Julia Brich**, **Michael Rietzler**, **Marcel Walch**, **Katja Rogers**, **Michael Haug**, and last but not least **Claudia Wainczyk**.

An influential time during my PhD has certainly been my internship with Microsoft Research in Cambridge, UK. I owe to **Steve Hodges**, **James Scott**, and **Nic Villar** for the experience of an awesome team culture and their imprint on me not to take the hardware as a given but create the UI experience from the ground up.

Finally, my deepest gratitude goes to my family, for their loving support throughout the years of study and research. Starting with my parents **Angelika and Detlev Winkler**, I'm utmost thankful for your unconditional support, care, and encouragement over the years, especially in stressful times. Further for many hours of your's proofreading this thesis regarding grammar and typos (I hope these acknowledgments turned out okay without that). I'm also indebted to my family in law (**Friedhelm**, **Ilse**, and **Stefanie Weking**), who more than once compensated my lack of time (because of upcoming conference deadlines) by helping out in the household and petting the abandoned souls of my wife and daughter. Furthermore, I thank my pal **Gideon Baumann** for having been such a caring friend throughout the years. Most of all, I praise the unwavering love of my beloved wife **Christine** that is beyond words. Thank you for your tolerance of countless weeks and weekends overshadowed by work, for your patience, for putting up to monologues about research projects and even showing real interest sometimes, but especially for the provided emotional stability and your devotion and love. And thank you **Marie–Sophie** and **Leonie–Salome**, my dear little daughters, for lightening up just any day.

I humbly thank God for blessing me with all of these wonderful people and opportunities and thank you all again for your support!

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

Appears light blue in text **First appears on page**

ANOVA	Analysis of Variance	13
ANSI	American National Standards Institute	28
API	Application Programming Interface	97
AR	Augmented Reality	7
CRT	Cathode Ray Tube	23
CV	computer vision	210
DLP	Digital Light Processing	26
DMD	Digital Micromirror Device	22
DoF	Degrees of Freedom	36
EMG	Electromyography	37
FOV	Field of View	19
GUI	Graphical User Interface	20
HCI	Human-Computer-Interaction	v
HWD	Head-Worn-Display	6
IMU	inertial measurement unit	175
IPC	Interactive Phone Call	132
IR	infrared	41
ISO	International Organization for Standardization	11
LBS	Laser Beam Steering	23

LCoS	Liquid Crystal on Silicon	22
lm	Lumens (definition in Subsection 2.4.1)	18
Lux	luminous flux (definition in Subsection 2.4.1)	18
lx	Lux	27
MEMS	Micro Mechanical System	22
MMDE	mobile multi-display environment	v
NPWR	Nomadic Projection Within Reach	84
NPWER	Nomadic Projection Within Extended Reach	160
OST	optical see-through	20
phablet	very large smartphone, almost the size of a tablet	88
PIM	Personal Information Management	4
ProCamS	Projector Camera System	37
PSSUQ	Post Study System Usability Questionnaire	13
PPI	pixels per inch	213
QR	Quick Response	41
ROI	region of interest	198
SAR	Spatial Augmented Reality	162
SLAM	Simultaneous Location And Mapping	32
SUS	System Usability Scale	13
UCD	User-Centered-Design	11
VR	Virtual Reality	19
VRD	Virtual Retinal Display	18
VST	video see-through	20
WoZ	Wizard-of-Oz	14

Part I

A CASE FOR NOMADIC PROJECTED INTERFACES

1 INTRODUCTION

1.1 MOTIVATION

Computing has not just become mobile, but nomadic. Whereas the invention of the laptop computer only increased the mobility of computing, today's "smart" devices such as smartphones enabled nomadic computing for everyone. Wikipedia defines nomads as members "of a community of people who live in different locations, moving from one place to another" [185]. Following this idea Leonard Kleinrock has coined the term "Nomadic Computing" in the '90s [142] and defined "nomadicity" as

> "the system support needed to provide computing and communications capabilities and services to nomads as they move from place to place in a way that is transparent, integrated, convenient and adaptive." [143]

During that time, Kleinrock and his colleagues were mainly concerned with nomadic use of Internet connectivity and it is in part thanks to them that switching networks while on-the-go, today is transparent, adaptive and thus convenient for the end-user. Like advancements on the network level enabled today's nomadic computing, all the same did the ongoing miniaturization that allowed to carry smartphones (today extended to smart watches, glasses, etc.) as a general companion in everyday life. Today people read and write their e-mails, check news, write messages, remotely control their smart home, share documents, etc. any time, any place, and independent of a specific environment. Through the addition of sensors like camera, GPS, and motion sensors, even context-aware information provisioning and interaction have been enabled that reach far beyond some of the capabilities of desktop or mobile computers. Hence, we see that changes at the software *and* the hardware level have been required to create the necessary convenience that eventually enabled the shift from mobile to nomadic computing.

However, these changes for convenience came at a cost: the comparably small display for output and input, the lack precise input controls (such as a mouse), single activity focus of the operating system, to name a few, render certain nomadic tasks cumbersome to perform, if not impossible for the time being. This is amplified for around 7-15% of the population, where the nomadic computing device is *the only* computing device they have easy access to [239] and which percentage could drastically increase if nomadic devices provided better support for traditional desktop tasks like accounting, video cutting, teleconferencing etc.

Although games, entertainment and navigation have increased, currently, functions of Personal Information Management (PIM) are and probably will continue to be primary tasks in nomadic computing in Germany [238] and the US [239] alike. These tasks include messaging, email, surfing, social networks, news and management of personal schedules, notes and task management. However, even when these task become more complex, e.g. planning a trip and doing research on hotels and flights, traditional computers are preferred as they, for instance, allow to compare multiple websites side by side. Naturally, this is even more true for the more traditional computing tasks previously mentioned and especially for collaboration where nomadic devices preclude necessary requirements like sharing a common screen, input area or easy means for sharing data.

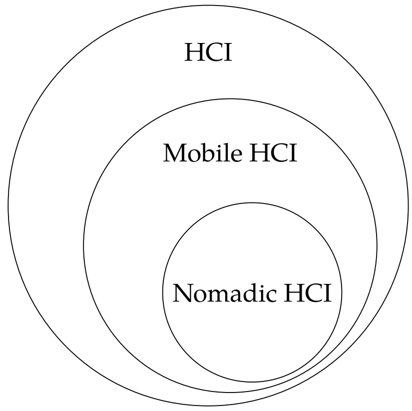


Figure 1.1
Mobile HCI considers the implications of the small form factor, mobility, and constrained resources of mobile devices on HCI. Although inspired by nomadic usage, mobile devices are similarly used in familiar places that are supported by infrastructure, like at home, at work, etc. Nomadic HCI is concerned with issues in Mobile HCI that pertain to nomadic usage such as being at unfamiliar or uncontrollable locations and on-the-go.

This leads us to the motivation of this thesis, which sets out to improve nomadic computing, particularly for personal and collaborative information management, by means of new device and interaction concepts. These are reflected in research on human factors and mobile nomadic HCI (see Figure 1.1) just the same as new soft- and hardware solutions. As the thesis' title suggests, particularly the application of mobile projectors to this problem is investigated. But before the thesis' research approach, its structure and its contributions will be outlined in more depth, we will have a closer look at the concrete *deficiencies* of today's mobile devices as related to nomadic computing and nomadic information management in particular.

1.1.1 Deficiencies of Mobile Devices

1.1.1.1 *Output/Input [1] Size Deficiency (D1)*

In terms of display real estate, mobile devices are strictly limited to maintain a small mobile form factor. This was not a problem before the advent of smartphones when Motorola, for instance, introduced their very successful StarTAC (display diagonal 0.25") in 1996 as the "smallest mobile phone of the world" [2] – a smartphone marketing strategy in sharp contrast to today's. But the more power and content were made available on our mobile companions, which could potentially evolve to our primary computing devices some day, the higher became users' needs to perform more complex tasks. This inevitably led to a constant—fourfold since 2009 as depicted by Figure 1.2—increase of screen size, up to 5.7" (Samsung Galaxy Note 4) diagonal within the last years. The fact that people *willingly* give up some of their nomadic convenience for the benefit of a slightly larger out- *and* input area highlights the importance of screen size even for nomadic computing. Yet, mobile devices can only grow so much further not to lose their mobile quality and justification and as such there remain tasks that cannot be (optimally) performed on nomadic devices.

Apart from deficiencies that will be discussed by the next sections, the small size directly leads to longer task completion times and higher error rates compared to traditional computing devices. These are caused by the fact that information has to be split across many screens on one hand, and fingers occluding much of the content during the interaction (the so-called fat-finger problem) on the other hand. Thus mobile devices diminish overview and accurate pointing, two important requirements of any interaction in HCI. The latter is alleviated by comparably large button sizes (the same buttons would be more than 5 cm high on PC monitors running at the same resolution), which again diminishes

1 Please note that the usual order of Input/Output or I/O has been reversed deliberately as projectors address much more easily the output than the input deficiency
2 http://en.wikipedia.org/wiki/Motorola_StarTAC, visited November 8th, 2015

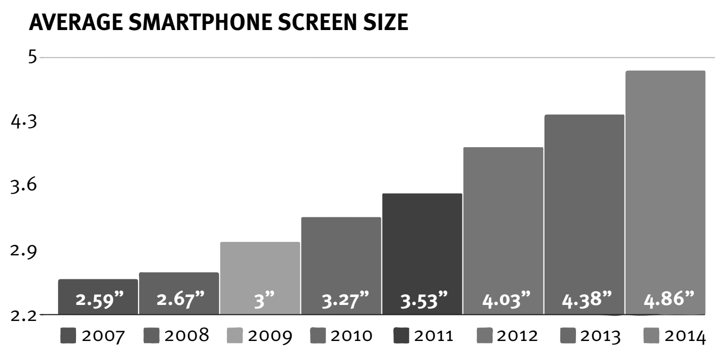


Figure 1.2: Evolution of smartphone screen sizes which increased fourfold during the last years. Based on image by PhoneArena [84]

the amount of content that can be shown at a single time and increases task completion times.

A solution to the inherent size constraint of mobile screens are mobile projected interfaces as they decouple the size of the offered display from the physical size of the device creating it. As such they allow the creation of large displays from very small mobile form factors. Today mobile projectors come at a large variety of different sizes, mostly depending on brightness and battery runtime of the projector. These mobile projectors, to which all small projectors belong that are powered by battery, are typically categorized into portable (can be carried with one hand), pocket-sized (fits in the palm of the hand), and pico/handheld-projectors (can easily be operated in the hand like a mobile phone or even smaller). Besides accessory devices, pico-projectors are also built into other devices like video cameras[3] and smartphones, so called projector phones like the Samsung Galaxy Beam [221] (a detailed view is given in the next chapter, Chapter 2). Apart from that, pico-projectors are also used in many Head-Worn-Displays (HWDs) for creating the virtual display on a surface in front of the eye, e.g. Google Glass [93], or by directly projecting onto the eye's retina [162]. The Future Work chapter (Chapter 12) will reflect on the future influence of HWDs and other wearable display technologies on the application of projection for nomadic computing.

1.1.1.2 *Multi-tasking Deficiency (D2)*

Multi-tasking currently is very cumbersome to perform with mobile devices at macro and a micro level.

At the micro level, mainly because of the small input/output space all mobile OS currently follow the "single activity focus" concept. Switching between applications requires multiple interactions steps like pressing a button, scrolling through a list of applications and then opening it through another touch selection. Switching between windows or activities *within* an application is entirely unsupported at the moment. Even considering new ideas like split-screen applications in Apple's iOS 9, mobile devices do not even come close to the support of overview and task switching provided by traditional operating systems. Apart from a larger display size, allowing for multi-modal input can support multi-tasking as the system does not enforce to be used through a modality that may be occupied at the moment. Mobile devices are very focused on the touch-input modality and allow only very limited control through alternative input modalities such as speech, which is not always appropriate to use.

At the macro level, switching from a real task in the environment to a digital task performed using the device is difficult as well. The mobile device has usually to be retrieved from a pocket or bag, then enabled and possibly unlocked, and finally the desired app has to be

3 Nikon S1100pj and S1200pj [184]

opened. This requires at least one free hand, depending on the hand and screen size of the user maybe even two. Putting the device away and switching back to the real world task only is a little quicker. Hence, micro-interactions such as checking the time are not possible to perform within a few seconds and in parallel to real world tasks. This second connotation is related to the awareness aspect of the upcoming Environment Deficiency (D4), but regards active task switching instead of subconscious multi-tasking.

1.1.1.3 *Collaboration & Privacy Deficiency (D3)*

Related to the small in- and output size, collaborating on content on mobile devices is extremely difficult. Although mechanisms to simplify sharing of data exist, for instance Android's BEAM[4] functionality, collaboratively working on information like it is possible on desktop and particularly tabletop computers is impossible. Oftentimes, the mobile device has to be handed away, giving up control over very private device information. This automatically leads to a deficiency of privacy support as well. Because the display of the device is coupled to the device and its input modalities, the display cannot be handed out separately, or at least moved in place, without giving away the device itself. Projected displays, in contrast, allow for independent placement of the display.

1.1.1.4 *Environment Deficiency (D4)*

The last deficiency mainly regards the connectedness to the environment, which is much more important in nomadic scenarios as it is more arbitrary as, for instance, in the office or at home. On one hand, because the nomadic user is usually more distracted by the surrounding environment, i.e. people nearby, ongoing discussion, finding the way, crossing a street, etc. Hence, it is more challenging to gain the user's attention. As phones and tablets are carried in pockets or handbags, they cannot easily make the user aware of new information apart from very limited information channels like vibration and audio alerts. At the same time, the user should not be distracted so much that they lose the connection to the environment. Focusing on a small screen display typically leads exactly to this disconnectedness and has been the topic of much debate when people ran into poles while texting or when they were "phubbing", a neologism comprised of "phone" and "snubbing" that an Australian publisher of dictionaries invented to describe the social discomfort that this disconnectedness creates in a group discussion.

Augmented Reality (AR) is believed to address this in a very suitable manner by combining the information with the user's environment, allowing the user to remain aware of the environment during interaction

4 By putting Android devices with BEAM enabled back to back, their NFC technology recognizes each other and initiates data transfer of the current screen content [25].

with information. But not only the user is more connected, bystanders are so, too, as they possibly can better comprehend the actions of the user which may aid acceptance. However, current mobile devices are not well suited for this type of AR. First, because they do not allow for hands-free operation, which also renders them unsuitable to support the user continuously during a complex or long-running task. Furthermore, their support for an augmented view is limited to the position, angle and size of the small screen. Projections, on the other hand, augment the environment *directly*, which not only creates believable augmentations, but also does not limit the user's hands or pose to perceive the augmentation. It should be noted, that the publicity of this type of AR impacts the privacy of the user, which should be considered and addressed, for instance, in combination with the previous deficiency.

1.2 RESEARCH AGENDA & THESIS CONTRIBUTIONS

Altogether, mobile projectors seem well suited to address these important deficiencies of information management in current nomadic computing. Complementing the identified deficiencies, the following research questions formulate very concrete questions rather than research domains, which will be answered in the conclusion of this thesis. Chapters (partly) answering these questions are shown on the margin:

Chapter 4
R1 Do larger projected displays support quicker task completion times and lower error rates for interaction?

Chapter 6
R2 Which new input modalities are enabled that aid information management? And what are their unique affordances, requirements, or limitations, respectively?

Chapter 7
R3 Which new types of collaboration are enabled by one or several additional projected displays?

Chapter 7 & Chapter 8
R4 Does an additional projected display increase privacy awareness and privacy management?

Chapter 8 & Chapter 9
R5 Can projected interfaces increase awareness of information through peripheral perception?

Chapter 9
R6 Can worn projected interfaces shorten lead-time for micro interactions?

Chapter 5
R7 Given the comparably low brightness of mobile projectors (two magnitudes lower than their static counterparts), which is often said to preclude nomadic usage scenarios due to out-of-control environmental light, can nomadic scenarios be realistically supported today?

Complementary to the previous research questions, the case studies presented in this thesis address the deficiencies as follows (corresponding chapters in parentheses):

Output/ Input Size	Multi-tasking	Collab. & Privacy	Environment
Penbook (6)	Penbook (6)	SurfacePhone (7)	Penbook (6)
SurfacePhone (7)	SurfacePhone (7)	IPC (8)	SurfacePhone (7)
IPC (8)	IPC (8)		IPC (8)
	AMP-D (9)		AMP-D (9)
	SpiderLight (10)		SpiderLight (10)

Table 1.1: How the case studies presented in this thesis address the deficiencies of current nomadic computing (chapters in parentheses).

As Chapter 3 will show, many previous works on mobile projection did not consider nomadic usage (Section 3.1). Of those who did, most applied out-of-reach projection, which as will be shown in Subsection 2.4.1 and further elaborated in Chapter 5, is unsuitable for nomadic scenarios. Of the few remaining works on nomadic within-reach interaction, almost none have focused on information management and the deficiencies described before (Section 3.2). This work is thus the first to methodologically research how mobile projection can aid nomadic information management. In particular, it provides the following conceptual and engineering contributions to the body of knowledge of HCI and practitioners in the fields of UI and UX design of mobile devices:

Theoretical & Conceptual Contributions

1. A thorough analysis of the deficiencies of mobile interaction in nomadic scenarios and a set of hypotheses (research questions) how mobile projection may be able to address these.
2. An overview of principles and related work on (mobile) projection and assessing their relation and applicability to nomadic usage scenarios.
3. A proposed framework called Nomadic Projection Within Reach that evolves from elaborated technical as well as human factor constraints and is tailored to nomadic usage. It promotes projected interfaces to be brought very close to the originating device respectively its user, such that projection gets both physically as well as figuratively "within reach".
4. The evaluation of the previously mentioned concept using three in-depth case studies on supporting mobile users in nomadic scenarios using Nomadic Projection Within Reach, addressing aforementioned deficiencies.
5. An extension of the former framework to on-the-go scenarios and situations where within-reach interaction alone is not sufficient. This is termed Nomadic Projection Within Extended Reach and advances the framework to a cross-distance continuous interac-

tion space. Two further case studies evaluate the extended framework.

6. A set of 12 concrete design guidelines that support practitioners in directly applying the knowledge gained through the case studies to their (new) device and UI designs.
7. An extensive prospect on future work that positions mobile projection within the broader scope of wearable displays for nomadic computing, highlighting both its unique advantages over other in- and output technologies as well as its limitations and how these may be addressed in the future.

Engineering Contributions

1. Novel *software algorithms*
    a) to recognize special gestures and walking steps (AMP-D) as well as touch interaction under extreme conditions (SurfacePhone);
    b) to adjust projector focus (AMP-D);
    c) for semi-automatic calibration techniques where autonomous calibration is not possible (SurfacePhone, AMP-D).

    Some of these have been made available open source (Chapter 7).
2. *Integration* of existing components such as projectors, (depth) cameras, inertial sensors, servo motors, power management to novel device concepts enabling unprecedented opportunities for nomadic interaction.
3. *Applications* for nomadic computing and information management such as sharing media (SurfacePhone, IPC, AMP-D), playing games (SurfacePhone), managing notifications (AMP-D), schedules and maps (IPC), applications for the hospital domain (Penbook) as well as quick lookup of location-aware information (SpiderLight).
4. *Design and construction* of optical light paths (esp. Penbook, SurfacePhone, SpiderLight), hardware cases (SurfacePhone, AMP-D) and carrying facilities (AMP-D) for standalone nomadic devices of which some have been made available open source.

Worthwhile mentioning are further the numerous empirical contributions from the 13 conducted user studies presented or related to in this thesis, which beyond the implications that have explicitly been drawn from them, provide interesting insights into HCI.

1.3 RESEARCH METHODOLOGY

The field of Human-Computer-Interaction has a twofold goal: on one hand it seeks understanding users' behavior with technology, technology's influence on users' behavior, and the synergies as well as contradictions in-between. On the other hand, as a field of computer science (the related field in psychology is called "human factors") it also seeks inventing new human-computer interfaces that incorporate the achieved understanding to improve the interaction in terms of efficiency, effectiveness, learnability, simplicity, joy of use, to name but a few. The commonly applied design methodology in HCI is the "User-Centered-Design (UCD) Process" standardized under 9241-210 [123] by the International Organization for Standardization (ISO) which comprises both goals to an iterative process. It can be viewed like an iterative V- or waterfall-model in software engineering which puts an extreme focus on the user as main source of requirements and usability as a central goal of the product. As such, it also defines three typical iterative stages in software and hardware development:

Context and Requirements Specification For any IT project, understanding the requirements of stakeholders and the context and limitations of use are essential to success. In the context of nomadic HCI, this phase is of special importance as the context of use can be arbitrary, ranging from office to home to on-the-go contexts and the diversity of users regarding their age, size, gender, previous experience or exposure to technology, to name but a few, all play into the final user experience.

Solution Design Are the requirements believed to be understood as far as possible at the current stage, the solution design tries to implement these in prototypes that will optimally suit further evaluation and analysis until the final product is achieved. That said, the goal to target at this state is not to come as close as possible to a final product design, but to design the prototype for the sake of the subsequent evaluation to answer open questions that previous requirements engineering was not able or even possible to deliver.

Evaluation The evaluation involves users with the prototype to collect quantitative and qualitative data that informs future iterations of the design process. One of the most important aspects of this state is to assemble a representative set of users who represent a sample of the final users of the product.

1.3.1 Methods

1.3.1.1 *Context and Requirements Specification*

Methods used in this phase, for instance, comprise focus groups, questionnaires, and in-situ observations with possible end-users or domain experts. After the first iteration of the development life cycle, most of the changes to the specification are drawn from previous evaluations. However, participants of the previous evaluation may have surfaced related domains or desired functionality that is completely new and may require new initial assessment using the methods described before.

1.3.1.2 *Solution Design*

The instantiation of solution designs highly depends on the time when they occur within the development life cycle. In early phases of development, usually low-fidelity prototypes are created using methods such as paper-prototyping or digital but non-functional UI flows using tools like Balsamiq [35] that support linkage between sketches of UI states. These prototypes are considered horizontal prototypes as they try to visualize the scope of interaction rather than an in-depth exploration of how the actual interaction looks like. More seldom, videos are used to present an interaction to users, when, for instance, an equal appearance of the prototype between participants is of utter importance (e.g. research on social acceptability).

However, considering mobile projection horizontal prototyping often leads to unrealistic imaginations of users as at least handheld projection involves a completely different type of interaction compared to standard graphical user interfaces on desktop computers or even mobile devices. Thus, vertical prototypes are often required to prototype parts of the interaction, demanding the complete chain of tracking users and visualizing their interactions. In case studies, which primarily focus on qualitative research questions (see for example Chapter 8) some technical simplifications that remain unnoticed by the user are possible, like mimicking an on-device tracking with tracking technology in the environment. But other projects require full vertical prototypes (Chapter 6 and Chapter 9) or a complement of both (Chapter 7). The latter is important if a realistic instantiation of the concept is not achievable by using rapid prototyping and the concept and technical research questions have to be evaluated separately.

1.3.1.3 *Evaluation*

The evaluation seeks understanding the complete user experience. This consists of how well a system works at executing the desired function (assuming the user performed the interaction right), how well users are able to perform the right interactions to achieve their desired goals, including aspects like clarity of presentation and documentation found in widespread usability measures such as IBM's System Usability Scale (SUS) [55] and the Post Study System Usability Questionnaire (PSSUQ) [158].

QUALITATIVE VS. QUANTITATIVE The data that is gathered through user studies must typically be structured into different classes of information. The first distinction to be made is usually between quantitative and qualitative data. Quantitative data typically results from objective observations like the time required to complete a task, the amount of errors occurring during the interaction and other types of data that can be measured on a continuous scale. For these types of data, statistical tests for parametric data are valid to apply, for instance to test for significant differences between study conditions using an Analysis of Variance (ANOVA). In contrast, users' subjective answers to questions are typically considered *non-parametric* information. At that, it is irrelevant whether the information has been gathered on a quantitative scale, such as a Likert scale [186] ranging from 1 (very good) to 5 (very bad), because users cannot be assumed to mean the exact same thing when they label something as *good*. Instead, only an ordinal, discrete scale can be assumed where relations per participant hold, but not between participants. Therefore, quantitative non-parametric information must be analyzed with statistical methods that acknowledge the included uncertainty using statistical methods such as Friedman's ANOVA and Wilcoxon signed-rank post-hoc tests that are considered more conservative than those used for *parametric* information. Other qualitative data, like eye dominance, profession of the participant, open ended questions, typically result in nominal data types which store a set of possible values which are unrelated. Especially for verbose answers, or transcribed observations from audio and video captured during the study, Grounded Theory [92] can be used to find categories and thereby cluster the data into quantitative scales which may also be evaluated using parametric tests. Moreover, for some specific situations, transformations from non-parametric data to parametric data have been proven valid. For instance, asking the same question in multiple, partly opposite ways, multiple Likert scales can be averaged to a continuous interval scale, which is robust to parametric evaluation [186].

INTERNAL VS. EXTERNAL VALIDITY An important distinction to make and decision to take is whether the user study should achieve a high internal or external validity as both usually is not achieved at the same time in HCI studies. Lab studies provide a high internal validity as the conditions of the study can mainly be controlled by the experimenter. However, the lab environment creates an artificial environment in which users cannot be expected to show their very natural behavior. On the other hand, field studies where users participate unattended and often in their natural environments, likely inhibit a huge variety of situations that may result in a lot of interesting qualitative findings. Resulting quantitative data is often hard to interpret or questionable in its comparability, though (cf. [141]). A common approach is to start with lab studies, also to assure that the prototype works well and as expected under controlled conditions. If so, and if the lab study could not entirely reveal the desired usage information, field studies may be conducted. Foundational research, targeting atomic interaction elements, may also directly conduct field studies, for instance using app stores to release questionnaires or simple games to the masses at once (cf. [110]).

Unfortunately, conducting field studies on nomadic projected interfaces presents itself very challenging as it either requires a large budget to create multiple instances of projector(-camera) systems consisting of multiple expensive hardware items. Or, it is very time consuming if many users are to take one or few devices to their homes for several weeks. In both cases, very polished prototypes that work without the experimenter's intervention and furthermore in many different scenarios are required, which alone because of the fact that mobile projectors only provide very limited brightness is difficult. Naturally, studies cannot be conducted by relying on app stores as participants do not have the required hardware. That said, so far the research community has failed to deliver long-term field studies on mobile projected interfaces and it will be important to carry them out in the future.

USER-ELICITATION Instead of designing interaction techniques and evaluating them afterwards, an increasingly widespread approach in HCI has become to elicit them from users by asking them how they would like to perform certain interactions. This approach, first described by Nielsen et al. [183] and later formalized by Morris et al. [177], Vatavu et al. [257], and Wobbrock et al. [275, 276] is often applied in interaction domains that inhibit a large variability, e.g., gesture interaction. When designing for such an interaction domain, biasing of the designer or expert users is likely to occur and existing mental models of users may be unknown leading to gestures that feel unfamiliar to users. On the other hand, interactions created by users may lead to a higher and easier adoption, although as Vatavu et al. [257] show, agreements above 30% between users are rather unlikely to achieve. A disadvantage is that users tend to propose familiar interactions without having a long time to test them which may lead to non-optimal or even completely unsuitable interactions (cf. Chapter 4). User-elicitation has been applied in chapters 4, 7 and 10.

WIZARD-OF-OZ For the sake of completeness, so-called Wizard-of-Oz (WoZ) approaches must be named as well. In a WoZ study parts of the system are replaced by human operators but without the participant to notice. For instance, the participant thinks the speech-to-text interface is recognizing their voice but actually the text is manually given to the system through an operator in another room who listens via a hidden microphone. For the same reason as horizontal prototypes are oftentimes unfeasible in research on projected interfaces, so are WoZ approaches which cannot provide a realistic visual experience without latency. For instance, for pedestrian navigation instructions in the AMP-D study (Chapter 9) it was previously considered whether a manual positioning (location and rotation) of the participant by a human follower was possible. However, in pilot studies and despite the development of very good WoZ tools, the right timing was never achieved to guide participants adequately. Hence, WoZ was not applied throughout this thesis.

1.4 THESIS OUTLINE

The rest of this thesis is structured in the following six parts:

Part I Chapter 1 has detailed the motivation and research approach for this thesis. Chapter 2 will present basic principles and methods regarding mobile projected interfaces, which not only aids a technical understanding but furthermore as motivation for within-reach interaction as promoted by the framework presented later. Afterwards, Chapter 3 classifies related works on projected interfaces according to the dimensions *mobility*, *interaction distance*, and *application domain* and presents closely related works on nomadic projection and information management in detail. Chapter 4 argues evaluated disadvantages of out-of-reach interaction for nomadic projection from a human factors perspective, which complement the technical motivations for within-reach interaction.

Part II Presents the framework of Nomadic Projection Within Reach based on arguments of previous chapters. This is followed by three case studies (chapters 6 to 8) that investigate the framework's application to the previously outlined deficiencies of current mobile devices. Each of them states the addressed deficiencies on the side of the beginning of the chapter and insights regarding the deficiencies and research questions at the end.

Part III Presents the extension of the framework to Nomadic Projection Within Extended Reach in the first section of Chapter 9 and two further case studies on applying the extended framework version to nomadic (and peripheral) interaction on-the-go, using the same structure as before.

Part IV The final part starts out with 12 design guidelines based on the lessons learned from the case studies (Chapter 11). Afterwards, in Chapter 12 a thorough prospect, both on directly related advancements to the framework, as well as on future applications of nomadic projection in light of the current trend of smart wearable devices is given. This closes the loop to the initial research question, how current deficiencies in nomadic information management can be addressed by the strengths of new technologies and which role nomadic projection may play in this new ecosystem in the future. Chapter 13 finally concludes the thesis, providing answers to the research questions previously raised by this chapter.

2

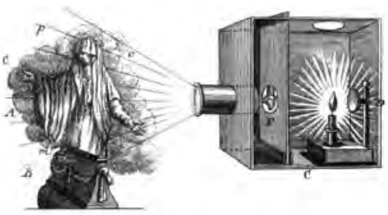
PRINCIPLES OF (MOBILE) PROJECTED INTERFACES

Projected interfaces have a long history in academic research dating back to the late 1980s. During these times, mobile (handheld) projectors as we think of them today (see their history in the next section) were not available. Nevertheless, many basic challenges of interaction with projected interfaces, which for the most part apply to mobile projection as well, have been investigated and solved. More recently, mobile projectors have become available and existing solutions for display, interaction, and feedback with (mobile) projections have been extended, adapted, or revised. This chapter will give an overview of basic research in these areas, introduce available devices for mobile projection, and directly as well as indirectly related interaction techniques. Each section will conclude with a prospect on how these principles have been applied throughout the projects presented in this thesis.

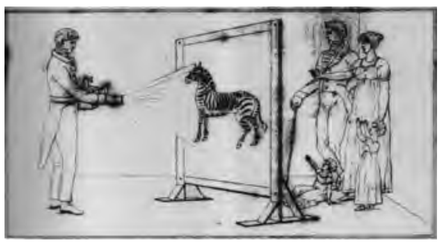
This chapter will further be complemented by the subsequent one, which will discuss and classify related works based on their direct relevance to the thesis' topic, i.e. Nomadic Projection Within Reach.

2.1 HISTORY OF MOBILE PROJECTION

Mobile projection has a longer history than one might think as it was already used in a comparable manner as the magic lantern (*laterna magica*) in Europe in the 17th and 18th century and as the *Utsushi-e* performance in Japan in the early 18th century [266]. A light source (candle, oil lamp) was placed behind a concave mirror which bundled its light towards a painted slide that would block some parts of light and let others pass through a magnifying glass (lens) to appear on steam (Fig-



(a) Inner workings of a magic lantern



(b) Performance with a magic lantern

Figure 2.1: Magic lanterns were used in Europe in the early 18th century as early forms of cinematic entertainment (images by Willis [265]).

ure 2.1a) or a canvas in front of an audience. Much like shadow puppetry, by changing the painted slide entirely or gradually, for instance by blocking a part of it, animated imagery could be created for the audience. More importantly, because the projectors were hand-held devices (or seldom times worn on a belt), imagery could be animated by moving the projection as well. Combining these techniques, large interactive animations like a horse galloping across a large canvas could be achieved (Figure 2.1b). By combining multiple projectors, interactive stories could be told much like comic scenes in later times. Willis [266] provides an overview and a modern adaption of this kind of storytelling [267, 269].

2.2 PROJECTED DISPLAYS IN CONTEXT

In recent years we have witnessed ever more mobile displays being introduced to the market. Starting from smartphones and later tablets, to now smart watches, smart glasses and handheld projectors. Most of these create the perceived image at different points in the visual pipeline between surface and human eye which inhibits different advantages and disadvantages and provide different abilities for AR, both of which this section will briefly discuss. We can distinguish whether...

VIRTUAL RETINAL DISPLAY (VRD) ... the display is formed directly on the viewer's retina, so no screen or surface as medium is involved. As such no light is lost during the transport and no pixel borders are visible which may lead to a better image than with other techniques (when comparing equal resolutions). Due to the short distance between emitter and eye, competing ambient light poses not much of a problem. As will be elaborated in the next chapter, a mobile projector requires something between 200 and 800 Lumens (lm) for a clear image in an office environment (typically 400 luminous flux (Lux), see Section 3.3). In contrast, a retinal display only requires around 0.25 lm, a thousandth, to achieve the same effect. This is owed to two differences:

First, although optical tricks make it appear several meters away, the display physically is at very short distance to the eye, which accounts for the majority of the difference. Second, while projections on typically diffuse surfaces reflect the light in all directions returning only about a seventh to the user holding the projector, virtual retinal displays can concentrate all light towards the user's pupil. Finally, the amount of desired ambient light passing through the glasses can be controlled, i.e. as done in sunglasses and even dynamically as implemented by the display prism of Google Glass. In theory, this speaks for a high superiority of this technique compared to mobile projectors. However, although researched for a long time, Virtual Retinal Displays (VRDs) are still in their infantry, with currently only one very expensive product by Brothers on the market [20] and another in pre-production [31]. As Hainrich and Bimber explain, that is because delivering the light stream precisely to the pupil and with a sufficient field of view is extremely challenging [101, pp. 457]. Optical see-through (cOST)–VRDs that are much more practical for nomadic use require the light to enter from the side to not block normal vision, which required complex optical setups and eye tracking to deliver the light precisely to the retina independent of the user's current direction of gaze. Existing systems such as the AirScouter by Brother [20] thus deliver only comparably small Field of Views (FOVs) of only 25°.

HEAD-WORN-DISPLAY (HWD) ... the display is formed on a screen directly in front of the user's eyes. HWDs (also called head-mounted-displays (HMD) or near-eye-displays (NED)) recently received a reincarnation through products such as Google Glass and Oculus Rift. Light modulation, fresnel lenses, prisms, or the display of defocused images trick the human visual system into thinking that the display is actually further away (2.4 m in case of Google Glass) although in fact it is very close to the eye (usually less than 10 cm) and could not be focused by the eye otherwise. A major problem with HWDs are the occurring visual rivalries that result from contradicting depth cues which the brain has constantly to balance and which can lead to eye strain and simulator sickness.

In monocular HWDs, *binocular rivalry* occurs because one eye sees and focuses on virtual content in mid-air, the other eye instead on something real in the environment and the brain steadily changes dominance between these influences [151]. *Accommodation rivalry* occurs because binocular HWDs usually mimic Stereopsis by displaying disparity images to each eye (stereoscopic display) that show the content slightly displaced to the left respectively right side to allow for depth perception as in natural 3D perception. Normally in the latter, the human visual system does two things simultaneously when looking at an object: (1) rotate (converge) the eyeballs to center the desired object, which results in a convergence point that lies in *convergence distance* to the eyes; (2) focus the eyes on the object by accommodating the lens to the *accommodation distance*; in natural perception, with correct eye-sight, convergence and accommodation distance correspond to each other (accommodation-convergence reflex). In contrast, when using an HWD the accommodation distance (lens focus) of the eye must remain fixed on the physical display to perceive a sharp image, whereas the convergence point is directed to "focus" at the distance of the object of interest. This mismatch is experienced in contemporary Virtual Reality (VR), AR and 3D cinema, leading to different degrees of motion sickness depending on the extend of induced rivalry, i.e. how far objects are virtually placed away from the projection screen and the overall depth (distance between farthest and nearest object) in a scene.

Apart from that, many of the advantages and disadvantages of VRDs similarly apply to HWDs: the small distance between display and eye allow for low Lumens output—although with HWDs around half the energy is lost to the display system. In exchange, the use of a display leaves more scope for light to enter the pupil, thus notably relaxes the constraints of VRDs. Still, existing optical see-through (OST)–HWDs which do not block perception of the real world and thus are much more applicable for nomadic interaction, support only field of views below 25°[1]. Both, VRDs and HWDs today are available as glasses or similar. As worn devices, they become part of the wearer's social appearance. Although devices like Google Glass are a big leap towards more unobtrusive devices, they are still very noticeable which may hinder their social acceptance. Further, it puts very high size constraints on the overall system to minimize its weight and bulkiness which also significantly limits battery power. On the other hand, the glasses form factor allows for the display to be quickly glanced at and provides short lead time to interaction and hands-free operation.

BODY-WORN-DISPLAY (BWD) . . . the display is worn somewhere on the body, e.g., as a smart watch or on a skiing jacket. Aligned between HWDs and handheld displays, these support short lead time to interaction and similarly quick glanceability like the former ones (as for instance studied by **harrison_wherelocatewearable_2009**, but not their capabilities regarding hands-free operation, AR and as always-available display.

HANDHELD DISPLAY . . . the *display* (thereby excluding handheld projectors) is on a handheld device as is standard for most contemporary nomadic devices (smartphones, tablets). Handheld devices typically provide Lumens in excess of 3000, yielding around 4500 lx in a typical viewing distance of 35 cm. They can be used for video see-through (VST) display by displaying the camera stream and overlaying it with additional digital information, albeit the wide-angle cameras of mobile devices and the monocular capture and display do not allow for a realistic see-through experience. Another disadvantage of handheld devices is, well, that they are handheld, occupying the hands that cannot be used for other tasks and leading to arm fatigue when held for a longer time. Cranes to alleviate this problem as we have added in [W14] are not available in nomadic scenarios.

PROJECTED DISPLAY . . . the display is formed directly on a surface in the environment, either to augment the surface such as an object with digital information (AR), to create a Graphical User Interface (GUI) to interact with digital information or mixed and related forms by using interaction metaphors like Motionbeam metaphor and Spotlight metaphor as described in Subsection 2.5.4. Previous points have already described the inferiority of mobile

1 Google's Glass 14.7°, Epson's Moverio 23°

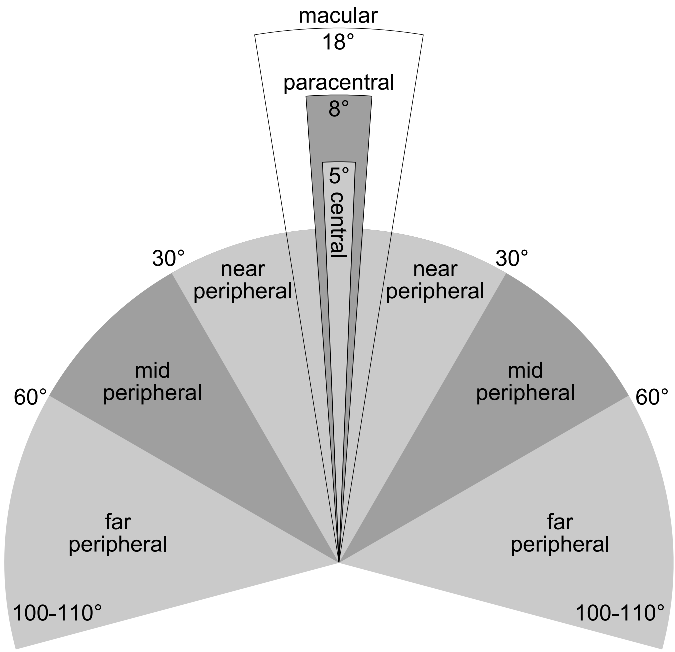


Figure 2.2: Human visual field. Image from [196]

projectors regarding brightness that make it unsuitable for some scenarios (Chapter 5 discusses this in detail). On the side of advantages, handheld and worn projectors provide a large field of view that comprises that of human near-peripheral vision (60°) that supports adequate color vision [196]. By using short-throw lenses, throw ratios below 0.3 can be achieved allowing for a FOV beyond 100°that is able to exploit a lot more of human's full 200-220° visual field (cf. Figure 2.2). Naturally, by steering a handheld or worn projector across the space, even areas beyond human's visual field can be leveraged as has been shown in the AMP-D study (Chapter 9). This is an advantage over other techniques which are bound to the user's arm span for handheld displays or would require for now impossible dynamic transformations of the optics used for retinal and head-worn displays. Another advantage of projected displays regarding AR is that because the display is generated at the object of interest, it shares the same environmental influence (ambient light, medium of light transport) as the augmented object, leading to much more believable augmentations than other techniques provide out of the box. For the same reason, visual rivalries (e.g., different focal planes) as discussed for previous systems do not occur. Finally, projected displays inhibit the greatest flexibility as they can appear independent from handheld or worn devices and can be integrated in other devices to provide them with an (additional) display.

Chapter 12 on future work will discuss possible implications of these differences and provide a more formal comparison (cf. Table 12.1).

2.3 CONTEMPORARY MOBILE PROJECTION TECHNOLOGY

Mainly three aspects in mobile projection technology affect its applicability to nomadic projection: the size of the projector (significantly affected by the battery size), the amount of light output (brightness) which are quasi coupled, and the mechanics used to form the image, all of which will be discussed in this section. Note that the focus on *nomadic* projection demands only battery-powered projectors to be considered in this section.

2.3.1 Image generation

At the heart of the projector, a light source and a Micro Mechanical System (MEMS) form the projected light beam. As light sources, for mobile projection only LEDs and lasers are used while incandescent lamps provide theoretically more Lumens but are not energy efficient enough (also leading to lots of unwanted heat dissipation) for mobile battery-powered use.

For all LED-based systems, a MEMS arranges it so that some of the emitted light of red, green, and blue LEDs is blocked and other light is let through to the lens of the projector which focuses the beam. As MEMS it is either used a Digital Micromirror Device (DMD), which consists of millions of nano-sized mirrors which each reflects the light to either pass through or to land on a light absorbing surface. The other, not very different option is to use a Liquid Crystal on Silicon (LCoS) chip, where each pixel is able to either block light or let it pass through to a reflective surface behind it which mirrors it to the projection lens. The LCoS technique is 50% less light efficient due to the necessary light polarization compared to the DMD technique, but allows for higher resolutions because the pixels can be packed more tightly than the mirrors. LED-based systems require a focus lens due to the high etendue of the emitted light. This requires the user to manually adjust the focus every time the projector is moved—which is impractical in mobile scenarios—or an automatic focus mechanism, which has been added in the AMP-D project (Chapter 9) but has not been available in consumer devices until the recently announced ZTE Spro 2 [286].

Laser-based systems, on the other hand, which feature a red, blue, and green laser module, can utilize DMD and LCoS MEMS as well with different pros and cons compared to LED-based systems. Speaking *for* laser light is its high etendue which, although a lens is used to form the image, provides always-in-focus images between several centimeters (it was 20 cm for AAXA's L2) and infinity. Moreover, the focus-free system does not only abandon a focus dial, but also allows to project at acute angles (which play a major role in nomadic use cases as can be seen later) and on non-planar, complex surfaces. Speaking *against* laser light is the so-called laser speckle that the highly focused light beams create on the surface and which is more uncomfortably to look at. Further on, as laser light is potentially more dangerous to the eye than diffuse light, selling and using laser-based systems is regulated by laser classes which limit its brightness. Because laser-based systems utilizing DMD or LCoS project the whole image at once, the laser light's intensity is split across the whole picture which renders several hundreds of Lumens light output still acceptable to not compromise eye safety.

Unfortunately, this is another story altogether for a third MEMS technique that is only applicable to laser-technology which is Laser Beam Steering (LBS). This uses a MEMS to directly scan the laser very quickly over the whole image, similar to how Cathode Ray Tube (CRT) monitors used to create an image. Our comparably phlegmatic visual system recognizes this as a single picture. But in the event of a failure, when the scanning process stopped in the very unlikely but possible event where a user looked directly into the motionless laser beam, its entire power could fall on a spot of the user's retina and damage it. To avoid that, the overall exposure to laser light until either the corneal reflex kicks in (roughly after 250 ms) or the person looks purposefully away, must be limited and the safest way to achieve this is to limit the maximum light output in the first place. Devices conforming to this specification fulfill the requirements to be classified as laser class 2 according to DIN EN 60825-1 or DIN VDE 0837 respectively. More precisely, they do not expose more than 1 mW of power and only across the visible light spectrum between 400 and 700 nm. In Germany, laser devices classified as class 3 can be brought to consumer market under certain conditions but to sell the product internationally, manufacturers have to settle on the least common denominator which currently is class 2 (cf. [152]). In a recent study conducted by the Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) only 17% of participants reacted within the 250 ms interval with the blink reflex when a laser class 2 beam was directed to their eyes [159]. That said, it is unlikely that these safety regulations will be softened in the near future. Software-based protections like presented by Kaufman et al. [135] have a too high latency to reliable aid eye-safety. Thus, LBS-based projectors are not allowed to provide more than 25 Lumens—a quarter from what other techniques achieve at a similar size, drastically limiting their applicability. Although, in this light, non-LBS laser based systems seem superior, it is for economic reasons that currently only the inferior LBS-based laser projectors are to be found on the market (and have been used frequently in the studies of this thesis).

2.3.2 Types, Sizes, and Their Light Output

Commercially available hardware is subject to quick change, thus this section will only provide high-level categories of available hardware

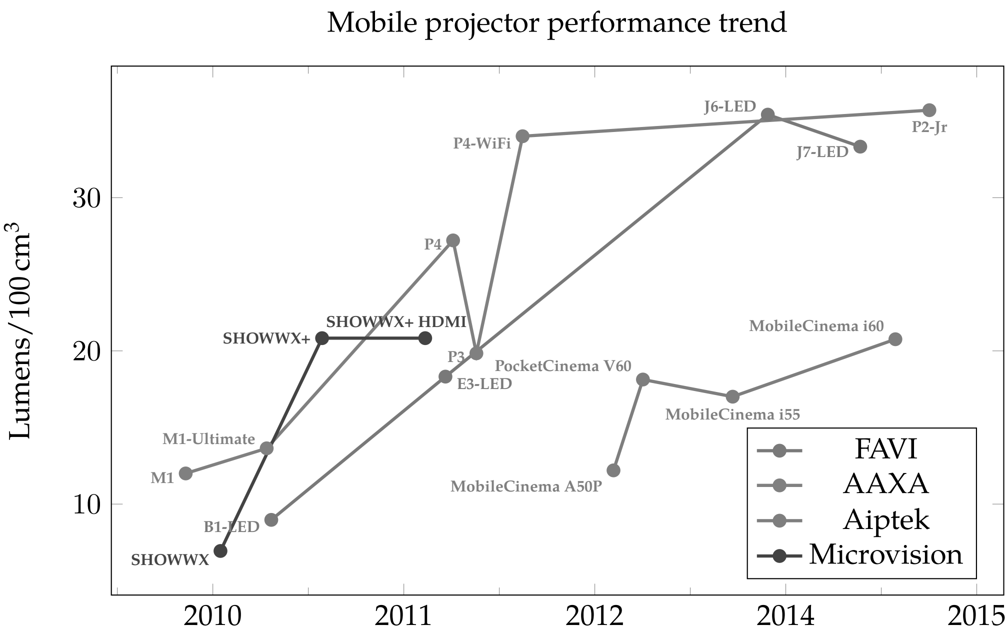


Figure 2.3: Advancement in energy efficiency of mobile battery-powered projectors, showing a linear trend of around 10 Lumens/100 cm^3 performance increase per year.

whereas Chapter 5 will reflect on these categories regarding their application to nomadic interaction.

The most important distinction to make regarding nomadic projection are the size of the projector (as it decides whether the projector can easily accompany the user throughout the day) and its light output as it primarily decides *when/where* the projector can be used. As all battery powered projectors are designed to provide 1.5–2.5 hours of continuous use (to allow for watching a movie), higher brightnesses lead to larger projector sizes to fit a larger battery (and sometimes additional elements for cooling) which makes these two aspects interdependent. Thus, besides Lumens/Watt as the basic performance metric for light energy efficiency, this renders Lumens/cm^3 a good performance metric from a usability perspective. In particular, the user does not care whether the small size of the projector is more due to a better energy efficiency in emitting light, because of a better packing of the components, or because of a more efficient built-in battery. Figure 2.3 depicts the advancement in brightness efficiency in Lumens/cm^3 of pico projectors. Mainly we can distinguish three classes of available hardware for nomadic computing, from larger to smaller: *pocket projectors*, *pico projectors*, and the *pico engines* themselves.

Pocket projectors, despite their name, are best defined as fitting just about in a (large) user's palm (see Figure 2.4). They weigh up to a kilogram, measure up to 2200 cm^3 and reach up to 400 lm (e.g. AAXA M4). Their size and weight make them less suitable for handheld usage or worn scenarios but they may be integrated in larger mobile devices such as tablets, digital (video) cameras, etc. Furthermore, for re-

Figure 2.4: Dell m110 pocket projector (300 Lumens, 1280px × 800px) that was used in the AMP-D project. Image courtesy of Dell.

Figure 2.5: The Microvision SHOWWX accessory series (SHOWWX 10 Lumens shown) using Microvision's PicoP engine (top left). The SHOWWX+ HDMI (15 Lumens) was used in all projects presented in this thesis except for the AMP-D project. Some projects used a stripped down version (middle left). All images courtesy of Microvision.

searchers they may act as a realistic prediction of next generation (5 years ahead) pico projector performance to explore their applicability to new domains (as was done in the AMP-D project in Chapter 9).

Pico projectors can be operated in one hand and typically weigh between 0.1 and 0.3 kg (1-2 times the weight of a smartphone), have volumes of 130 (Rif6 Cube) to 250 cm^3 (AAXA P4) and reach up to 100 lm (Favi J6-LED-Pico). They come as pure peripheral devices that take a video signal as input (see Figure 2.5) or as smart projectors, which run an operating system that allows them to provide streaming functionality over Wi-Fi and to project pictures and videos of attached USB devices. As the latter is, besides the missing screen, not very different from a smartphone, we have also seen smartphones that integrated projectors, so called projector phones (e.g. the Samsung Galaxy Beam, Figure 2.6) and other smaller devices such as compact cameras.

Lastly, although not a category of end-user products like the ones before, the size of available optical **pico engines** which are at the heart of pico projectors, can tell us about the minimal size requirements for mobile projection devices. In some scenarios, for instance, components like battery, controller boards and chips, and cooling measurements can be left aside because the integration with an existing device already provides for all of these. On the other hand, 1.5–2.5 hours continuous use may not be required at all for many scenarios different to watching a movie. For the time being, the market is mainly dominated by two pico light engines: the DLP engine by Texas Instruments with its smallest reference design (TI LightCrafter EVM) consisting of RGB LED and DMD chip, delivering 20/50 lm (without/with cooling measures) at a volume of 25 cm^3 (only a fifth of the smallest available product featur-

Figure 2.6: The Samsung Galaxy Beam projector-phone and its projection under different ambient lighting conditions (also see Subsection 2.4.1). It features a 15 Lumens projector, its successor the Galaxy Beam 2 (not shown) a 20 Lumens projector. Images courtesy of Samsung.

ing this engine). Similarly, the laser-based PicoP engine by Microvision delivers 25 lm at a size of only 5 cm^3. The end-consumer product of Microvision that uses this engine, the SHOWWX+, which was used in many projects presented in this thesis, has a volume of 144 cm^3. Again, like with Digital Light Processing (DLP), this shows the room for improvement beyond the scope of light sources and imagers.

The majority of all these projectors currently supports resolutions about HD ready (720p) with a transition to full HD (1080p) being imminent.

2.4 VISIBILITY OF THE PROJECTION

Different to other display systems, projections require a lot of adaption to the environment both by the system as well as the user to become a projected interface with satisfying visibility. Depending upon the setup whether it is more static or more mobile, responsibilities shift between the system and the user, but overall, mainly four aspects of the projected display have to be accounted for: Projection Distance, Position and Surface Selection, Geometric correction, Visual Compensation, and Focal Correction.

2.4.1 Projection Distance

Firstly, and most relevant to the remainder of this thesis, the projection distance not only decides about the size of the projection (together with the projector's throw-ratio), but more importantly about its luminance. The strong rivalry of projected displays with ambient light has already been touched upon. This section provides a more formal view of it. Speaking about luminance of projections, some terms have to be defined:

LUMENS (LM) The light that is emitted from a light source in a given direction (angular span).

ANGULAR SPAN (STERADIANS) OF A LIGHT SOURCE

$$\Omega = 2\pi \left(1 - \cos \frac{2\theta}{2}\right)$$

The angular span Ω is calculated from the apex angle θ that is defined as twice the 2D angle between the angle with the highest luminous output, and the angle where this output is reduced to only 50%. For projectors, the apex angle is easily calculated from the throw-ratio (tr) as $\frac{45^\circ}{tr}$.

ILLUMINATION (LUX) The amount of light that illuminates a surface. Light of 1 lm, perfectly illuminating a surface of 1 m^2, yields 1 Lux (lx) on the surface.

LUMINANCE (CD/M^2) To be able to compare the brightness of projected displays against those of screen displays, the measure of luminance and their relation to illumination is interesting as well. Luminance is the amount of light emitted from the display and is measured in candelas per square meters (cd/m^2). Assuming a perfectly diffuse reflecting surface, the relation between cd/m^2 (also called "Nits") and Lux (lx) is given by

$$Lux = Nits \times \pi$$

As ambient light is measured in Lux, we can now calculate the required Lumens a projector must provide to at least match the ambient Lux from a chosen distance or the size of the targeted area, respectively. From previous equations it follows

$$\text{ANSI Lumens} = \frac{\text{targeted Lux on screen } (lx) \times \text{screen area } (m^2)}{\text{screen gain}}$$

Because the screen area's width is equal to the projection distance d divided by its throw ratio tr; further the screen's height equals the width multiplied by the inverse aspect ratio; and screen gain must be ignored[2] as it occurs arbitrarily and uncontrollable in nomadic environments, we can derive the following equation which is more directly applicable knowing the characteristics of a projector:

$$\text{ANSI Lumens} = alx * \left(\frac{d}{tr}\right)^2 * \frac{arh}{arw} \tag{2.1}$$

with d = distance, tr = throw ratio of the projector, alx = ambient Lux to match, arw = aspect-ratio width, arh = aspect-ratio height.

2 Projection screens typically provide a gain around 1.3 by reflecting more light back to the projector's origin and an ordinary diffuse white surface around 1, i.e. neither a gain nor a loss. As even the latter is unrealistic to presume in nomadic environments, the calculated values must be taken as best case estimates.

However, this equation only provides the required Lumens to *match* ambient light, which is not sufficient as the ambient light still diminishes the projection's contrast which is exemplified by the following example: we assume a 50 Lumens projector with a very good contrast of 5000:1 (which LBS projectors achieve) roughly providing $100\,lx$ in a coffee house with typical ambient lighting of $100\,lx$. Then, the brightest point becomes $100\,lx + 100\,lx = 200\,lx$, the darkest $100lx + \frac{100\,lx}{5000} = 100.02\,lx$, leaving only a contrast ratio of $\frac{200\,lx}{100.04\,lx} = 2$ which is not sufficient. According to the effective ISO [124] and American National Standards Institute (ANSI) [26] standards, the absolute minimum of contrast ratio is already higher with 3:1. Most of all, these standards only consider the bare minimum required for text to be discernible at all, for instance as basis to Web Accessibility Guidelines [254]. For pleasant user experiences, we have to consider our typical indoor surroundings that already provide ratios above 20:1 and the screens that we are accustomed to, for instance the screen of an iPad 2, which adjusted to full brightness achieves not less than 47:1 even in a much brighter 1,000 lx environment [249] (not only because of its bright display but mainly because its screen does not reflect more than a tenth of the ambient light). This may explain why a recent ANSI/Infocomm standard [121], especially developed for contrast ratios of projected displays, defines **7:1** as a target for *passive viewing* (e.g., following a simple PowerPoint presentation) and **15:1** for *basic decision making* (making sense of a complex graph or spreadsheet).

Two things can be taken away from these considerations: First, the innate contrast ratio which the manufacturer measures in a completely dark room, has almost no effect in surroundings of considerable ambient lighting. Second, only matching the luminance to ambient light results in poor-contrast images. This can be of an advantage if the projected image was only to augment reality in a very believable way. But for mimicking a display for presentation, entertainment, or information management it is not sufficient. A general rule in selecting projectors for theaters is thus to choose their Lumens such that the luminance of the projected image at least doubles that of ambient light. For presentations, the fourfold is advisable and in the previous scenario would boost the contrast ratio from 2:1 to at least an acceptable 8:1 (slightly above the lowest task level of the aforementioned ANSI/Infocomm standard [121]).

Based on these considerations, Table 2.1 lists the required Lumens—both low contrast for passive consumption as well as optimal (4×) for information management—in typical everyday environments.

At first, the enormous range of Lumens strikes the eye, ranging from ≈3 lm for supporting video consumption at a coffee place to almost 150,000 lm to support vivid presentations in full daylight at a distance ≥1.5 m. Looked at more closely, we further see that even in comparable low-light environments (400 lx office), above a distance of 1 meter we require a projector featuring between 272 lm and 1089 lm, which

Environment	Lux (lx)	Distance (m)	Projection size (diagonal inch)	Min. required Lumens (lm)	Optimal Lumens for vivid display
Full daylight/ Direct sun indoors	35,000	1.5	≈61	36,610	146,440
		1	≈41	23,821	95,284
		0.25	≈10	1018.5	4074
High ambient/ Overcast day	1000	1.5	≈61	1046	4184
		1	≈41	680.6	2722.4
		0.25	≈10	29.1	116.4
Office/Sunset[3]	400	1.5	≈61	418.4	1673.6
		1	≈41	272.24	1088.96
		0.25	≈10	11.64	46.56
Living room	200	1.5	≈61	209.2	836.8
		1	≈41	136.12	544.48
		0.25	≈10	5.82	23.28
Coffee place	100	1.5	≈61	104.6	418.4
		1	≈41	68.06	272.24
		0.25	≈10	2.91	11.64

Table 2.1: Minimum required Lumens for low-contrast passive consumption and high-contrast information management (×4) under different ambient lighting conditions and between *out of reach* and *within reach* projection distances. Calculations based on Equation 2.1, assume a throw-ratio of 1.1 (based on Microvision SHOWWX+ HDMI), 16:9 aspect ratio and no screen-gain (typical nomadic surfaces will likely even have a screen-gain below 1 as they are not perfectly white, thus these numbers must be taken as best case estimates regarding the projection surface).

is a magnitude more than what mobile projectors currently offer (cf. Subsection 2.3.2).

These considerations, together with the steady yet conservative (10 Lumens/cm^3 per year according to Figure 2.3) trend in advancements of mobile projector technology, strongly promote applying nomadic projection in a *within reach* distance, rather than an out-of-reach distance. Because of that, the next chapter uses the interaction distance as a discriminating factor among related works. All the same, the competitive luminance of within-reach projections serves as strong motivation of the later presented Nomadic Projection Within Reach framework.

2.4.2 Position and Surface Selection

Projections typically are used where traditional screens do not easily fit in terms of size, fixture or required mobility. Because of that, except for the standard office presentation scenario, it is unrealistic to expect a perfect projection surface delivering optimal reflectance in a position that is optimal for all interested viewers. Thus, trade-offs have typically to be made.

For simple planar projections, Siriborvornratanakul et al. [237] present a system that is able to automatically select the largest uncluttered planar area within a cluttered projector's FOV in a static setup. Handheld projectors are often steered manually, but techniques have been presented to automatically adjust the position based on privacy impacts [67] using a motor-steered mirror. Such automatic projection movement has also been shown for static setups [272].

Given sophisticated pre-warping of the image (see next section), non-planar surfaces can be used for projection as well, which drastically increases the number of available projection areas. When the projection surface is moving, projections can still appear static given a fast enough tracking and projection system [191] or a successful motion prediction [144].

An issue regarding positioning the projection is that people looking into the projector beam are getting blinded. With static projections, this typically only happens because of improvident behavior of the viewers and is less of a problem. With mobile projection, however, this becomes a social problem when people are actively blinded through the movement of mobile projectors. In case of laser projectors, this might even escalate to a medical risk. This can be avoided by intelligently combining multiple projectors in fixed scenarios [248] or suppressing parts of the projected image in mobile scenarios [135].

Most projects presented in this thesis have taken a hybrid approach to surface selection, depending on the usage scenario: devices like the Penbook (Chapter 6), SurfacePhone (Chapter 7), and AMP-D (Chapter 9) prescribe the general surface to use (i.e. back of lid, table, floor and wall) but allow free positioning of the projection by moving the device itself or the body in case of AMP-D respectively). Devices presented in chapters 8 and 10 instead are handheld or –worn and thus require more active surface selection.

2.4.3 Geometric correction

For a projector to be able to create an image that is larger than the size of its imaging unit, the projected image leaves the projector in the form of a cone. The throw-ratio of the lens will decide upon the exact proportion between distance to the projection surface and the resulting size of the image. Generally speaking, the further the distance the larger the image and the majority of people seems to be familiar with this concept through their experience with spotlights or setting up presentations.

Naturally, this relation is not only true for the image as a whole, but for each individual ray of light. Consequently, when the projected image reaches the projection surface at non-orthogonal angles, the projected image is distorted, with parts of the image closer to the projector appearing smaller and others farer away appearing larger than the center of the image. As a result, the resulting image is not only distorted from a rectangle into a quadrilateral but pixels are also shifted towards or away from the center of the image (projective transformation). This is important to note as it means that the resulting 2D image cannot be described or corrected by a simple affine transformation. Raskar et al. [208] and Raskar et al. [210] and many of the works cited in this section describe the underlying math to geometrically correct for handheld projection. It is not completely trivial, as first the inverse of the homography between projector and projection surface has to be computed and multiplied on the vertices of the projected image texture and then the counter-distorted image has to be fitted within the largest inscribable rectangle (of the projector's aspect ratio) of the projector's image plane in world coordinates [210]. The latter results from the fact, that only too large projected parts can be made appear smaller, but smaller parts cannot appear larger at the same position as they already fill the entire projection plane. When the Spotlight metaphor (Subsection 2.5.4) is applied, the second part (fitting to the projector's bounding box) of the previously described process can and should be left out. Fortunately, it is then almost sufficient to leverage the standard capabilities of 3D graphics engines. As the occurring process is the inverse of the perspective foreshortening performed by the human vision system, or more generally speaking, that of any 2D camera creating a 2D image of a three dimensional observation, correcting an arbitrary image to appear undistorted on a planar projection surface requires the image only to be projected as seen from a virtual camera placed at the exact position and rotation of the projector in the scene and looking at the scene (e.g. the ground) using the same FOV as the projector. This approach was used, e.g., for the AMP-D project. To receive perfect results, the so called intrinsics of the projector must be taken into the computation as well to account for possible lens aberrations as well as a possible off-axis alignment of the projector-lens.

Raskar et al. have been one of the first to automatically pre-warp the projected image such that it appears undistorted on the projection surface [210]. They equipped a projector with a camera to allow for automatic calculation of the projector intrinsics and two tilt-sensors to compute roll and pitch of the projector against gravity. The solution was suitable for planar wall projections and required a separate calibration phase. Raskar et al. later extended the concept to curved surfaces and an automatic alignment of multiple projectors to form a single image without discernible transitions [208]. While these works required a separate calibration phase, Dao et al. presented the idea of a semi-automatic calibration that requires the user to start from an orthogonal, undistorted projection and press a button for a subsequently automatic correction [81]. Less computationally expensive but at the same time less flexible is to rely on fiducials in the environment to denote rectangular projection areas [40]. Correct projection on arbitrarily shaped objects has later been presented by Sugimoto et al. [246]. Systems that have access to a surface mesh of the environment can adapt the content to arbitrary surfaces [129, 272] and progress towards mobile solutions using Simultaneous Location And Mapping (SLAM) has been made [176].

When two people are to interact by means of projected interfaces, *dyadic projection* concepts exist that maintain a correct view for two people on arbitrary surfaces [42, 45]. Finally, geometric warping is also used to project slightly offset real textures on top of the real environment to create special effects like shaking of the surroundings after an explosion [128]. This effect is also heavily used in arts (see Subsection 3.2.2.1).

Related video

The projects presented in this thesis have taken different approaches to geometric correction as demanded by the respective use cases. This ranges from leveraging infrastructure calibration (IPC), over one-time manual calibration (Penbook) and continuous semi-automatic calibration (SurfacePhone[4]) to automatic calibration (AMP-D prototype). Apart from AMP-D that could directly leverage a 3D engine as described before, other projects in this thesis followed the already mentioned approach described in Raskar et al. [210].

2.4.4 Visual Compensation

Besides geometric correctness, projections heavily depend on the surface they are cast upon. Important aspects to distinguish are, obviously, color and visible structure, but also micro-structure and content of the surface.

2.4.4.1 *Radiometric Compensation*

In terms of color it is often thought that white surfaces are most suitable for projection but that is only half the story. White surfaces not only reflect the projection very well but also the surrounding ambient light, even the ambient light that is created by the projector itself in the environment. This means that deep black colors are often hard to achieve on white screens and thus gray screens are often favored over white ones, for instance, in cinemas where very bright projectors are affordable and a high contrast is desired. In mobile scenarios, gray surfaces reflect less ambient light and may therefore provide a better contrast and viewing experience than white ones in certain situations.

4 Penbook and SurfacePhone would not require calibration when sold as a commercial product in a pre-calibrated rigid body without any moving components.

When the projection surface contains color, the additive light transport will lead to color blending and the colors and intensities of the projected image will look distorted. Bimber et al. demonstrate how projected images can be adapted to almost cancel out the underlying image or structure or reveal only certain parts of it respectively [49, 51]. Later Grundhofer and Bimber show how this offline technique can be extended to real-time radiometric compensation [96] that tries to maximize the clarity of the projected image on planar surfaces of arbitrary color and structure. Of course, the level of possible adaptation directly depends upon a superior brightness of the projector over the existent ambient lighting.

Oftentimes, the perceived clarity of a projection depends especially on the perception of projected edges. Sajadi et al. show how to drastically improve edge appearance by overlaying a hi-fi edge image over the standard image, either time-multiplexed using one projector [218] or using multiple projectors simultaneously [23].

2.4.4.2 *Diffuse and Reflective Materials*

Regarding the micro-structure, projection surfaces can be distinguished in diffuse, reflecting, and retro-reflecting materials. Most surfaces in our environment are rather diffuse, i.e. they diffuse incoming light in all directions. This allows for a large audience since the projection can be viewed from a large range of angles. However, if there is no audience, at least not across the whole range of reflected angles, much of the light is lost. Reflective materials in contrast, act more like a mirror[5] which means that they diffuse and reflect the incoming light only around the incident angle. Thus, if the viewer is standing, for example, in a line with the projector, almost all light from the projector is reflected to the viewer and almost no light from ambient light sources, rendering a much brighter image than from diffuse screens. Reflective materials typically contain metallic elements, for instance, early cinema has used silver-coated screens to increase the brightness of the images in a time where projectors did not deliver sufficient brightness even for dark rooms. Today, silver screens are used to preserve the light's polarization for 3D projection with passive 3D glasses. Finally, retro-reflective materials are synthesized that reflect incoming light back to its origin, no matter what the angle of the incident light is. They are heavily used in transportation to make road signs and pedestrians visible at night by directing the light of headlights back to the car. For projections, they offer a very high screen gain from arbitrary angles as long as the viewer and projector are close together. Krum et al. have exploited this effect to present mobile projection in daylight environments [148] which is impossible with current mobile projection technology and diffuse surfaces. Unfortunately, both reflective materials and retro-reflective materials are not commonly available in the

5 A perfect mirror is unsuitable for projection as some form of light diffraction is required to make the image visible to the viewer.

environment to allow for a general mobile projection strategy but instrumented environments may exploit their advantages.

2.4.4.3 *Three-dimensional Perception*

To project three-dimensional content, three approaches exist. The traditional one separates the image to different color channels and allows for a three-dimensional perception when viewed through corresponding glasses that filter out one of these colors to present different images to each eye. Although this technique distorts the color perception it neither poses requirements on the projection surface nor the network infrastructure and thereby enables a mobile application as purely optical technique as shown by Chehimi [71]. Other techniques familiar from the cinema require either active shutter glasses which block one eye to show each projected picture only to one eye at a time and which require a high synchronization between the glasses and the projection system. Or 3D projection using orthogonally polarized light where images for left and right eye are polarized differently and passive glasses with corresponding polarization filters ensure that each image reaches the correct eye of the viewer. As said before, this technique requires a special projection surface that preserves the polarization and therefore precludes this setup from mobile applicability as well.

While no color-maintaining, glasses-free technology for nomadic 3D projection exists, geometric and radiometric compensation should probably be employed by any nomadic projection device—at least if it dealt with information presentation. But so far we have only seen geometric, but no radiometric compensation been applied which might be explained by the following reasons:

- Radiometric compensation requires a fair competition between the brightness of the projector and environmental light to achieve believable results. It has been presented with projectors offering thousands of Lumens in environments with controlled lighting (darkened rooms) or instrumented surfaces [51]. In stark contrast, current mobile projectors still do not offer more than a hundred Lumens in totally uncontrolled environments. This renders the possible effect of radiometric compensation almost neglectable.
- In addition, truly mobile systems are automatically much more resource-constrained. While the geometric compensation already introduces a critical delay to the interaction loop, radiometric compensation would require an additional image correction on a per pixel-basis which would add another detrimental latency to the interaction fidelity of the system.
- Finally, radiometric compensation requires the system to feature a camera and the camera view to fully overlap the projected area, both not requirements of mobile projection systems per se.

Consequently, it must be concluded that radiometric compensation is strongly desired for nomadic interaction with projected interfaces, especially since tailored projection surfaces such as white, gray or reflective ones are rarely available. However, different to geometric compensation, radiometric compensation is too computationally expensive for the small possible visual effect that currently available projector technology would allow for. As in related work, I therefore decided to refrain from applying radiometric compensation throughout the research presented in this thesis.

2.4.5 Focal Correction

Another important aspect to achieving a good visibility of the projection is the control of focus of the image. As Subsection 2.3.2 has already laid out, laser-based projectors provide an alway-in-focus projection at the expense of being more constrained by health-safety regulations. Other light types require the lens system to focus the light on a specific focal distance. Typically this spans not more than a small range up to a few tenths of centimeters. In controlled environments, projectors with a very small aperture can be used to increase this range up to a meter as utilized by CastAR[6]. Other approaches include preceding image compensation to diminish defocus blur within a small range, shown for static setups [56, 195] and later for dynamic setups by Naoki et al. [109]. Another option, although only feasible for static scenarios, is multi-focal projection using multiple synchronized projectors with different focal lengths [50].

Because of the partly acute projection angles and sometimes repeatedly mirrored light paths occurring in nomadic projection, many of the later presented projects relied on laser-based projection to circumvent focus correction in the first place. Where laser-based systems were not an option, an automatic focus control for the projector has been developed to support focal correction (for AMP-D in Chapter 9 and for UbiBeam in [W3]).

2.5 PROVIDING INPUT

Unlike other mobile displays that are used today (smart phones, smart watches, tablets, etc.), projected interfaces neither come nor are easily equipped with a touch sensing layer. Thus, a variety of interaction techniques were presented in the past. As a plethora of research has investigated interaction with *static* large projected screens, in the following we must therefore limit the scope of related work to providing input to *mobile* projections, including interaction with handheld, worn or at least portable projection systems. Rukzio et al. [216] identified *Input on*

6 http://castar.com

the projector, *Movement of the projector*, *Direct interaction with the projection*, and *Manipulation of the projection surface* as four building blocks of interacting with mobile projections. An additional type of interaction that has emerged are Pointing & Gestures performed in front of the projector in mid-air. Together with Device Motion and direct Touch interaction, these are the most important interaction techniques as far as related to this thesis and will thus be the focus of the next sections. Similarly, certain recurring interaction metaphors that have emerged will be discussed. These subsections will focus on the enabled interactions from a user perspective, whereas the subsequent section Section 2.6 will then discuss the required tracking technologies from a system perspective.

2.5.1 Pointing & Gesture

Distant pointing at projections has a long tradition since people usually use finger pointing or laser pointers to point at certain parts of a slide to which they are referring verbally. Mistry et al. have presented simple gesture tracking using colored markers for pointing and gesture interaction with mobile projections [173]. Cowan et al. used shadow casting for multi-user remote interaction with mobile projections [78]. Molyneaux et al. use a mobile depth camera to support mid-air annotations and physical interaction with projected objects by finger shadows [176]. Pointing and Gesture interaction has been used for the AMP-D prototype (Chapter 9).

2.5.2 Device Motion

Interacting with the projection by moving the device is one of the most widespread interaction techniques with mobile projections, possibly due to its ease of implementation. Inertial sensor units, delivering up to 9 Degrees of Freedom (DoF) by means of accelerometer, gyroscope, and magnetic sensor, often come on-board on projector phones—whether as one device or a stack of phone and projector accessory—or can easily be attached and interfaced via Bluetooth. Inertial measurements provide the device’s orientation against gravity and magnetic north, respectively. Given a pre-calibrated or established reference frame to the projection surface, the exact orientation of the projection device against the projection surface can be calculated. This can then be used to prewarp the projected image (using the inverse orientation) for scale-invariant geometric compensation against a two-dimensional planar surface, as explained earlier (Subsection 2.4.3).

Device motion can further be used to provide simple commands to the system, for instance, by performing small gestures a ball can be thrown in a game as we have proposed in [W10]. Large gestures, without a stabilization as for example supported by the soon explained Interaction

Metaphors for Implicit Interaction, are suboptimal because the projection moves randomly across the room. With PICOntrol Schmidt et al. [228] show one of the rare use cases where larger gestures are useful as they allow to remotely control appliances which are beneath the projection. Within the scope of this thesis, explicit device motion has only been applied to the SpiderLight in Chapter 10, which is the only handheld device presented in this thesis.

2.5.3 Touch

Touch interaction with projectors usually refers to the idea of performing touch interactions directly on the projected display similar to interacting with a touchscreen. This has to be distinguished from using the touchscreen or touchpad of the projector itself as provided by many commercially available products like Samsung's Galaxy Beam, or the FAVI A3-Wi-Fi, which by indirectly steering a cursor mimics pointing but not touch interaction.

For mobile handheld projection, touch interaction implies that the distance between projector and surface cannot exceed an arm's length, at least when performed by the same user. Thus it would result in a small projection area which in multi-user applications, in addition, is partially occluded by the user. Therefore, touching projections coming from *handheld* devices is largely ignored in research.

HANDS AND FEET More often, touch interaction is employed with body-worn projectors. With OmniTouch Harrison et al. present an example of the former, where a shoulder-worn Projector Camera System (ProCamS) allows multi-touch interaction on the user's palm and arm, planar objects (e.g. a sheet of paper) held in hand, and nearby surfaces at arm's length. A larger camera FOV would allow for additional gestures [105]. In contrast, worn sensors such as Electromyography (EMG) (cf. [224] and the Myo sensor [180]) or acoustic [106] sensors can render a camera unnecessary. Nonetheless, not only the upper body has been explored for touch interactions but also foot and toe interaction [30, 166] and other parts for interaction with floor-projected [54] interfaces have been studied.

OBJECTS DEPLOYING PROJECTORS The other discernible trend is to include projectors into mobile devices which stay fixed during interaction but provide an interactive projected display. Wilson et al. present a mobile system of the size of a small suitcase that creates a table-sized touch-interactive display using shadow tracking of infrared images [270]. Cai et al. achieve similar results without additional IR illumination [60]. Bonfire augments the periphery of a laptop computer with small virtual screens and object augmentations and allows for simple touch operations [133]. Linder et al. propose an autonomously moving desk lamp equipped with projector which projects a touchable interface. LampTop [22] improves touch detection in this scenario.

PENS AND STYLI Touch interaction with projection has also been investigated using pens and styli. Song et al. proposed a pen equipped with a projector for paper annotations [242] and a mouse equipped with projector for spatial annotations with a pen on, e.g., blueprints [243]. With the TouchJet Pond [251] a mobile version of a smartboard has become commercially available.

FLUID, SHAPE CHANGING, OR FLYING Not handheld but still mobile are a number of related works which use projection on non-rigid surfaces. The AquaTop display [146] projects on a white salted bath and tracks interactions above the bath as well as poking out fingers from a submerged hand. Given more sophisticated tracking techniques as the used Kinect depth sensor, completely new types of touch and gesture interaction could possibly be realized. Similarly, acoustic levitation of tiny white particles as presented by Ochiai et al. [188] allows the projection to be penetrated by the user's hands as well. InFORM [89], a successor of [154] projects on a surface consisting of 900 mechanically actuated pins, making the surface a shape changing interface. These pins allow to be touched, pushed, and pulled to provide three dimensional touch or haptic interaction, respectively. The Flying Display [187] allows interaction on a 2D projected flying display in mid-air, provided by a pair of drones one of which features a projector, the other one a projection screen.

Touch interaction, including direct touch interaction (chapters 7 and 8), using a pen (Chapter 6) or using tangible (Chapter 9) interaction, denotes the main interaction theme used throughout this thesis.

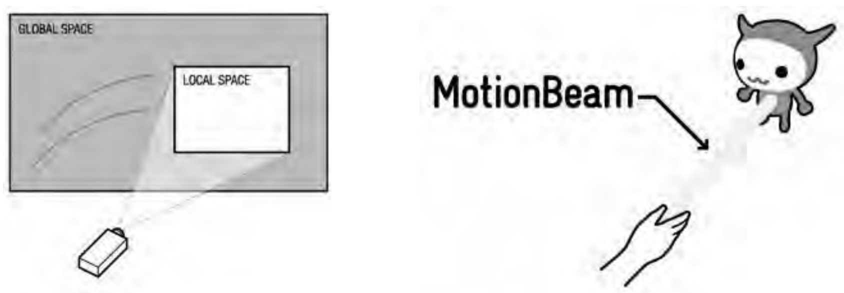
2.5.4 Interaction Metaphors for Implicit Interaction

Two common concepts for interaction with projectors have emerged, both of which rely on device motion, but both of which allow for an rather implicit than explicit usage. Schmidt defined implicit interaction as

> "an action performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as input." [227]

SPOTLIGHT METAPHOR The spotlight metaphor had already been used by Cao et al. [63] but was introduced as a concept by Rapp [207] and is a mere extension of the peephole metaphor [283] to mobile projection. The idea of the peephole metaphor is to treat the actual display as a window (or peephole) to a much larger virtual background behind it. For mobile projection, it is well conceived by treating the projector as a spotlight whose beam unveils the content that otherwise remained hidden in darkness (see Figure 2.7a). Different to peepholes, the spotlight metaphor does not only relate to movement in two dimensions, but as the size of the projection shrinks or grows depending on the distance between projector and surface, it must be implemented as three-dimensional window to achieve an immersive effect even on 2D surfaces—let alone 3D surfaces. When adding a cursor to the center of the projection, the concept is also well suited for pointing interaction and selection within the environment as we have done in WallPlay [W10].

Within the scope of this thesis, only the AMP-D project (Chapter 9) employed the spotlight metaphor. More precisely, by applying the spotlight metaphor to body- instead of device movements, the concept was extended to a wearable lantern instead of a handheld spotlight.



(a) Spotlight interaction **(b)** Motionbeam interaction

Figure 2.7: Spotlight and Motionbeam interaction metaphors as illustrated and defined by Willis et al. [267].

MOTIONBEAM METAPHOR

> “[Whereas the] *spotlight metaphor* is primarily concerned with navigating a virtual *background* space [in contrast, the *motionbeam metaphor*] is focused on interaction with characters in the foreground.” [267]

It was defined by Willis et al. [267] and introduces 8 principles—one of which is the spotlight metaphor—for interacting with animated characters that stick to the center of the projection as if they were affixed to the end of a virtual beam (Figure 2.7b).

Both interaction metaphors support implicit interaction with the projection as the user is freed from thinking explicitly about how to achieve navigation through the system. This is enabled by exploiting users' familiarity with existing interaction metaphors (spotlights and rod puppets). The Motionbeam metaphor provided some inspiration for the design of the tangible interaction used in the AMP-D prototype (Chapter 9).

2.6 TRACKING TECHNOLOGIES RELATED TO MOBILE PROJECTION

While the previous sections presented interaction concepts with mobile projectors, this section focuses on the technology required to support previous interactions. Further on, important related technologies for tracking humans, objects, and the environment shall be presented which go beyond previous use cases or may be used to improve such in the future. Again, the scope must be limited to technologies applicable to *mobile* projection, i.e. not relying on instrumentation of the infrastructure, but may well include related technologies that have not been applied to mobile projection, yet.

2.6.1 Projection Device

Allowing the projector to track its own rotation has formerly been realized by built-in or attached inertial sensors. When adding a standard 2D camera, relative positioning to a pre-calibrated camera frame [49] can be established to extend to 6 DoF tracking. Other works equipped the projector with an additional range finder instead of a camera, e.g. ultrasonic, to receive the distance to the projection wall [267]. In WallPlay [W10] we used a manual calibration procedure where the user manually aligns two images to the wall's top and bottom. With mobile depth cameras becoming available on the market (e.g. [203]), these can provide rotation and distance to the surface directly without the need for additional sensors (although inertial sensors may offer more precision and less latency to acquire some of the rotational axes). Using mobile SLAM approaches on the basis of surface reconstruction like KinectFusion [125], absolute mobile self-positioning (6 DoF) within a room is becoming feasible [176].

2.6.2 Gestures & Touch

For mobile projection, the user's limbs (head, arms, hands, fingertips, feet) and location are the most important data sources, providing performed actions and the position of the user. Computer Vision and electrical sensing have been used to track the user and also combined approaches exist [133].

COMPUTER VISION Earlier works used 2D cameras combined with infrared light for tracking [54, 213, 225, 270] or identifying shadows in the projected image [60, 78] which do not require extra hardware, but are susceptible to improper lighting environments. Hu et al. rely on the deformation of the projected image to detect touch [119].

Lately, depth cameras have been integrated into mobile ProCamSs [176] or attached to the user's body [104]. Body-worn cameras are limited in their FOV (e.g. shoulder-worn to seeing arms and hands only), which Chan et al. [69] address by presenting a pendant-worn camera with ultra fisheye lens recognizing several gestures and postures through a machine learning approach. On the other hand, Fanello et al. allow existing 2D cameras equipped with infrared (IR) LED and IR bandpass filter to perceive short-range depth similar to existing depth cameras [85]. Sharp et al. present very robust, fast, and accurate tracking of the complete hand posture [234] using modern depth cameras (Kinect V2) which future generations of mobile devices may be able to run in real-time.

Sahami et al. exploit thermal reflection for interaction that would allow body-worn ProCamS systems to grab a larger area of user interactions as a direct line of sight is not required [217]. Flexpad allows interaction with the handheld projection surface by deforming it [245]. Kohler et al. present a ProCamS for self-localization within a room [145].

ELECTRICAL SENSING If equipping the user with a projector and camera was acceptable, replacing the camera by a different device for electrical sensing might be, too. Besides worn EMG [180], acoustic [106] or inertial sensors, the capacitive sensing of Touché [224] can be combined with projections on a large range of grounded objects to facilitate touch interaction.

2.6.3 Objects, the Environment and the Projection itself

Tracking objects for 2D augmentation has been achieved by recognizing their shape [133] (requires learning process) or by attaching simple LEDs to them [228, 267]. Three-dimensional augmentation requires estimating pose and orientation of the object through a series of invisible LEDs (like used by the Wii remote), structured light [280], or invisible printed IR patterns [268]. This is not very different from surface reconstruction by means of depth cameras [125]. An approach that learns object features over time has been presented by Molyneaux et al. [175].

Objects can well be other projection devices for multi-user collaboration. In SideBySide, the red channel of the projector is replaced by an IR projector and this channel is used to project Quick Response (QR) codes for invisible optical transportation of user actions and relative projector position alongside the visual content [269]. Cotting et al. use imperceptible patterns on top of the visible projection to achieve a similar relative positioning [76]. Twinkle recognizes thick drawings as physical borders for projected characters in a mixed reality [284].

Nomadic devices cannot rely on external tracking infrastructure. From a commercial perspective, based on contemporary mobile devices, they are also more likely to provide cameras and inertial sensors than muscle sensing facilities. Projects in this thesis thus mainly employed computer vision and inertial measurements for interaction tracking. Nevertheless, biometric sensors, especially as part of smart watches and wristbands, are on the rise and will provide further alternatives in the future.

2.7 PROVIDING TACTILE FEEDBACK

One of the biggest challenges in ongoing research on projected interfaces is how to achieve adequate tactile feedback. With handheld projections, device vibration can be used similar to how tactile feedback is provided for touchscreens on smartphones and tablets. But many projection scenarios are not suitable for handheld projection, thus feedback on the projection device is not available. Touch-operated projections provide a natural haptic feedback on the surface. However, as previously presented touch recognition technologies for projected interfaces are not as accurate and robust as contemporary capacitive touch technology, the system might miss touch events despite the natural haptic feedback.

Recent works have investigated generating tactile sensations remotely through the air. AIREAL [241] and AirWave [99] use directed air vortex rings, compressed actuated air, to create remote sensations. The systems are currently too large for mobile application, though. Remote sensations can also be generated through ultrasonic pulses, allowing for smaller systems [64] with a commercialization on the way [253].

Another way of providing tactile sensations is to put a low current anywhere on the user's body through a body-worn device [37]. When touching conductive objects the user perceives a controllable texture at their fingertip. However, to many everyday objects the conductive layer must be added to work with REVEL. The technology has also been applied to back-projected tabletops [38].

Finally, finger augmentations and implants such as simple magnets have been proposed to simplify the creation of force feedback by changing magnetic fields [264, 274].

As much as these systems have potential to solve the lack of tactile feedback for projected interfaces in the future, so far they are too bulky to be applicable to nomadic scenarios. The approaches presented in this thesis have thus to rely on natural haptic and visual system feedback only.

The next chapter will present closely related works on (nomadic) projected interfaces.

3

RELATED WORKS ON MOBILE PROJECTED INTERFACES & THEIR CLASSIFICATION

Mobile projections are very popular across a variety of application domains, at least in research. A hand full of related overview articles and theses [61, 216, 265] have already classified certain aspects of *personal* and *mobile* projection, including implications of the form factor and in- and output techniques. However, they did not yet consider the requirements and implications of *nomadic* projection and nomadic information management. The foremost of these is that the device is required to work not only in a purely mobile sense (allowing the user to move with the device) but completely independent from the immediate infrastructure around, i.e. on the go and in unfamiliar, even inappropriate locations. Recalling the introduction on nomadic computing, the experience in these locations should still be "transparent, integrated, convenient and adaptive" [143].

As a consequence, the next section (3.1) will distinguish related works on projected interfaces regarding their goal or suitability to support nomadic interaction. In a next step, Section 3.2 will classify related works regarding their support of the previously identified deficiencies, Personal Information Management (PIM) in general, and further intentions why projection has been applied. As the previous chapter (Subsection 2.4.1) has already explained, for now and the foreseeable future nomadic projection is only feasible at short projection distances. Hence, in a final step, Section 3.3 is going to look at the interaction distances the respective works support.

The taxonomy on the next two pages gives an overview of the classification of related works according to these three aspects, uncovering a largely ignored research area that is the focus of this thesis and distinguishes it from previous research on projected interfaces. The following sections will then discuss each of these aspects in detail.

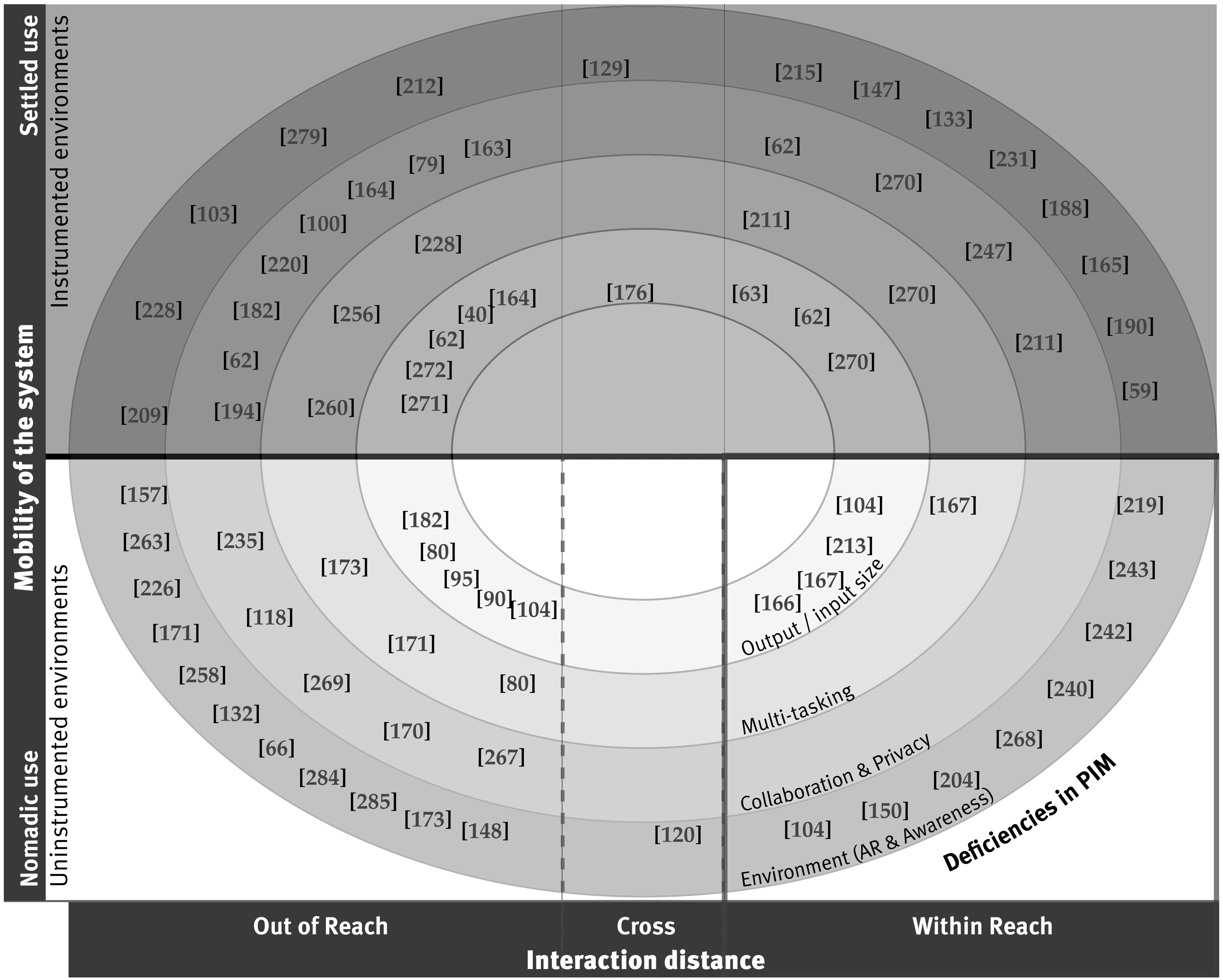


Figure 3.1
Taxonomy of related works, depicting *interaction distance* on the x-axis (three areas), *mobility* on the y-axis (two areas) and the *deficiencies* addressed by projection on the radial axis (four areas).
Clearly, the area of *realistic*[a] nomadic support by using *cross*- and *within reach*-interaction (red borders) is considerably underrepresented. When considering the fact that none of these aim at general support of Personal Information Management (PIM) and the almost non-existent support of multi-tasking and collaboration, it is further obvious that this thesis explores a novel area in the larger field of mobile projection. At the same time, this area seems very rewarding to investigate because of the various advantages of *within-reach* projection and interaction[a].

a as hinted by Subsection 2.4.1 and concluded in Chapter 5

TAXONOMY-RELATED REFERENCE LIST (FOR CONVENIENCE OF THE READER)

To maintain a single page, the list only shows citation keys which consist of first author, the first three words of the title (minus conjunctions), and the publication year (except for online references). The full list of references is to be found on page 270.

[40] beardsley_interactionusinghandheld_2005
[59] _burtoninc_
[62] cao_multiuserinteractionusing_2007
[63] cao_interactingdynamicallydefined_2006
[66] cassinelli_skingames_2012
[79] cowan_projectorphoneuse_2012
[80] dancu_smartflashlightmap_2014
[90] _fujitsustartselling_
[95] greaves_evaluationpicturebrowsing_2008
[100] gurevich_teleadvisorversatileaugmented_2012
[103] hardy_toolkitsupportinteractive_2012
[104] harrison_omnitouchwearablemultitouch_2011
[118] hosoi_visiconrobotcontrol_2007
[120] huber_lightbeamnomadicpico_2012
[129] jones_roomalivemagicalexperiences_2014
[132] kajiwara_clippinglightmethodeasy_2011
[133] kane_bonfirenomadicsystem_2009
[147] korn_potentialsinsituprojectionaugmented_2013
[148] krum_augmentedrealityusing_2012
[150] laput_skinbuttonscheap_2014
[157] leung_designingpersonalvisualization_2011
[163] martinezplasencia_reflectoslatespersonaloverlays_2014
[164] mathur_exploratorystudyuse_2011
[165] matoba_splashdisplayvolumetricprojection_2012
[166] matsuda_wearableinputoutputinterface_2013
[167] matsumoto_embodiedwebembodied_2008
[170] mcfarlane_interactivedirtincreasing_2009
[171] mcgookin_studyingdigitalgraffiti_2014
[173] mistry_wuwwearurworld_2009
[176] molyneaux_interactiveenvironmentaware_2012
[182] ni_anatonmefacilitatingdoctorpatient_2011
[188] ochiai_pixiedustgraphics_2014
[190] okude_rainteriorinteractivewater_2011
[194] oswald_igeexploringnew_2015
[204] qin_dynamicambientlighting_2011
[209] raskar_rfiglampsinteracting_2004
[211] rekimoto_augmentedsurfacesspatially_1999
[212] robinson_haptiprojectionmultimodalmobile_2010
[213] roeber_typingthinair_2003
[215] rosenthal_augmentingonscreeninstructions_2010
[219] sakata_mobileinterfacesusing_2009
[220] samosky_bodyexplorerarenhancingmannequin_2012
[226] scheible_interactivesnowsculpture_2011
[228] schmidt_picontrolusinghandheld_2012
[231] seah_sensabubblechronosensorymidair_2014
[235] shilkrot_pocomoprojectedcollaboration_2011
[240] sodhi_lightguideprojectedvisualizations_2012
[242] song_penlightcombiningmobile_2009
[243] song_mouselightbimanualinteractions_2010
[247] sugimoto_hotaruintuitivemanipulation_2005
[256] vatavu_theresworldoutside_2013
[258] virolainen_burntosharecontentsharing_2010
[260] waldner_displayadaptivewindowmanagement_2011
[263] weigel_projectorkiteasingrapid_2013
[267] willis_motionbeammetaphorcharacter_2011
[268] willis_hideoutmobileprojector_2013
[269] willis_sidebysideadhocmultiuser_2011
[270] wilson_playanywherecompactinteractive_2005
[271] wilson_combiningmultipledepth_2010
[272] wilson_steerableaugmentedreality_2012
[279] xiao_worldkitrapideasy_2013
[284] yoshida_twinkleinteractingphysical_2010
[285] zhao_picopetrealworld_2011

(a) Lightspace [271] **(b)** RoomAlive [129] **(c)** Aware handheld projectors [176]

Figure 3.2: Settled projection examples facilitating collaboration (a) and entertainment using whole-room (b) and handheld-projection (c).

3.1 MOBILITY OF THE PROJECTION

While the efforts in nomadic computing started by focusing on the constant availability of network connectivity, Vartanpiroumian was right in arguing that today supporting nomadicity only at the network level is too short-sighted and must be considered at the application level as well [255]. However, this still does not go far enough. The limitations of mobile devices that have been outlined before (small output/input size, carried in pockets/bags) cannot be adapted by software alone but require new device concepts and hardware configurations to adequately support nomadicity.

Following on that idea, we should first differentiate projection systems in those which do *not* support nomadicity (which we will call "settled projection") and those which do:

SETTLED PROJECTION This category comprises two scenarios: (1) The traditional application of projection to business presentations or movie screening. As these are not interactive, they are not closely related to HCI and therefore ignored in this classification. (2) Interactive systems that allow one or several projection(s) and one or several user(s) to move within the confined space of an instrumented environment, e.g. a smart room or a smart building. Typically, projectors are affixed to ceilings and walls of the interactive rooms to support this type of interaction. However, even handheld projection devices are sometimes included, but which then rely on the environment to achieve their functionality (e.g. [176], see Figure 3.2).

NOMADIC PROJECTION In contrast, in nomadic projection the whole system is standalone and mobile in such a way that it can be carried by the user as a general companion, placed or held for interaction at various, uninstrumented places. Recalling Kleinrock's definition of nomadic computing [143], we can argue that mobile projectors fit the requirements particularly well because:

1. if they are pico-projectors, they can be easily *integrated* to the device;
2. they allow, as any projector, for easy *adaptation* to the current context regarding position and size of the display;
3. if they recognize their context of use, they can further *transparently* adapt to it (e.g. automatically enable/disable, automatically adapt to geometric or color distortion) which adds to the *convenience* in nomadic device usage.

The taxonomy on the previous pages shows how related works have been classified in either one or the other category. As only works on nomadic projection are closely related to this thesis, in the following, only these will be discussed in detail.

3.1.1 Related Works on Nomadic Projection

Works on nomadic projection further branch out in those where the projector has been integrated into another device to enhance its capabilities and those where the projector—or a new type of projection device—enabled a new functionality on its own.

3.1.1.1 *Device-integrated Nomadic Projection*

For instance, Song et al.'s PenLight [242] and MouseLight (Figure 3.3a) [243] focus on augmenting existing information layers, such as paper sheets like blue prints. Positioning the device on the paper gives access to otherwise hidden digital information like heating, ventilation, and pipes. In comparison a smartphone would only allow for a smaller AR view and more importantly, would not allow for bi-manual interaction. Kajiwara et al. [132] integrated a projector to a digital camera to visualize the camera's viewfinder for easier free-hand shooting or taking pictures with the self-timer. Roeber et al. [213] presented a projected hardware keyboard and this idea has lately been integrated to the Lenovo Smart Cast phone [156]. Projector phones have further been used to share pictures solely through an optical channel to (supported) public displays [258] or for health education in rural india [164].

3.1.1.2 *Standalone Nomadic Projection Devices*

Standalone devices often support a more specific use case. McFarlane et al. [170] (Figure 3.4b), for instance, support military mission planning in the field. Map navigation on the street while riding a bike was presented by Dancu et al. [80]. Furthermore, games have been a well researched topic with mobile projection devices. Willis et al. support single-user gaming using the motionbeam metaphor [267] (Figure 3.4a).

(a) MouseLight [243] (b) Browse pictures [95] (c) Bonfire [133]

Figure 3.3: Device-integrated (nomadic) projection examples enabling (a) bi-manual interaction with digital information overlays, (b) picture browsing using a projector phone setup and (less nomadic in comparison) peripheral projections to increase display space and awareness (c).

This is extended to multi-user gaming, which does not require any network infrastructure, once using IR-projected invisible markers for optical communication [269], once using small visible markers in the corners [235]. Further games include a personal projected pet [285] and jump'n run interaction using AR and physical boundaries [284]. Leung et al. [157] use a mobile projector to constantly show one's own online-social identity on the ceiling above to spur conversations with people nearby.

A broad vision of nomadic projection is painted by Mistry et al. [173] and the Sixth-Sense concept and prototype (see Figure 3.4c). Some of the included ideas, like everywhere interaction with a projected display, may be feasible without wearing colored fingertips in the near future, when advanced gesture tracking systems such as [104, 134, 166] become small and mobile enough to wear them as general companion. Other ideas of Sixth-Sense, like augmentation of flight tickets, require a degree of world knowledge about arbitrary nomadic situations which seems further off in the future—although, for instance, Google's "Now on tap" and Apple's "Proactive"-technologies are making constant progress towards this goal [181].

3.2 INTENTIONS FOR USING PROJECTION AND APPLICATION DOMAINS

In their survey articles, Cao [61], Rukzio et al. [216], and Schöning et al. [230] consider personal and collaborative information management as one of the three main intentions of nomadic projection. This opinion is supported by a social study by Cowan et al. [79] on the possible usage of mobile handheld projectors and projector phones that revealed PIM as frequently mentioned application domain for nomadic usage sce-

(a) MotionBeam [267] (b) Interactive Dirt [170] (c) OmniTouch [173]

Figure 3.4: Standalone nomadic projection devices designed for gaming (a), military mission planning (b), and leveraging the environment to increase awareness (c).

narios (also much in common with the general smartphone usage as mentioned in the introduction). Furthermore, mobile (handheld) projections have been shown to provide an information throughput comparable to mobile screens [201].

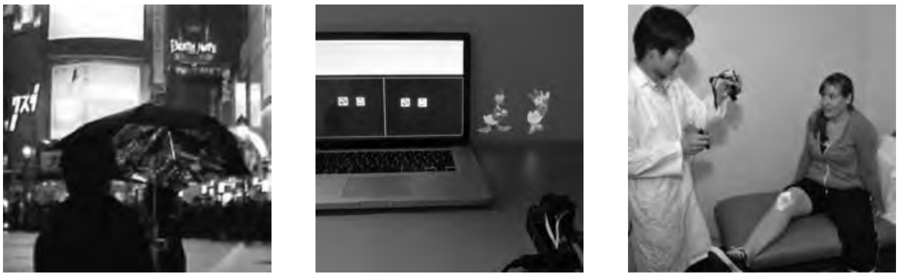
Important aspects pertaining to nomadic PIM have already been identified in form of the four deficiencies (Subsection 1.1.1) in the introduction. These are closely related to the topic of this thesis and related works having created support in one of these domains will be the focus of the next four subsections and are also classified on the radial axis of the taxonomy. Further on, related works, of course, pursued intentions apart from information management and subsequent subsections summarize the most prominent remaining ones.

3.2.1 Related to Deficiencies in Nomadic Information Management

3.2.1.1 *Increasing Display Output/Input Size*

Such as it was the basic use case for static projection, mobile projection is oftentimes used to increase the display real estate. As the ultimate vision of pervasive computing offering interactive displays anywhere anytime has not come true, yet, display real estate is still sparse in most mobile environments. Larger displays by means of mobile projection have been used to make personal data exploration—such as browsing the picture gallery of the own smartphone—more convenient [94, 95] (see Figure 3.3b) and add a large display to the street in front of bikes [80].

In settled scenarios, being able to use all walls and the floor of a room as display either through steering a self-contained display across the room [197, 272] or using the spotlight metaphor (Subsection 2.5.4) [62, 63, 271] can be regarded as significant increase in output/input space (cf. Figure 3.2).



(a) Pileus umbrella [167] (b) SideBySide [269] (c) AnatOnMe [182]

Figure 3.5: Projection addressing intentions related to information management (in spite of being partly applied in other domains). (a) A projected umbrella display for nomadic information access such as wayfinding on a map, addressing display size and multi-tasking. (b) A device concept for gaming allowing ad-hoc collaboration without any reliance on infrastructure. (c) Using AR to increase the patient's awareness in the communication.

3.2.1.2 *Multi-tasking*

Projection has been used for increasing multi-tasking capabilities in settled environments to place documents and windows across the room [63, 211, 256, 260] and interactive tables [270]. Regarding nomadic scenarios, some allow to pursue real-world tasks while using the device to follow information in the umbrella interface [167] (Figure 3.5a), looking at digital street graffiti [171] or riding the bike [80]. SixthSense [173] supports micro-interactions using quick gestures for taking a picture. However, digital multi-tasking between several applications on one device has not been researched so far.

3.2.1.3 *Collaboration & Privacy*

An innate purpose of projection has always been to simplify collaboration. Especially smart spaces (Section 3.1) today are able to provide an unprecedented level of collaboration. But not much inferior, mobile projections have been used for multi-user games [118, 194, 235, 269] (see Figure 3.5b), information exchange [62, 182] (Figure 3.5c), augmented learning [220], remote assistance [100] and of course for collaborative media presentation and management [79, 164, 247, 262]

Another picture is painted when looking into the privacy support of mobile collaborative projection. Here, only Cao et al. [62] provide rudimentary privacy controls, most other works do not consider privacy. Privacy has only been considered much for static projection scenarios, here especially for projected tabletop interaction [163, 232] and it is to question how mobile and even nomadic systems can support privacy.

3.2.1.4 *Environment: Augmented Reality & Awareness*

AUGMENTED REALITY Projection, in particular mobile projection, is especially well suited for creating AR due to its physical combination with the augmented artefact, which when correctly applied (i.e. using compensation techniques as in Section 2.4) leads to a very believable mixed reality.

Raskar et al. [209] envision a tagged world where additional information is revealed by handheld projectors directed towards them. Schmidt et al. [228] and Huber et al. [120] extend this to augmented controls for interaction with objects and appliances in the environment. Wear Ur World [173] provides many examples of daily life where such dynamic information would be valuable, ranging from live delay information on flight tickets, videos and live weather data in the newspaper, to in-situ product and book ratings. Sakata et al. [219] and Sodhi et al. [240] propose hand augmentation for pedestrian navigation. Song et al. [242, 243] present registered paper annotations, for instance on top of blueprints. Molyneaux et al. [176] present an exhaustive range of examples of AR using projection within a personal room such as a living room. Ni et al. [182] allow doctors to augment patients with annotated x-rays simplifying communication and understanding between doctors and patients. Roeber et al. [213] use projection to augment a keyboard in front of the device. Apart from augmenting objects, augmenting the own world, for instance through a projected pet [285], is another form of AR. The twinkle game is affected by edges in the real environment [284] and Cassinelli et al. [66] analogously use projection for "playable clothing".

A special type of AR may be considered to use the emitted light of projectors directly for input to another system, like an optical and thereby analog transport of information. Burn To Share [258] allows to point the projected image to a back-projected digital bulletin board. A camera watching the bulletin board from the rear captures difference images between projected and observed image and is thus able to capture the front-projected image of the user to store it permanently on the board. The optical transport is further used for invisible control in the already mentioned PICOntrol [228], Lumitrack [280] and SideBySide [269] systems.

Another vision of AR enabled by projection are displays seemingly floating in mid-air. Usually, a display medium is required to reflect the photons and make the projection visible. Nevertheless, also soap bubbles [231] and cooperative drones [187] have been proposed as well as particles levitated by ultrasonic waves [165, 188]. Recently, the controlled creation of plasma (tiny lightning bolts) in mid-air has been presented. A very bright variant exists for emergency cases [59] but is too dangerous for interaction. Conversely, Fairy Lights [189], creates only very small displays (size of a coin) but is less dangerous and allows for touch interaction with the projection (albeit still using class 4 lasers whose direct exposure leads to eye or skin damage). Nevertheless, this vision can be considered the ultimate goal of nomadic projection as it would break its current dependence on suitable surfaces in the environment while retaining its advantages compared to screen-based displays.

SUPPORTING AWARENESS & PERIPHERAL INTERACTION Static projection has a long history in providing peripheral display and interaction [e.g. 52, 172, 192, 198]. Comparably, there is very few related work on awareness and peripheral interaction with mobile projectors. Qin et al. [204] do not exactly use a projector, but the projected aura around a mobile phone to increase awareness of notifications is not all that unlike. HaptiProjection [212] allows serendipitous information encounter but requires more active interaction than typical peripheral systems. The Pileus Internet Umbrella attached a projector to the handle of an umbrella with white interior that is used as projection surface to provide information access while on-the-go (cf. Figure 3.5a). Because of the umbrella that shields ambient light, it can be used outside and therefore adapts to different lighting environments, granted that the user considers wearing a sun-umbrella socially acceptable.

3.2.2 Further Intentions and Application Domains

For obvious reasons, application domains apart from information management are not as closely related as previous categories. Nonetheless, they are worth mentioning to provide a thorough picture of the state of the art in mobile projection.

Previously mentioned application domains for applying projection already encompassed domains like gaming [57, 66, 74, 118, 122, 128, 129, 131, 160, 174, 235, 251, 267–269, W10, 282, 284, 285], learning [21, 139, 147, 220], military support [170], smart homes [223, 256], smart domes [44][43], and new approaches to industry workflows [147, 215]. Three intentions/domains that remained unmentioned so far are projected interfaces in the arts, toolkits, and where projection was used as expedient instrument.

3.2.2.1 *Art*

Besides the early works of augmenting pictorial artwork [51], "projection mapping" has become an indispensable instrument in the toolbox of artists. Early works of Scheible [226] projected on snow and facades. The website `http://projection-mapping.org` presents art projects like floor-projected piano playing [24], interactive restaurant experiences by means of storytelling directly on the table [153], or ad-hoc visual artistry with a body-worn interactive projector system [261]. Projection further allows to augment places which are otherwise out of reach, sometimes not only physically but also legally. When the Spanish government prohibited demonstrations in front of the Spanish parliament in 2015, a virtual demonstration party was projected in front of the parliament that provided a correct 3D perspective for a filming camera nearby [116].

3.2.2.2 *Toolkits*

Two sorts of toolkits for projected interfaces have emerged. Commercial systems usually provide very good support for perspective projection mapping using videos or bundled effects even across separated surfaces (see [46] for a good overview). In contrast, research projects offer less maturity but bundle the projection with interaction facilities, for instance, by means of depth cameras. Hardy et al. [103] provide a toolkit that allows multiple planar surfaces in the FOV of the ProCamS to be defined as touch-enabled display and with geometric compensation. The RoomAlive-Toolkit by Microsoft can be regarded as an extension of the former [214]. WorldKit [279] adds support for several types of interaction widgets and allows to define these interactive surfaces interactively. While these toolkits support static smart spaces, Weigel et al. [263] target mobile projection including multiple projectors.

3.2.2.3 *Expedient Instrument*

Finally, projection technology has oftentimes been applied as an expedient instrument: in any case as a substitute for very large physical screens or floor displays [54], but also to prototype and evaluate new device concepts for which the required display technology did or does not exist, yet. Prominent examples of such are shape-changing interfaces such as bendable [206, 282] and rollable interfaces [137].

3.3 INTERACTION DISTANCE

The low brightness of mobile projectors particularly poses an issue to nomadic projection which must work under arbitrary lighting conditions—direct sunlight let aside—as they are typically out of the user's control. Huber et al. [120] use a ProCamS consisting of laser pico-projector and Kinect depth camera for nomadic interaction with objects at arbitrary distances. Such focus-free LBS projectors cannot provide enough brightness for the envisioned remote projection, though. Similarly, the pervasive graffiti by McGookin et al. [171] may not be visible in most

(a) PlayAnywhere [270] **(b)** Dyn. information spaces [63] **(c)** Omnitouch [104]

Figure 3.6: Examples of within-reach projections: (a) Leveraging the environment like the palm and walls for interaction; (b) Not nomadic, but mobile everywhere table-top interaction; (c) defining information spaces within a room using a pen (within reach) and interacting with them from an out-of-reach distance.

outdoor environments, which is why they also provide an alternative AR view on the mobile device. In this regard, the usage of special material like retro-reflective material [148] has been investigated and makes even low-brightness projectors usable outdoors. At the same time this precludes nomadic applicability as these materials do not occur naturally in the environment. This leads us to the next differentiator among related works.

3.3.1 Out-of-Reach Interaction

Unfortunately, the majority of existing and previously mentioned works on nomadic projection employed out-of-reach projection for their interaction design (cf. the taxonomy on page 44 again). As we have learned earlier, in history projections used to be far away from the viewer and have been interacted with—if at all—only through out-of-reach interaction (e.g. by pointing with laser or mouse pointers). Hence, it does not surprise to find many more works using *out-of-reach* than *within-reach* projection and interaction (Figure 3.1 shows almost twice the amount of references on the left compared to the right side). Many of these use the distant projection for gaming on nearby walls and the floor [66, 118, 122, 128, 129, 160, 174, 235, 251, 267, 269, W10, 284, 285], learning [21, 131, 147, 220], military support [170], smart homes [223, 256] or smart domes [43, 44]. As projection used to be mostly applied in spaces that offer control over the ambient light, out-of-reach projection offered large images at very low cost, which made it so appealing.

3.3.2 Within-Reach Interaction

In nomadic scenarios, the opposite is true: The user typically has no control over ambient light and thus the natural quadratic light attenuation (following the inverse square law) leaves only a faint image of the remote projection in bright environments. Of course, within-reach interaction has been used, but mostly in instrumented environments (settled use) in previously mentioned domains [59, 62, 63, 133, 147, 165, 188, 190, 215, 231, 270] (see Figure 3.3c and Figure 3.6a-b) as well as in further domains, e.g., new industry workflows [147, 215]. Only very few works applied a *within-reach* distance to nomadic interaction, for instance, to make the user more aware of incoming calls [204], to project on the palm for micro-interactions [104, 166, 219] (cf. Figure 3.6c) or pedestrian navigation [240], nomadic reading of interactive textbooks [268] and the already mentioned PenLight [242] and MouseLight [243] systems (Figure 3.3a).

3.3.3 Cross-distance Interaction

The cross-distance space is not as much lying somewhere *between* out-of-reach and within-reach interaction distances—although the taxonomy may give this impression—but is meant to encompass both distances, preferably *adapting* the style of interaction to the user's preferred distance, instead of dictating a certain position to take to interact with the device. Especially in nomadic scenarios, users might not always have the room or freedom to take arbitrary positions to their devices, which is why they should consider adapting to different interaction distances. As depicted by the taxonomy (Figure 3.1), current support for adapting to different interaction distances is sparse. Huber et al. [120] use the projection distance as input gesture, e.g., to flip through a stack of documents, but do not yet change the input method based on distance. The handheld projector system by Molyneaux et al. [176] automatically switches between shadow-based interaction for out-of-reach distances, and touch interaction when coming into within-reach distances. RoomAlive [129] supports out-of-reach (like shooting with a gun) and within-reach interaction (punching a bug) simultaneously within an instrumented environment.

Until now, we have learned about technical arguments for within-reach interaction such as increased brightness of the projection, which Chapter 5 will later use as argument for the Nomadic Projection Within Reach framework. Nonetheless, human factors regarding nomadic information management with projected interfaces are equally important to look at. This is the purpose of the next chapter, which researches nomadic information management using *out-of-reach* projection.

4

INVESTIGATING OUT-OF-REACH INTERACTION WITH PROJECTOR PHONES

Related video

The previous chapters, so far, provided a technical overview as well as a classification of related work, both of which motivate more research to be conducted on *within-reach* interaction. However, human factors related to mobile projection, interaction distances, and nomadic information management have not been considered so far. That is the undertaking of this chapter. In particular, it compares existing and new techniques for distant pointing with projector phones—as they are the most likely future platform for nomadic projection—in various application domains. The main result of this is that mobile projection offers advantages for some application domains, including information management, but *within-reach* interaction outperforms *out-of-reach* interaction at least regarding information management tasks.

This chapter is based on the previously published refereed conference paper:

[W9] **Winkler, C.**, Pfeuffer, K., Rukzio, E., "Investigating mid-air pointing interaction for projector phones." In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*. ITS '12. New York, NY, USA: ACM, 2012, pp. 85–94

In addition, the following partially related thesis was supervised by the author:

- "Development and Evaluation of Mid-Air Interaction Techniques for Projector Phones". Ken Pfeuffer. Bachelor's thesis. 2011

4.1 INTRODUCTION

An inherent issue of mobile phones with touch screens is their small size, which is on the one hand essential for their mobility but on the other hand significantly limits the available space for input and output of information. The emergence of pico projectors and in particular projector phones, i.e. phones with built-in projectors (e.g. Figure 2.6 on page 26), provides a versatile solution for this issue: users can project and interact with a large display almost anywhere and at any time. Such projector phones also support various forms of colocated media viewing, browsing and interactions, which are not possible with conventional mobile phones.

Projector phones (e.g. Samsung Beam 1 and 2, Sharp SH-06C) or accessory projectors (e.g. the SHOWWX+ from Microvision) that can be connected via TV-out to a conventional phone only mirror what is usually shown on the touch screen [216]. Projecting the touch screen user interface while maintaining the same interaction style must lead to suboptimal interactions as it requires many context switches during operation and because those interfaces were designed for high resolution screens with small dimensions operated through direct touch input.

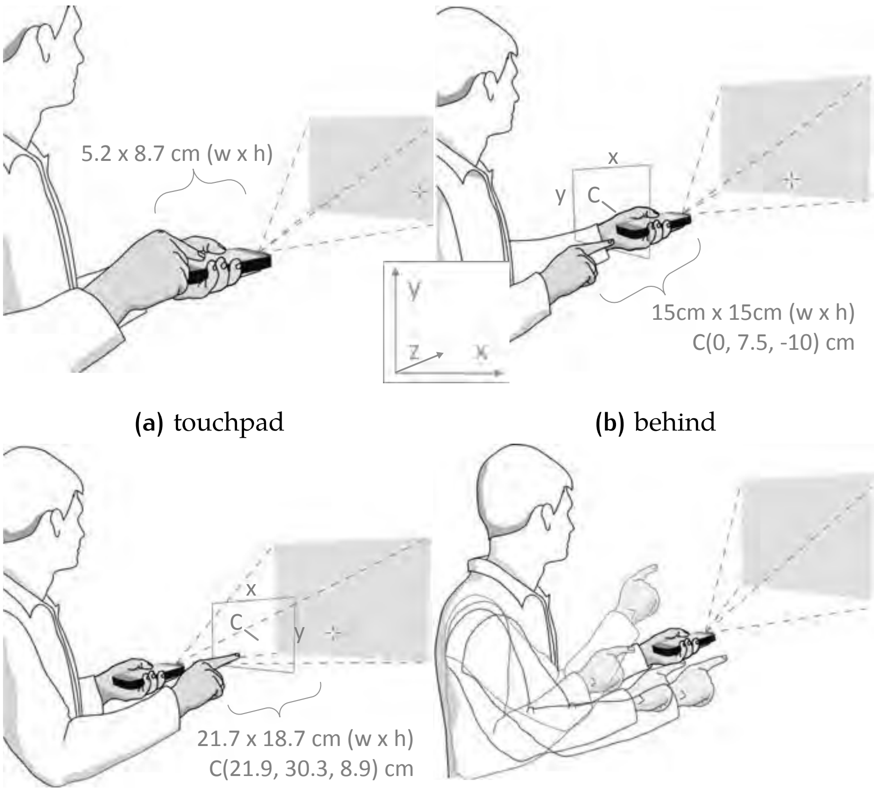
Using a pointer as an intermediate that marks the current position on the projection is a basic way to overcome some of these problems. In particular, using the touchscreen of the mobile phone for indirectly controlling a mouse pointer on the projection requires no additional hardware and has been the focus of various research projects and products, e.g. [179]. The conceptual disadvantages are the indirectness and the unavailability of the touchscreen for interaction or as information display since it is occupied as a touchpad.

Using mid-air finger-pointing techniques is an interesting alternative due to the more direct interaction and the possibility to use the mobile phone screen as secondary in/output to the projection. Further, these techniques neither require the user to carry additional hardware nor do they require movement of the phone that interferes with the projection as accelerometer based interactions[1] would do for instance. Thus they seem very suitable for typical ad-hoc mobile scenarios. However, the mid-air space around the user is quite large and unexplored considering the bi-manual and interdependent control. So far it is unclear which interaction area will be optimal and how well users will be able to manage the dual-display, bi-manual interaction.

To open this area of research we investigated the performance of three mid-air finger-pointing techniques leveraging different interaction areas (see Figure 4.1b-d) compared to the existing *touchpad* technique (Figure 4.1a). We compared the techniques through an experimental user study based on the ISO 9241-9 tapping task. Our results indicate that the interaction technique in which the user points *behind* the mobile phone to control a cursor on the projection performs significantly better than other mid-air techniques in all scenarios. Even more, it also performs ≈15% faster than the *touchpad* option despite yielding ≈2.5 times more errors. This makes it an interesting alternative interaction technique for a variety of application scenarios, even without considering its aforementioned advantage of keeping the touchscreen free.

In a follow-up experiment, we compared the superior interaction techniques of the previous study, *touchpad* and *behind*, in common nomadic usage scenarios such as browsing, map navigation, gaming, and annotating (drawing) in order to analyze their performance in realistic

1 if not using the Spotlight metaphor (Subsection 2.5.4) metaphor but which requires a calibrated environment



(a) touchpad **(b)** behind

(c) group defined **(d)** user defined

Figure 4.1: Compared pointing techniques. 'C' denotes the central point of the interaction area in relation to the projector phone position (0, 0, 0).

contexts. Also, we included a standard smartphone without a projector in our comparison in order to analyze the performance of the projector phone interaction techniques in contrast to the current usage of a smartphone. Results of the second study confirm the familiarity of users with standard *touchscreen* phones but also highlight various advantages for the projector phone interaction techniques, e.g., in terms of not occluding targets on the screen, improved visibility, the usage in collaborative settings, and joy of use. Having said that, PIM-related applications did not benefit, but instead achieved better performance on the touchscreen which will be discussed at the end of this chapter.

4.2 SPECIFIC RELATED WORK

Most available solutions used (and still use) the touchscreen of the mobile phone for input which requires no additional hardware but suffers from the separation of input (phone) and output (projection). This could lead to a large number of context switches as the user has to switch their focus constantly between the projection and the phone in case the phone screen is used for displaying information [68, 95]. The usage of the touchscreen as a *touchpad* is an effective approach for controlling a cursor on a remote screen (Figure 4.1a) [169]. The advantage of this concept is that it is already very familiar from touchpads found on laptops. Conversely, it has the disadvantage that while using the screen as a touchpad, there is no easy alternative to interact with the content on the mobile phone screen at the same time. One possible solution would be a hardware button on the side of the device to toggle between touchscreen and touchpad mode. While this seems feasible it still would not allow for interactions where both displays are simultaneously active, e.g. for seamless dragging of pictures from phone display to projection.

Unfortunately, related approaches on pointing and gesture tracking for projectors (Subsection 2.5.1) cannot easily be applied to the handheld scenario. These so far require worn cameras or shadow-based interactions that are not applicable for single-user scenarios because of too-large shadows and the lack of fine-grained control. Neither applicable are touch-based interactions for handheld projection as had been outlined in Subsection 2.5.3. However, with depth cameras' integration to handheld devices being imminent (e.g., Google Tango[2], Structure Sensor[3]), it seems interesting to explore the area around the handheld device for possible pointing interaction.

In the first study we aimed for investigating how well simple pointing tasks and target selections can be performed on the projection from a projector phone. Remote pointing has extensively been researched on large *fixed* projections. The *Pointable* facilitates remote interaction with distant targets on large tabletop displays through perspective pointing and ray-casting [36]. Pointing on vertical displays has been researched in regard to the influence of effects like parallax and control type, different ray pointing techniques [130] and different devices like laser pointers [193] or bare hands [259]. However, the findings from this research strand can only partially be applied within the context of personal projectors as the mobile scenario is substantially different: the projection is constantly moving with the device; the user has to hold the projector phone during the whole interaction, which introduces jitter to the projection and the interaction, limits the possible movement area per hand and makes the interaction bi-manual by nature. Further, mobile users usually do not want to carry or use additional hardware like a laser pointer or air mouse, which is why mobile interaction techniques have to get by with the user's bare hands. Since interaction happens in unaltered environments, the gaze of the user cannot easily be made available. Hence, image-plane ray-pointing techniques are unpractical

2 `https://www.google.com/atap/project-tango/`, visited November 16th, 2015
3 `http://structure.io`, visited September 24th, 2015

and pointing must usually be based on the relation to the projector alone.

Remote mid-air pointing nevertheless has shown good performance, which is why we decided to compare the usage of the touchscreen as a *touchpad* against mid-air pointing techniques. Both have much potential to enrich interaction in various situations and do not interfere with the projection.

4.3 INTERACTION TECHNIQUES

The aim of our research is a first exploration of mid-air pointing for projector phones. Since the area around the user is quite large, we considered different spaces around the device including behind, before, above, below, and to both sides of the device. All spaces have substantial implications on the usability (cf. [113]) and technical feasibility of the approach.

Through two preliminary user studies we discovered that interacting in front of the projector phone is not a well-suited space. While this might work for projectors worn around the neck [173], interacting with the right hand in front of the projector that is held with the left hand requires that the right hand must be held very far away from the body. Additionally, the shadow on the projection created by the finger close to the projection occludes large parts of the projection. In contrast, when pointing with the index finger *behind* the projector phone (Figure 4.1b) to control a cursor on the projection, the user does not interfere with the projection. Further, it might allow for a convenient posture as the user is able to rest the upper arm of the pointing hand on the upper part of the body. Also, this technique is more independent of the user's girth.

In contrast, interacting to the right side of the device (respectively left side for left handed users) as well as interacting above or below the device pose a more difficult challenge for a real implementation: the necessity for maintaining an input space that is planar to the projection surface ($x \times y$ in Figure 4.1) assumed, the device would require a depth camera facing to the side of the device. In theory, this could capture the finger's horizontal-movement via depth sensing. Similarly, an upward facing depth camera could provide vertical movement via depth sensing for interaction above the device. However, in practice, depth cameras have two limitations: the first is their minimal detection range, which usually lies above 10 cm. The second is their inaccuracy, typically lying around a centimeter or more. For these reasons, it seems favorable to use the depth camera only for depth segmentation of the user's hand and not for precise position tracking. Hence, to control for

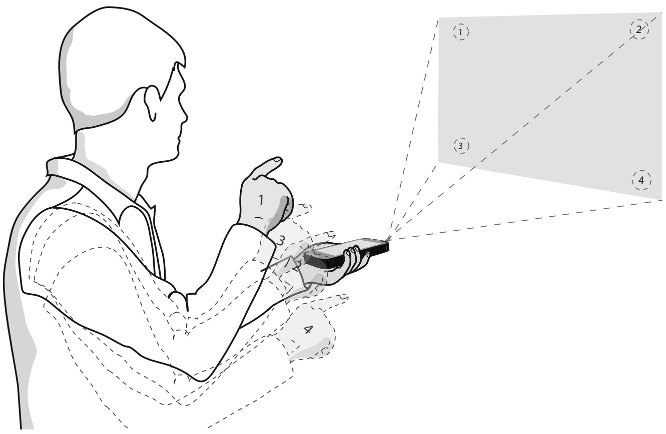


Figure 4.2: Pointing gestures performed for the definition of the input space for *group defined* and *user defined*.

a feasible amount of techniques to be tested in the user study, interaction spaces to the sides and above/below the projector have been left out.

Instead we wanted to learn and also lay more stress on participants' own likings for input (which could include any desired space around the device) following a user-elicited approach (cf. Subsection 1.3.1.3). Therefore, we derived the *group defined* technique from a separate previous assessment where we tracked pointing preferences of 27 people. Participants (7 female, 20 male) of this study were undergraduates with an average age of 23 and have not had any prior experience with our work. Each participant was asked to define their preferred input space by showing the pointing gestures they would perform when selecting the four corners of the projection by pointing at each corner three times while holding the projector phone (see Figure 4.2). Those pointing interactions were observed and measured by an optical tracking system. We calculated the average of those readings that led to an input space as specified in Figure 4.1c, which is on the top right side of the projector phone.

For the *user defined* technique (Figure 4.1d), the same procedure that has been applied for *group defined* was conducted. However, this time not in a separate study, but before the actual experiment of the main study. Thus, each user defined and used their very own input space so that no common input space can be derived for *user defined*. However, the average of users chose a 16.0 cm in width ($\sigma = 7.5$) and 14.2 cm in height ($\sigma = 7.3$) interaction space with its center lying at 6.8 cm x, 20.3 cm y, 9.7 cm z ($\sigma = 9.8$) away from the phone. In terms of size this would be similar to *behind* whereas the position would rather resemble *group defined*. Based on the different input sizes, the four techniques have slightly different control-display (C-D) gains. However, findings of Casiez et al. [65] indicate that C-D gain has a less important impact in studies modeled after Fitts' Law which our results will confirm as *behind* and *user defined* involved a very similar C-D gain but yielded significantly different results. Moreover, *touchpad*, in spite of having the highest C-D gain, was the slowest technique.

The actual selection of a target shown on the projection is in *all four* interaction techniques performed by a tap on any position of the touchscreen of the smartphone.

4.4 FIRST EXPERIMENT: TARGET SELECTION

The main goal of this experiment was to investigate whether finger pointing based techniques (controlled and user-elicited types) provide a similar performance in terms of target selection times and error rate when compared with *touchpad*. In addition, the experiment should clarify whether users perceive these techniques as beneficial. This experiment compares the previously described four techniques through a two-dimensional target selection task based on the ISO 9241-9 tapping test.

Participants

12 right-handed participants (6 female) took part in the experiment and were rewarded 10€ afterwards. Most of them were undergraduate students and aged between 15 and 27 ($\bar{x} = 23$ years). Their academic backgrounds were humanities, economics, and computer science.

Experimental Design

The experiment used a within-subjects design, i.e. all participants participated in all conditions of the experiment (in counterbalanced order). The first independent variable *technique* contained four levels: *touchpad*, *behind*, *group defined*, and *user defined*. The second independent variable *size* of targets contained three levels: small, medium, and large (Figure 4.3). The smallest target size was defined through a preliminary test where we had looked for the smallest size that could be comfortably selected with *touchpad*.

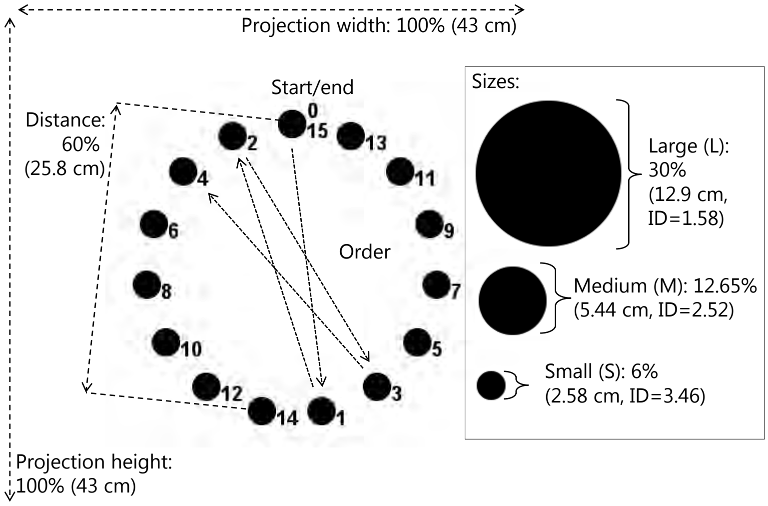


Figure 4.3: ISO 9241-9 task. Visualization of size and height of projection in relationship to the three target sizes.

4.4.1 Prototype and Setup

We assume that the three finger pointing-based interaction techniques can be realized through an additional camera on the bottom and / or side of the projector phone. Corresponding algorithms and approaches like coarse-grained depth tracking for background removal or IR-camera sensing have been reported previously, e.g. [104, 173, 270]. We used an external optical tracking system (OptiTrack V100:R2, 100Hz from NaturalPoint) and infrared markers attached to the user's finger and the projector phone in order to support accurate tracking of the index finger in relation to the phone (Figure 4.4). With this approach it is possible to compare the interaction techniques independently from a potentially inaccurate tracking solution. A SHOWWX pico laser projector from Microvision connected to a Samsung Galaxy S was used as no projector phone has been commercially available in Germany when the study was conducted.

The software used for conducting the study was written in Java and executed on the Android phone. Apart from running the study tasks and logging phone properties such as acceleration sensor values, the software also performed the pointer calculations based on the input from the tracking system in real-time. Pointer movement worked instantly without any noticeable delay. For the *touchpad* technique we implemented pointer acceleration similar to the algorithm used in Mi-

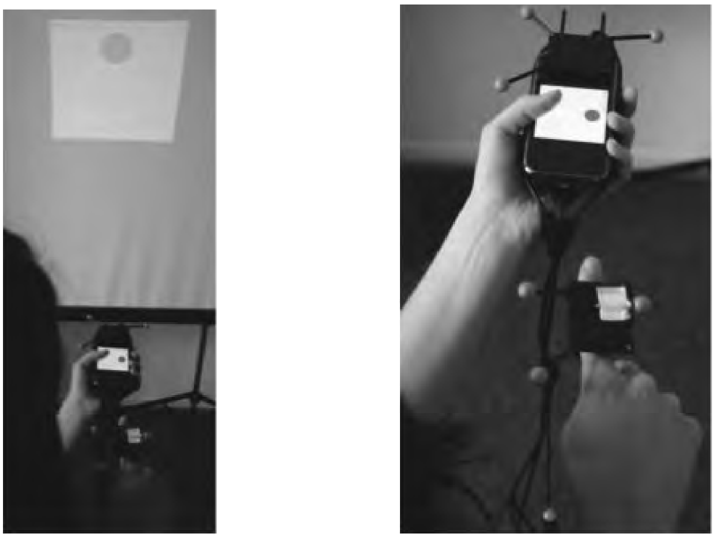


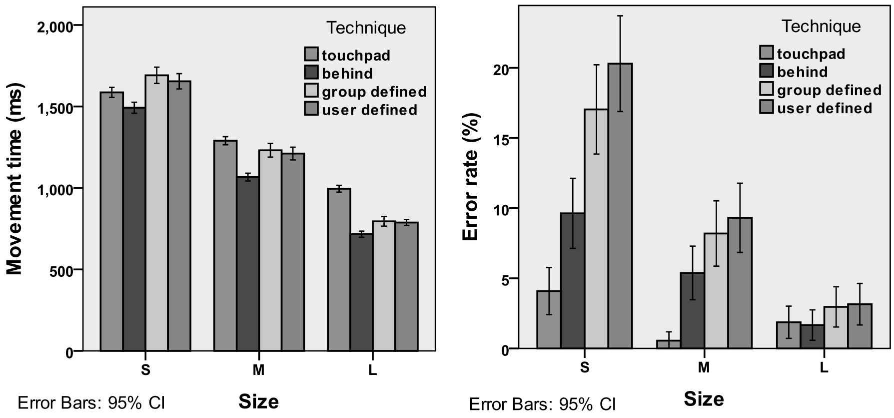
Figure 4.4: Study setting (shown for *behind* technique) and used hardware (participants did not look at phone display)

crosoft Windows [199]. Thus, and because the screen size was notably bigger than the farthest distance between targets, clutching was not required with *touchpad* in the first study.

4.4.2 Procedure

The experiment was conducted in a light dimmed laboratory room. The position of the participant was marked with an X on the floor, facing a wall 100cm away, resulting in a projection screen size of 43x43cm (see figures 4.3 and 4.4). A user standing in front of a nearby wall is considered as typical scenario for mobile usage of projector phones. Participants were asked to stay at this location throughout the study. Participants were holding the projector phone with their left (non-dominant) hand (figures 4.1 and 4.4), and pointing with the other hand. Participants were allowed to freely move the projector and their finger as only their spatial relation defined the position of the pointer.

Participants took part in the study individually. Initially, to define the *user defined* technique, the participant was instructed to point three times at each corner of the projection as they would want to point at them in the subsequent experiment. After that the experimenter explained and demonstrated the four interaction techniques and asked participants to rate each interaction technique on a 10-point Likert scale (1 – very bad to 10 – very good) based on their sole expectation. Then, each technique was tested with three different target sizes. For each of the 12 possible combinations of interaction technique x size the participant performed 1 test and 3 consecutive study rounds, each including 15 targets (see Figure 4.3). In each round, the user started with a click



(a) Movement times. (b) Error rates

Figure 4.5: Results of 1st study.

(tap any position of the phone screen) on the circle in the middle, then went to the first circle at the top from which on the time taken to every subsequent target was measured. After the user's click the target turned to green (hit) or red (miss) for 150ms and only one trial per target was allowed. After each round the user was shown their time taken and the percentage of hit and missed targets. In addition we logged hit locations and jitter of the phone using the built-in accelerometer. After each technique, participants rated the technique regarding perceived speed, precision, satisfaction, difficulty and fatigue. Finally, participants were asked to rate each interaction technique again on the 10-point Likert scale from before based on their actual experiences.

4.4.3 Results of First Experiment

Movement times and measured error rates are depicted in Figure 4.5. Movement time (*MT*) is defined by the duration between the occurrence of the target on the projection and the selection of the target. An

		Technique	Size	Tech x size
Movement time	n^2	161.850	52.744	445.843
	$\varepsilon_{G.\text{-Geisser}}$	.864	.914	.767
	df	2.591	1.828	2462.472
	df_{Error}	1386.343	978.017	2462.472
	F	73.587	2251.249	20.886
Error rate	n^2	162.925	106.886	489.437
	$\varepsilon_{G.\text{-Geisser}}$	.857	.846	.777
	df	2.570	1.693	4.663
	df_{Error}	1374.708	905.705	2494.854
	F	38.147	95.670	10.029

Table 4.1: ANOVA and post-hoc analysis of measured data.

error is defined as click outside the target area. The results reveal *behind* to require a 15.4% shorter average movement time than *touchpad* when considering all sizes. The results also reveal a 2.55 times lower error rate of *touchpad* compared to the second best error rate of *behind* that we will discuss later.

Movement times (MT) and error rates (ER) were analyzed using a factorial repeated-measures ANOVA. Since sphericity had been violated for all effects, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity (Table 4.1). According to this the main effects and the interaction effect were reported as significant ($p < .001$). The main effect *technique* and the interaction effect *technique x size* (split by *size*) were further post-hoc analyzed using pairwise comparisons of means with Bonferroni correction (for 6 and 18 comparisons respectively):

$MT \times technique$ There were significant differences in movement time between all techniques (p < .01) except for *group defined* vs. *user defined*. Hence, users performed fastest with technique behind and slowest with touchpad ($MT_{touchpad} = 1291\,ms$, $MT_{behind} = 1092\,ms$, $MT_{groupdefined} = 1239\,ms$, $MT_{userdefined} = 1217\,ms$).

$ER \times technique$ The error rate significantly differed (p < .001) between all techniques except for *group* vs. *user defined*, revealing that users made the most errors with *group* and *user defined*, less with *behind*, and least with *touchpad* ($ER_{touchpad} = 2.2\%$, $ER_{behind} = 5.6\%$, $ER_{groupdefined} = 9.4\%$, $ER_{userdefined} = 10.91\%$).

$MT \times technique \times size$ No significant differences were found between *group* vs. *user defined* (S/M/L), *touchpad* vs. *group defined* (M) and *touchpad* vs. *user defined* (M). *Touchpad* vs. *user defined* on sizes S, M were reported significantly different (p < .05). Remaining differences were reported as significant (p < .01).

$ER \times technique \times size$ On target size S, all pairs revealed significant differences (p < .01) except *group defined* vs. *user defined*. On size M, only *touchpad* vs. all other techniques showed significant differences (p < .01). Size L revealed no significant differences.

For further evaluation of the results we used the Fitts' Law model and calculated throughputs (TP) as described in [244, 277]. First all measurements of the circular tapping task were rotated to horizontal 0° and 16 of 6480 targets (0.25%) were filtered out as spatial outliers. Then we calculated the effective index of difficulty (ID_e) individually for each subject and condition (technique and target size) based on the users trials (successful or not) over all test rounds of the condition (3 rounds x 15 targets) using equation

$$ID_e = \log_2(A_e/W_e + 1) \tag{4.1}$$

where A_e is the average actual movement distance over all rounds for a particular combination [244] and W_e reflects the standard deviation of endpoints as

$$W_e = 4.133xSD_{x,y} \tag{4.2}$$

where $SD_{x,y}$ is the bivariate endpoint deviation calculated as the spread of hits $\langle x_i|y_i\rangle$ around the center of mass $\langle \bar{x}|\bar{y}\rangle$

$$SD_{x,y} = \sqrt{\frac{\sum_{i=1}^{n}(\sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2})^2}{n-1}} \tag{4.3}$$

Having ID_es for each subject, technique and target size, we calculated the individual throughput for each subject and technique using the mean-of-means approach [244], and the grand throughput by averaging individual throughputs. The grand throughputs, depicted in Figure 4.6, show a similar picture as the movement times. *Behind* outperformed other techniques, especially showing a 28.5% higher TP than *touchpad*. Our measured throughput of 1.957 bits/s for touchpad is in line with measured throughputs of traditional touchpad usage in the literature, which agrees on values between 0.99 and 2.9 bits/s [244]. As pointing on movable displays has not been studied before we can relate the throughput of *behind* to fixed pointing only. The *fixed-origin* pointing described by Jota et al. [130] shares with *behind* the similarity that the pointing ray depends on the user's finger *and* another point in space, which albeit is fixed. They measured throughput of ≈3.4b/s for *fixed-origin* pointing – for one-dimensional tasks only, though. In this light, the throughput of *behind* pointing might be slightly smaller than similar pointing on fixed projections, which can be explained by the increased complexity of the bimanual control.

A factorial repeated-measures ANOVA on throughput revealed a significant main effect of *technique* (F(3,33) = 6.219, $p < .01$, η^2 = 7.104). Post-hoc pairwise comparisons with Bonferroni corrections showed no significant differences except for *touchpad* vs. *behind* ($p < .001$). Finally, we created Fitts' Law models of the form

$$MT_{technqiue} = a + b \cdot ID_e \tag{4.4}$$

using linear regression. The average model fits (Pearson *r*) and parameters (*a*, *b*) are given in Table 4.2 and fit the measured results well: In

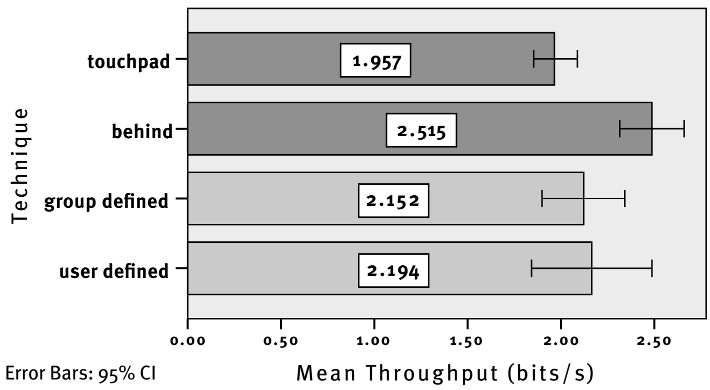


Figure 4.6: Grand throughput of interaction techniques (since *group* and *user defined* yielded a comparably high error rate close to or above 10% their calculated throughput values may be less meaningful).

particular, it shows the lower initial time required to start moving in mid-air (a of *behind*) as well as the smaller slope b of *touchpad* that indicates faster movement on the touchscreen for targets with a higher *ID*.

Technique	a	b	r
Touchpad	442.33	333.5	.953
Behind	-222.17	492.67	.937
Group defined	-735.42	795	.862
User defined	-563.25	741.58	.938

Table 4.2: Fitts' Law parameters and model fits.

After the study we asked participants to rate each interaction technique again on the 10-point Likert scale (1 – very bad to 10 – very good interaction) from before. Here *touchpad* performed best (average rating prior experiment 7.31, after the experiment 7.38) directly followed by *behind* (6.31, 7.31) that increased an entire point. Conversely, decreasing differences were found between *user defined* (7.46, 6.31) and *group defined* (6.46, 5.85).

We collected participants' ratings (Likert scale 1 – 7) after completing the tasks with each technique. Participants rated perceived speed, accuracy, fatigue of different body parts and selected questions from the Nasa TLX [107]. Participants' feedback delivered an overall similar picture to quantitative results in terms of speed, precision, difficulty, user satisfaction and precision, with the latter being experienced slightly better for *touchpad* ($\tilde{x} = 6$) than *behind* ($\tilde{x} = 5$). As expected, overall fatigue was the lowest and almost non-existent with *touchpad*. *Behind* was rated the second best on fatigue scales overall (left/right finger, hand, wrist, and shoulder) but was rated one point worse than other point-

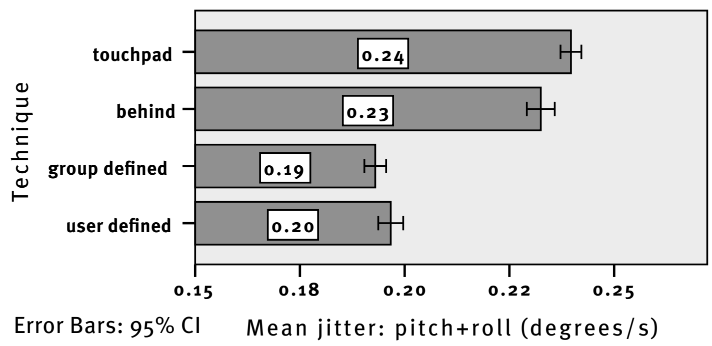


Figure 4.7: Mean phone jitter (pitch+roll) measured as the sum of differences over time.

ing techniques for left hand, left arm, and left shoulder fatigue – the body parts involved in holding the projector. This can be explained by the fact that the projector had to be held slightly further away from the upper body not to interfere with the pointing right hand. The latter is supported by our analysis of phone jitter (Figure 4.7) which shows *behind* to cause the highest jitter among the mid-air pointing techniques. For right-sided body parts fatigue was rated as almost non-existent in contrast.

4.4.4 Discussion of First Experiment

Contrary to our initial expectations, the experiment revealed a significant difference between *behind* and the other techniques. The difference between *behind* and *group defined* / *user defined* can mainly be explained by the fact that the independent group of 27 people who provided the information for the input space of *group defined* and the participants of our study preferred on average an area on the right top side of the projector phone. Users seem to choose this area because it allows them to move the right arm freely, unrestricted by the upper body or the projector phone. The negative implication of this area is that upper arm, lower arm, and finger have to be controlled simultaneously. Based on our results it seems that most participants were not able to control the attitude of their pointing arm exactly and steadily enough in those two interaction techniques. This caused pointing jitter, inaccurate pointing and arm fatigue.

In contrast, when using the *behind* technique participants were able to rest their upper arm of the pointing hand on the upper part of their body. Therefore, they had to control only their lower arm and index finger, which allowed accurate and steady pointing and led to lower arm fatigue. The results show that those advantages outweigh the disadvantage of *behind* that is the slightly limited input space. For instance

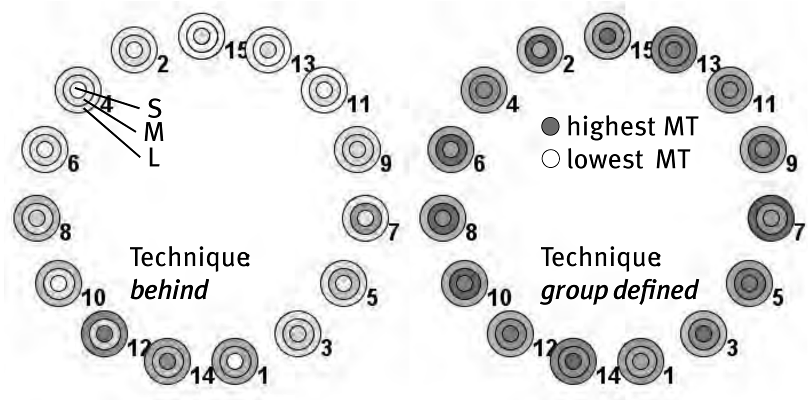


Figure 4.8: Target heat maps for *behind* and *group defined* averaged over all users depicting movement times (MT) for the target sizes (S, M, L). The overall superiority of *behind* is clearly visible, despite the problematic area at the bottom left.

it is more difficult to select areas on the bottom-left of the input space (see Figure 4.8), especially for corpulent and female users.

Compared to *touchpad* the interaction technique *behind* has the significantly lower movement times because the user needs less time to start moving in the air whereas *touchpad* requires to place the finger on the screen and overcome the initial resistance on the surface. However, *behind* is more vulnerable to errors for small and medium sized targets because of hand jitter and arm fatigue. This is less an issue with *touchpad* because it is easier to brake or rest the finger on the touchscreen surface. In real usage scenarios it will likely depend on the type of application whether the faster movement time or the higher error rate will have the higher impact. For instance, *behind* will likely perform worse than *touchpad* for text entry on the projection as errors are very frustrating for the user in this scenario. During browsing a website on the other hand, being 15% quicker in general might easily compensate for missing every 18th link (5.6% error rate). Furthermore, if the application made good use of the dual-display setup enabled by mid-air techniques like *behind*, e.g. a browser showing an overview of open tabs on the touchscreen and the currently active tab on the projection, the user interface could benefit further in terms of speed, clarity and user satisfaction.

4.5 SECOND EXPERIMENT: APPLICABILITY

Before we can study dual-display mobile applications with projector phones, though, we need to test how the superior mid-air technique *behind* compares to the *touchpad* technique in *nomadic* real world application scenarios with *unaltered* mobile applications. We further added

(a) Setup for *behind* technique (similar to *touchpad*).

(b) Setup for *touchscreen* technique (user sitting at desk).

Figure 4.9: Setup for second study.

the standard mobile touchscreen usage as third technique to the comparison that would allow participants and us to distinguish between the impact of the projection and the interaction technique.

4.5.1 Participants

For the second experiment we recruited the same 12 participants from the first experiment to ensure they had the same amount of practice with the projection techniques. All of them owned a laptop and were hence familiar with touchpads and all but P4 and P7 (10 of 12 participants) owned a touchscreen phone themselves (only a few featured multi-touch or VGA+ resolution screens, though).

4.5.2 Experimental Design

The second experiment comprised two independent variables *technique* and *application*. Techniques consisted of *touchscreen* (the application was used on the mobile touchscreen without projection, Figure 4.9b), and *touchpad* and *behind* (application was used on the projection, controlled via a cursor, Figure 4.9a). We decided to test four specific applications that are likely to benefit from the larger projection or the different input technique instead of fielding the projection in tasks that are optimized for and advantageous (like, e.g., private text entry) on the touchscreen. The four applications and reasons for choosing them were as follows:

4.5.2.1 *Browsing App*

Browsing has become one of the most common tasks performed on smartphones. With mobile phones reaching display resolutions comparable to laptops, websites can be used in "full site" or "desktop view" mode instead of their usually very restricted mobile versions. However, due to the small physical display size, this requires several zooming and panning operations by the user. In contrast, on the projection even small text can easily be read without zooming.

We used the standard Android browser in full screen mode (Figure 4.10) in all three techniques. The participant always started with a Wikipedia article about San Francisco. Starting on this web page the experimenter asked the participant to follow one of three predefined paths (counterbalanced). On every path the participant had to scroll down to the table of contents of the article, and then navigate to one of three predefined sections (e.g. museums). Then, the participant had to perform twice: following a link (e.g. to the Wikipedia article of the Museum of Modern Art) and finding a certain piece of information (e.g. when the museum was established). All tasks required roughly the same amount of scrolling, reading, clicking and time.

Figure 4.10: The Android Browser in the two projected conditions

4.5.2.2 *Map Navigation App*

Figure 4.11: The Google Maps application with larger zoom buttons in the projected modes.

Maps are another very prevalent mobile application. Often, after having used a text-based search for initial navigation, a subsequent action is to orientate oneself around the found location, which we took up for the second task. Starting from "Schützenbahn, Essen" where the user study took place, the user had to go approximately 70km directly north, east, or south to another large city like Bochum, Dortmund, etc. (the order was counterbalanced). Participants used the standard Android maps application with gesture support and zoom buttons in touch screen mode, and the standard Android maps widget without gesture support but with twice as large zoom buttons in projected mode (Figure 4.11). We anticipated that orientation was quicker and less demanding on the projected display since smaller city names and icons could be read more easily. Since we could not provide advanced gestures like pinch-to-zoom with the projected techniques (at least not with behind), we also anticipated navigation on the projection to be slower than on the touch screen alone which provided these gestures.

4.5.2.3 *Gaming App*

Figure 4.12: The game "Drunken Hunting" from the Android Play Store in the two projected conditions.

As games are ultimately diverse we acknowledge that a single game cannot be representative. However, it can provide a preliminary sense for a particular group of mobile gaming applications. Since a shooting game resembles much of the Fitts' Law tapping task, yet in a completely different setting, the popular app "Drunken Hunting" seemed to be a reasonable candidate. The goal in this game (Figure 4.12) is simply to shoot flying ducks by touching or pointing and clicking on them, respectively. In contrast to other similar simple shooting games, it features targets at different sizes with shooting smaller ones yielding more points than larger ones. We anticipated that smaller targets would be easier to see and hit while displayed on the (large) projection than on the (small) screen because of the bigger size and the eliminated fat-finger problem. Every participant played two levels with each level comprising 10 shots.

4.5.2.4 *Painting App (for Accurate Steering)*

Figure 4.13: The "Paint Joy" app from the Android Play store which allows tracing outlines (shown for projected conditions).

With mobile phones taking over increasingly more traditional PC tasks, accurate pointing *and* steering gains importance. Painting combines both requirements very well and the considerable number of downloads of painting applications in the app stores shows their increasing distribution. One obvious problem with painting, though, is the lacking accuracy caused by the fat-finger problem. With this application we want to research if the

(a) *Behind* with browsing (left) and gaming.

(b) Device equilibration **(c)** Touchpad

Figure 4.14: 2nd study prototype and techniques in use.

usage on the projection with the presented techniques increases the accuracy during the task.

We used the in 2012 most widespread painting application "Paint Joy" from the Android Play store. The task of the participant was to post-paint the outlines of a snail with house (Figure 4.13). This image was chosen because it combines horizontal, vertical, and circular lines – the basic subset of every more complex painting task.

4.5.3 Prototype and Setup

4.5.3.1 *Hardware Setup*

For the second experiment we employed a different prototype as we wanted to maximize the user experience of different display sizes between phone and projected display. We therefore used a Samsung Galaxy Nexus Android phone featuring 720p HD resolution. This phone offered the highest physical display size and resolution available on mobile phones at that time. Hence it seemed to be the strongest competitor against a projection. Similarly, we wanted to provide a large, bright, and high quality projection. Since none of the available pico projectors supported HD resolutions or a brightness beyond 50 Lumens, we opted for the palm-sized projector Qumi Q2 from Vivitek. This projector provides the same 720p HD resolution and a brightness of 300 Lumens while still only weighing 617g. Phone and projector thus weighed 742g together. Even though this was possible to hold in one hand and use for a short time, we decided to additionally uphold the projector from a rod affixed to a tripod moving freely in all directions. Thereby we equilibrated the weight of the projector to some extent, but it still had to be uphold and steered by the user as it would have without equilibration (Figure 4.14b).

The phone was attached in landscape mode (Android's default when connected to a projector) on a flexible plastic attached to the bottom of the projector. This construction allowed the user to hold the "projector phone" with one hand in *behind* mode (Figure 4.14a) and two hands in *touchpad* mode (Figure 4.14c). Participants could stand and hold the device comfortably while looking on an almost leveled projection, yet were required the typical balancing to preserve the levelness and position of the projection and cope with hand jitter as with a real projector phone.

4.5.3.2 *Software*

The pointing software was realized as an Android background service, which showed a shiny green cursor on top of all other Android windows and applications and intercepted all user touch events. Our background service processed these events and based on the current mode of interaction (*touchpad* or *behind*) sent them as new touch events to Android's input system. The latter was accomplished using Android's built-in monkey service, which we hijacked on our rooted device to send arbitrary touch events to the system. Additionally we attached to the native Linux events from the touchscreen. Overall, this gave us full control over Android's touch input handling to send our own events to the Android system and its built-in applications.

In both projector interaction modes clicking anywhere on the device resulted in a click at the current position of the cursor. In *touchpad* mode the cursor position was changed relatively to movement of the finger on the device (same as in the first experiment). Scrolling in the browser application and painting in the paint application were initiated with a double click from where on movement of the finger was passed through to the application until the finger was lifted up again (in *browsing* the cursor position remained fixed during scrolling). In *behind* mode the cursor was moved by moving the finger in mid-air just as in the first experiment, relying on finger position data acquired from the OptiTrack tracking system. In this case, scrolling and painting was executed while the finger was down on the touchscreen, i.e. the website was "grabbed" with the left hand's finger and moved up or down by moving the right hand's finger in the air. The game only required positioning and clicking to shoot.

4.5.4 Procedure

We employed a within-subjects design as in the first experiment. Each participant tried each of the three techniques with each of the three applications (counterbalanced). Each application was used with each technique between 2 and 3 minutes. We followed a qualitative analysis approach that would reveal differences that have not become apparent in the first study. We instructed participants to think aloud during all interactions, which we recorded for later analysis. After having tried all 9 combinations we asked participants for their feedback about speed, accuracy, liking, joy of use (each on a 7 point Likert scale), advantages and disadvantages of each technique and the projection in general. We were also interested in when, where, and for which applications participants would favor using the projection over using the touchscreen alone.

4.5.5 Results of Second Experiment

Overall, the projection techniques were liked much by participants and more fun to use than *touchscreen* as reported by 9 of the 12 participants. Partly, this has to be attributed to the novelty effect. Nevertheless, it indicates a positive user experience with both projection techniques, albeit being highly dependent on the application type.

In the browsing and map navigation tasks *touchpad* was perceived as slower than other techniques by at least four participants since *touchpad* required a double-click to initiate scrolling/panning which four participants perceived to slow down the interaction. In contrast, *behind* was reported to be very fast (P4, P5, P9) and precise (P3, P6, P10) as was *touchscreen* (7 and 8 participants respectively) despite the required zooming and panning steps. Especially in these tasks, those participants owning a high class smartphone and therefore being trained on getting by with the small screen for browsing and navigating performed much better with *touchscreen* while for novice smartphone users both systems seemed to perform equally well.

In the gaming task participants scored most successfully with the *behind* technique, which also felt intuitive (P1, P2, P8), but also became more aware of the freehand pointing jitter. P8 and P9 said "it was difficult to keep still". *touchpad* was more affected by clutching than in other scenarios, as moving the pointer over long distances from a previous shooting target to the next required more than one movement. 4 participants said they felt constricted by the small touchscreen compared to the large projection (P6 said "I didn't know where I was on the screen with my finger"). But *touchpad* was on the other hand perceived as the most precise (6 participants) since targets that did not move too fast could be hit more accurately than with any other technique. The *touchscreen* also performed fast but showed the problems that very small ducks could not be recognized on the small screen and that the finger occluded the targets at the expense of accuracy.

In the painting task 9 participants reported the fat-finger problem to hinder accurate painting on the *touchscreen*. Yet, *touchscreen* (Figure 4.9b) performed much better than *behind*, which was very unsatisfactory to use because of the comparably high jitter. Despite *behind*'s lower caused jitter compared to other mid-air techniques, the jitter is still too high for the technique to be qualified for steering tasks. Here, *touchpad* showed its huge advantage in that it, as P10 said, "combines the advantages of projection and touch-screen", namely the elimination of the fat-finger problem on one side and the haptic affordance of the touchscreen on the other that improves precision.

Further comments, independent of application, included that *behind* is an interactive performance like playing Wii (P5), which can be liked or disliked (as by P3 and P10 in our case). 3 participants also stated they would like to perform the click in the air, too, which we had thought about before but decided to stick to bimanual input as this will likely be the standard use case in future projector phone interaction. With *touchpad* participants liked that it feels familiar from laptops (P1, P3, P7) and requires little space (P2, P4) as well as little effort (P3, P5, P10) and therefore is more versatile in its application than *behind*. But it also requires a lot of movement on the touchscreen surface, which got uncomfortable over time for P7 and P9.

4.5.6 Discussion of Second Experiment

The second experiment has shown that mobile applications indeed can benefit from a mobile projection. Despite private or public media broadcasting and collaboration, the projection can even enhance unaltered mobile applications that originally have been designed for touchscreens. Further, the advantages of the projection are very co-dependent on the usage scenario and can for instance be very useful to overcome the fat-finger problem on touchscreens or to increase the visibility and ease the selection of small objects on the display. Based on these findings we predict that new application-specific interaction techniques that sensibly integrate touch and mid-air interaction on both displays will largely enrich the projector phone experience. Nevertheless, the tasks related to information management (browsing and map navigation) suffered from the imprecise mid-air or slow touchpad technique, respectively. Here, the direct touch control provided by ordinary touchscreen interaction was favored by the majority of participants.

4.6 LIMITATIONS

Even with comfortable arm postures such as with the *behind* technique, mid-air interaction might lead to higher fatigue than traditional solutions. Luckily, mobile situations rarely entail lengthy series of interactions. Further investigation of fatigue, especially on larger mobile projections, also in respect to different C-D gains that might affect speed, accuracy, and fatigue, is required. In our studies we did not experiment with different C-D gains: the lower bound of C-D gain was set by the physical size of the touchscreen that we did not want to exceed to maintain comparability. Higher C-D gains in contrast might have decreased the accuracy of mid-air techniques further.

When testing the applicability of the techniques in common mobile scenarios we did not include all mobile factors such as interacting on-the-go or sudden breaks. In contrast to touch input, the mid-air techniques forbid pausing of the cursor as long as the user's hand is within the input area. Furthermore, we only evaluated existing mobile applications specifically designed for touch input. Studying applications designed for dual-display mid-air interaction will deliver further interesting results. Finally, people used to multi-touch performed the tasks of the second study more quickly on the touchscreen, albeit acknowledging many advantages of the projection. However, the majority of participants were unacquainted with multi-touch for why we implemented *touchpad* interaction similar to laptop touchpads.

4.7 CONCLUSION

Projector phones raise various questions regarding their interaction design due to the large remote display, availability of various sensors, and movement of the projection. Most available commercial projector phones only mirror the information displayed on the phone display on the projection, involving frequent context switches and unsuited user interfaces. Using the *touchpad* interaction technique already provides a significant advantage as the user can focus primarily on the projection. But in real world scenarios the indirectness, the effect of clutching and the occupation of the screen diminish its applicability.

Our first study showed that more direct pointing using *behind* provides distinct advantages in terms of movement time and throughput when compared with *touchpad*, in particular when considering medium and large targets. The notably higher error rate of *behind* however makes it more suitable to application scenarios such as browsing and gaming and less to painting or text input. Interestingly, *behind* performed better than *group* and *user defined* although the latter two were gathered through a user-elicited approach (Subsection 1.3.1.3).

The second study analyzed the user experience of *behind* and *touchpad* in relationship to conventional *touchscreen* usage. Here, we have seen that projection-based techniques (*behind, touchpad*) received overall equally good feedback as *touchscreen* despite not having been explicitly designed for the chosen applications. Furthermore, participants saw various disadvantages of touchscreens such as the fat-finger problem and numerous advantages of using a projection, e.g., for collaboration and application areas such as gaming.

> *R1 page 8*

In regard to research question R1, considering remote handheld projection we must conclude that a performance increase is highly dependent on the type of application. For information management tasks, we have seen that participants preferred and were more efficient using direct touch interaction for various reasons such as hand tremor during longer operations of *behind* or the indirect control of *touchpad* involving clutching. These human factor constraints affirm previous technical considerations towards within-reach interaction for nomadic projected interfaces.

As outlined before, the following part will introduce Nomadic Projection Within Reach more formally and propose and evaluate several device concepts for nomadic information management which precisely support this direct touch interaction style.

Part II

NOMADIC PROJECTION WITHIN REACH

5 THE NOMADIC PROJECTION WITHIN REACH FRAMEWORK

Previous chapters have been the motivation for defining a framework that enables nomadic information management through mobile projection. The investigation on out-of-reach projection and interaction in the previous chapter has already hinted to some advantages of projections like providing a better overview, a better input accuracy in some scenarios, and solving the fat-finger problem. However, *within-reach* interaction showed better performance for information management tasks.

Apart from that, Table 2.1 (page 29) revealed the high amount of Lumens that mobile projectors required to achieve distant projections with an acceptable contrast under even the most moderate ambient lighting conditions. A projection at a coffee place—an example for a typically rather low-lit environment to support a relaxed atmosphere—from only one meter away already exceeded the light output of currently available projectors (cf. Figure 2.3) by the factor 3–5 and from $1.5\,m$ distance even by the factor 4–8. In more typical indoor lighting as living or office rooms, this mismatch even doubles or quadruples (factor 16–32). As mobile projector efficiency increased only linearly, at a slow rate of 10 Lumens per year (Figure 2.3), it is unclear when and if at all mobile projection will be mature enough to cover these projection distances adequately.

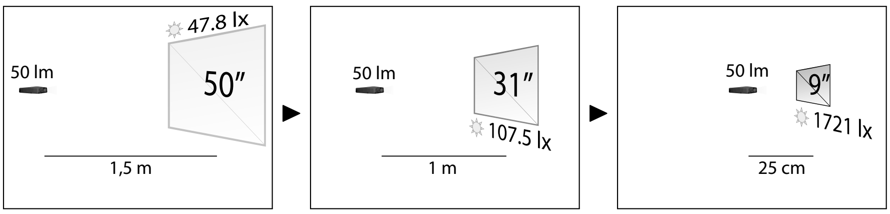


Figure 5.1: Decreasing the distance to the projection surface diminishes the size of the projection, but increases its luminance and thereby contrast and legibility. Relations assume a throw ratio of 1.1, based on the Microvision SHOWWX+ HDMI (which only provides 15 Lumens though).

Conversely, when the projection distance comes *within reach* at about $25\,cm$ away from the device[1] the projection becomes comparably small,

1 This considers handheld usage, thus the *within reach* distance is not defined by a typical arm length but by the distance between a projector held in hand in front of the body and the reachable distance to all corners of a projection in front of it, which is significantly shorter.

only the size of a tablet display, but its luminance increases, for instance, ≈36 times compared to a distance 1.5 meters away (see Figure 5.1). Assuming a 50 Lumens projector, the resulting projection provides a luminance exceeding $1600\,lx$ at this distance, thus providing more than the fourfold of office lighting ($400\,lx$) resulting in an acceptable display contrast. Even in spite of high ambient indoor lighting ($\geq 1000\,lx$), it will still provide an exceeding amount of light ensuring the display's visibility. At a size of, e.g. about 9" diagonal, it has the size of tablet displays, the fourfold size of typical smartphone displays, and most importantly, can be used *in addition* to them.

These two motivations, technical reasons speaking for within-reach *output*, and human factors speaking for within-reach *input*, lead us to defining the Nomadic Projection Within Reach (NPWR) framework for nomadic interaction and information management, which promotes:

1. *not* to use handheld projection, as it occupies the hands and does not support touch interaction in a meaningful way which is advantageous to PIM-related functionality. Instead, the projection device should either be worn or arranged for it to be easily put down, for instance on a table. If it is worn, counter measures must be taken against shaking of the projection during interaction.
2. to use projection distances *within reach* between 25 cm (with current hardware) and 50 cm (with hardware available in the near future) to allow for comfortable touch interaction and to achieve a sufficient luminance of the projection even in uncontrolled lighting environments. Even as projector brightness will increase in the future, so will other display technologies and thereby the rising expectations of users to the displays they use. Thus we can assume that this relation will even hold for the foreseeable future.
3. to leverage affordances in the environment such as tables, paper, cups, the floor or affordances of the device such as its lid—not only for output but all the more for new and expressive input modalities.

Looking at related work on settled and out-of-reach projection, regarding the deficiencies of current nomadic information management (Subsection 1.1.1), the projected display bears the following hypothesized advantages:

1. If it is considerably larger than typical mobile device screens, it might increase overview and decrease interaction steps and task completion times (e.g. [W9]).
2. It might enable multi-tasking possibilities such as available on projected tabletops (e.g. [233]) or in smart rooms using projection (e.g. [271]).

3. If it depicts an *additional* display second to an existing one, it might enable privacy-respectful interaction using one display for public and another one for private display (e.g., [62]).
4. It might enable new display locations and surfaces leading MMDEs and unprecedented user experiences through the use of AR (e.g., [104, 129]).
5. It might increase information and privacy awareness in single-user (e.g. [150]) and multi-user scenarios (e.g. [100]).

The following three case studies investigate this potential by applying Nomadic Projection Within Reach to nomadic information management scenarios. Each case study will explain the addressed deficiencies at the beginning as well as state them on the right hand margin (blue color means addressed). Each chapter's conclusion will relate the findings of the case study to the achieved mitigation of the deficiencies and the initial research questions (Section 1.2). These findings will later be summarized and generalized in design guidelines located at the end of this thesis (Chapter 11).

6 CASE STUDY ON NOMADIC DUAL-DISPLAY PEN+PAPER INTERACTION

This chapter presents the Penbook, a tablet that is extended with a built-in projector and integrated with a wireless pen and that uses real paper on the back of its lid as projection surface. This allows using the pen to write or sketch digital information with light on the projection surface while having the distinct tactility of a pen moving over paper. The touch screen can be used in parallel with the projected information turning the tablet into a dual-display device. Without considerably enlarging the device, it provides a larger output/input area (>D1) that can be leveraged to use multiple applications (such as browsing and taking notes) or multiple windows (such as different views on patient data) simultaneously for multi-tasking (>D2). The augmentation of real paper with projected ink shows a further unique advantage of a projected mobile interface, allowing to include parts of the environment into the interaction (>D4).

Deficiencies ***addressed by this chapter***

Output/input size *(D1)*

Multi-tasking *(D2)*

Collaboration & Privacy *(D3)*

Environment *(D4)*

Related video

This chapter is based on the previously published refereed conference paper

[W8] **Winkler, C.**, Seifert, J., Reinartz, C., Krahmer, P., Rukzio, E., "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, **Best Note Award**. New York, NY, USA: ACM, 2013, pp. 283–286

and extends this in particular by providing more details on the design considerations, envisioned usage scenarios, as well as implications of the case study on Nomadic Projection Within Reach.

6.1 INTRODUCTION

Thinking of mobile devices that can easily fit a projector, tablets quickly come to mind. To support Nomadic Projection Within Reach, they require the ability to stand on their own, for instance on a table. When work on this case study started, the first tablet cases (like those of Apple's iPad) already allowed for placing tablets upright by using a special folding mechanism of the screen cover's iPad. Later, tablets appeared that included a stand in their body, like Microsoft's Surface RT and Surface Pro series. A tablet device standing on its own has a quite large space around it that could be leveraged for projection. When considering single-user scenarios, the space in front of the tablet seems to be very suitable as it can easily be reached, is partly shielded against sunlight both by the device itself as well as the operating user. The height of the standing tablet allows for a projected display that can easily double or triple the display area of the device at almost no increase of the size of the device. Projecting a *touchable* display in front of the user is possible, for instance by adding a depth or at least infrared camera to the device (cf. Subsection 2.5.3). However, it would likely not be able to mirror the accuracy of the tablet's capacitive touchscreen. Adding an area for pen and handwriting is another compelling alternative, especially since it would provide an additional input modality and thereby an overall more expressive interaction with the device.

In this regard, it is interesting to note how natural handwriting still impressively withstands the digital revolution. Even today, pen and paper are used for its

- *flexibility*, for instance when taking notes during a presentation or sketching ideas,
- its *input fidelity and resolution* that outperforms digital imitations such that legal bindingness almost exclusively relies on real hand writing,
- ability of using their personal chirography instead of a keyboard.

Of course, the disadvantages of handwriting in today's digital world should not go unmentioned:

- undo using Tippex is a lot harder than in software,
- the content is not immediately digitally available (with all its detriment effects on sharing and remote collaboration),
- and readability depends on the handwriting, to name a few.

Nowadays, ever more tablets and very large smartphones, almost the size of a tablet (phablets) are equipped with pens and digitizers that allow for some of the advantages without entailing all of the disadvantages. However, they do not afford the haptic feeling and input fidelity of natural pen and paper as the thick glass of mobile devices hinders accurate writing, a comparable fidelity (range of pressure and contact angles), and comparable resolutions.

Based on these considerations, the EU project "Hospital Engineering"[1] sought for a nomadic device that patients, nurses, and doctors use throughout their day and which affords a large output space and a multi-modal (including handwriting) input space without sacrificing mobility. The result of this is the Penbook, the first tablet device with integrated projector. It is a novel multi-display device that besides touch

1 http://www.hospital-engineering.org/, retrieved November 19th, 2015

interaction provided by the tablet supports hand-written input using a digital pen on real paper in the lid of the device using augmented projection (see Figure 6.1). The device is the result of several interviews and discussions my colleague Christian Reinartz undertook with three nurses and two doctors within the project that involved about 50 German hospitals and companies working in the domain.

The following sections will explain the assessed requirements, the derived design considerations and concept, the implementation of a fully functional standalone prototype and several applications in the hospital and private domain together with their evaluation.

6.1.1 Pen + Paper in the Hospital Domain

Especially in the hospital domain, the low cost and flexibility of paper-based forms helped them to impressively withstand the digital revolution to the greatest extent. Their low cost is the most prominent of reasons for their persistence, i.e. they can be developed much cheaper than software. They also provide greater flexibility, i.e. forms can be altered and unplanned annotations can be added, and accommodate an innate legal bindingness when patients fill and sign these forms. For instance, patient registration, patient anamnesis, or the signing of prescriptions are just a few use cases that rely on the flexibility and ease of use of pen-based input. However, information written on paper documents needs to be transcribed into a digital form for further storage and quick access. This process is expensive and time consuming. In addition, paper-based forms are highly limited in terms of interactively supporting users to fill in the required information in each field. Yet this is easily possible with digital devices that can display additional help instructions in any circumstances. Hence, this raises the challenge of how to bridge the physical and the digital world while preserving the benefits of both.

6.1.2 Requirements analysis

Semi-structured interviews with three nurses and two doctors led to the identification of three typical scenarios, which requirements were assessed and analyzed in a subsequent step. The scenarios are patient registration, patient anamnesis, and prescription signing, each of which will be discussed in the following.

6.1.2.1 *Patient Registration*

In the context of patient registration, experts told us that, for instance, surgery forms can become very long and complex and many patients

Figure 6.1: The Penbook setup: a tablet computer with an integrated projector augments a paper on its lid, transforming it to a display that supports digital hand-writing with light instead of ink.

need help at least once while filling them out. Often, input space is too small for more difficult portrayals and together with the lacking support to undo written text, the final appearance of these forms often becomes unaesthetic and therefore hard to read for employees. Digital forms have the potential to resolve all these issues, but would imply that patients receive digital devices with e.g., soft keyboards that not all patients know to handle and could thus demand time-consuming support by the staff and lack support for signature. To combine the advantages of both worlds, we assessed that the registration form should still behave and feel like real paper, but with the added possibilities of undoing actions, optional space enlargement, and in-situ help through a connected device.

6.1.2.2 *Patient Anamnesis*

Patient anamnesis is in principal well suited for digitalization. Many parts of anamnesis forms carry simple information, e.g. aching in left knee, and can hence be mapped to check- and radio-boxes. However, they usually require the doctor to mark impairments or changes on x-rays or organ schemata. This demands very precise and flexible hand-writing which cannot be supported on digital mobile devices with the same fidelity as on paper. Even devices that support digital pens still suffer from the missing haptics, the optical misalignment, and reflections introduced by the thick screen glass, and usually sensor precision lying much below 600dpi. We assessed that handwriting precision was a critical supporting feature.

6.1.2.3 *Signing of Prescription Forms*

Finally, the process of signing prescription forms depicts a big disadvantage of the physicality of paper. In hospitals, patients often have to wait a long time for their prescription because the doctor does not have time to retrieve the prescription forms from the printer and sign them. A personal digital device of the doctor could address this problem by alerting to and presenting prescription forms to be signed. But as the unique signature of the doctor is key to prevent malicious usage of prescription forms, support for precise handwritten input is indispensable.

6.1.3 Specific Related Work

In the medical domain, many systems employed pen interaction to provide intuitive interaction to physicians [200], but did not try to mimic real paper and did not employ multiple displays. Research prototypes such as Hinckley et al.'s Codex explored the design space of mobile dual-screen devices with pen and touch input [112], yet without including real paper or a projected display that benefits haptics and the device's form factor. Research on pen and touch modalities was extended by Hinckley et al. by considering the combination of both modalities which yields new interaction techniques [115]. Many works have further dealt with the support of information gathering e.g. [111], while our work focused on the support of existing paper-based workflows.

6.2 CONCEPT OF THE PENBOOK

From the requirements assessed with experts, we developed the Penbook. It consists of two main components: an upright standing touchscreen tablet with a mounting support and an integrated pico projector

Figure 6.2: Interacting with the Penbook: (left) touch on the upright screen as well as pen-based input on the paper-based projection screen; (middle) pen-based input by using the rear side of the pen; (right) bi-manual interaction enables interaction with pen and touch simultaneously.

at its top, and a paper canvas, layered by an Anoto pattern and attached to the Penbook's cover that is used as a projection screen and for input via a digital pen (see Figure 6.1). This setup retains the mobility of a tablet device, yet enables different new options for interaction. In the following, we detail interaction techniques and illustrate their application using the aforementioned application scenarios.

6.2.1 Options for Interacting with Penbook

Penbook enables user input in four different configurations:

1. *Pen-based handwriting*. Unlike conventional digital pens that write with ink and additionally store a digital version of the user's handwriting, Penbook's pen leverages the projection to *write with light* instead of ink (see Figure 6.1). That is, the pen does not create permanent drawings on the paper canvas; rather the system traces the pen's movements and projects the paths onto the paper canvas. It is thereby not bound to conventional limitations of digital paper pens. For instance, drawing parameters such as color and stroke thickness can be changed. Also, the user can scroll within the drawing area by moving the pen over a scrollbar at the side of the paper, and further, drawings can be undone. At the same time, it keeps the haptic affordances of working with real pen and paper. For instance, this interaction is well suited for annotating content on the paper or the touch-screen (through annotation links) with hand-written notes. When real ink is desired, techniques such as PhotoScription [108] could be added in the future to make drawings permanent.
2. *Touch-based input*. The touch screen allows users to perform multi-touch operations (see Figure 6.2 (left)).
3. *Pen-touch input*. The rear side of the pen allows for touch input on the upright touch screen (see Figure 6.2 (middle)). Hence, users are not obligated to put aside the pen when operating the system with only one hand available.
4. Bi-manual input. When using Penbook while sitting, users can interact bi-manually with the system. That is, while using the pen for input on the projected screen, the other hand is available for touch-based input on the touch screen (see Figure 6.2 (right)).

The flexible coupling of the displays allows for seamlessly transitioning the Penbook between a dual-display laptop-like posture and a standard single-display touch screen posture by folding the cover behind the device (see second of left image in Figure 6.4). These distinct postures cover a large amount of interactions and usage scenarios.

6.2.2 Application Scenarios

In the following, we describe how Penbook solves many issues of the analog workflows in the aforementioned hospital scenarios without constricting their flexibility. We further describe applications for private nomadic usage.

6.2.2.1 *Patient Registration*

Penbook presents patients with a paper-like form on the projected screen, which does not look much different from traditional paper forms. Therefore, patients should know how to fill in the form without prior training, as the changing of colors and tip sizes is not required in this scenario. Deleting strokes by crossing them out is detected by the system automatically. The touch screen further informs the user about an available interactive help feature which is triggered by touching information circles next to the corresponding input field. Touching the information circle brings up a large and thorough explanation regarding the corresponding input field and possible input examples on the touchscreen (see Figure 6.3 (left)). Further, the digital input allows for optical character recognition (OCR) and digital storage in the background, without the user noticing it or having to deal with its implications, e.g. correcting writing errors. The patient's signature whose penmanship fits the rest of the form's content assures the same legal bindingness as traditional paper forms. Thus, Penbook does not constrain the existing advantages of patient registration forms, but augments the experience with useful features such as undo, interactive help, and additional space because forms can also be scrolled horizontally to reveal space beside the paper boundaries.

6.2.2.2 *Patient Anamnesis*

Penbook offers a digital version of typical anamnesis forms. Instead of shuffling paper stacks to find the correct form, doctors choose from a list of available forms. Most parts of the form consist of check- or select-boxes. But as in traditional paper forms, every digital form also has a schematic organ view that is shown on the projected screen when a form is selected (see Figure 6.1 (right)). Doctors use it to precisely annotate impairments or changes to the medical condition on the paper, also using different colors and brush sizes. Here, the process is changed more radically, but retains and surmounts the level of precision and flexibility provided by paper, as required in the annotation context.

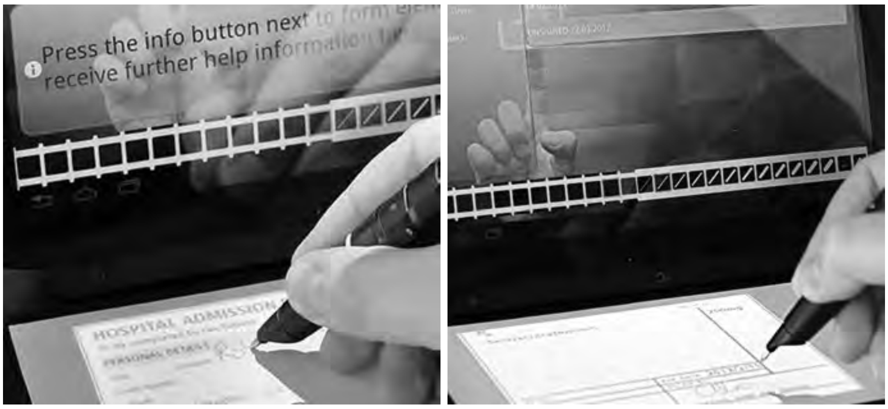


Figure 6.3: Penbook supporting diverse tasks in the hospital context: interactive patient registration (left); the issuing of prescriptions (right).

6.2.2.3 *Issuing Prescriptions*

Prescription signing requires a hand-written signature for authorization purposes. Using the Penbook enables natural signing of prescriptions as the prescription form and area for signing can be projected onto the paper (see Figure 6.3 (right)). When finished, the complete signed form is digitally available and can be encrypted and transmitted to the front desk to be printed automatically. As the Penbook is highly mobile, doctors can sign prescriptions directly after treatment.

6.2.2.4 *Scratchpad for Semantic Notes*

Apart from applications for the hospital domain, an obvious functionality of the paper-area is to provide semantic note taking, i.e. notes that are linked to their origin. As one example, we have implemented a scratchpad browser addon that links any taken note to the currently opened website. While doing an online research, for instance, to compare different tour, hotel, or flight operators for an upcoming trip, excessive amounts of open websites and information quickly pile up and are oftentimes difficult to overview in the end. With the scratchpad application, the user maintains a *shared* notepad area on the paper that works across all websites, but links any taken note to the currently opened website. After all options have been explored and corresponding notes taken, a simple tap on the note brings the user back to the corresponding website, eventually allowing more overview and a better comparison and decision at the end.

Instead of a shared scratchpad between multiple documents, another obvious functionality is to provide a per-item notepad to an application. A note taking facilities of PDF viewers, for instance, could easily be integrated with the Penbook to allow per slide annotations on the paper area. Here again, we see the advantage of the larger display space

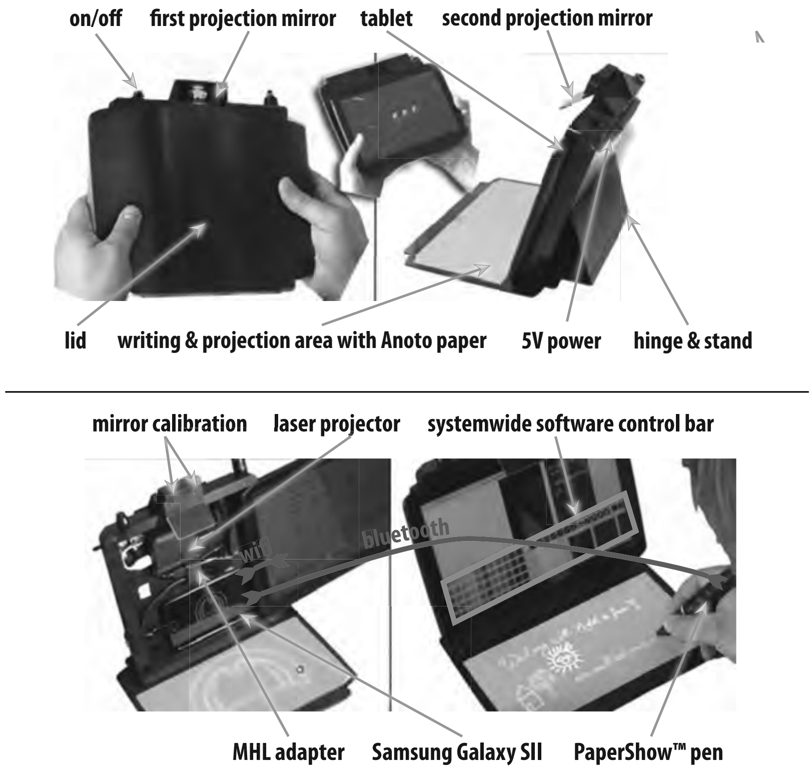


Figure 6.4: Hardware and software components of the Penbook protoype. The tablet and pen are both connected to the smartphone behind the tablet, which acts as server and source of the projector.

such that even large notes not longer interfere with the slide (as they do in current PDF viewers) but can be taken and viewed alongside the slide. Visual lines would link the note and the point of reference between the displays.

6.3 IMPLEMENTATION

In the following, first the custom-built hardware setup and its components will be explained. Afterwards, the software architecture will be detailed.

6.3.1 Hardware Design

As there neither was a tablet device with built-in projector nor a case with a paper-like cover available at the time of implementation, such a system was designed and built. It consists of four main components, depicted by Figure 6.4, which are described in the following.

1. A tablet computer (Motorola Zoom 2) that serves as the upright touch screen.
2. A smartphone (Samsung Galaxy SII) connected to a laser pico projector (Microvision SHOWWX+ HDMI) in order to control the projected screen and its content. The optical path of the projection passes two mirrors in order to increase its overall length to about 25 cm. Despite the low brightness of the projector, the projection can be seen very well in standard indoor lighting due to the short distance to the projection surface and the innate sharp focus of the laser projector (cf. Section 3.3). These two devices are powered by an external USB battery (that is enabled by the button at the top of the device) to ensure a sufficient runtime of the device.
3. An Anoto pen, which supports hand-written input on paper that is equipped with a specific pattern. The pen does not dispense ink; instead, strokes made with the pen are projected onto the paper. The pen tracks the Anoto pattern on the paper to track its position on the projection screen using a built-in infrared camera. The information is sent via Bluetooth to the mobile phone. The communication protocol of the pen was reverse-engineered and a driver written to connect the digital pen to Android that allows fine-grained control over its features. Additionally, a capacitive cap was added to the tail of the pen and the pen wrapped with capacitive seam to enable pen-touch input on the touch screen. The abovementioned components are integrated into an aluminum case with a flexible stand and a foldable cover that contains the projection screen.

The commercial availability of projector phones (e.g. Samsung Galaxy Beam) indicated that the device could be built in a form factor similar to standard tablet devices. Recently, several tablet devices with integrated projector appeared on the market (*Aiptek ProjectorPad P70*, *Lenovo Yoga Tablet 2 Pro*, *YF- X9 7,0*, Isonic Protab 7 HD) whose development may have been motivated by the Penbook.

6.3.2 Software

The Penbook software components are jointly distributed between the devices.

The server component executed on the mobile phone is connected to the digital pen. As soon as the pen is activated (by taking it out of its holder that is fixed to the case), the connection is initialized and the pen continuously sends location data. Depending on the current setting for pen color and stroke thickness, new strokes are added to the projected image. The image is pre-warped to appear correctly on the projection screen (which is not orthogonal to the projector) after passing through the optical path that includes two mirrors. The necessary onetime calibration is facilitated by an interactive application that defines pre-warping and mapping of the reference system/position of the pen to the projection area. Pre-warping consists of mapping the projection corners to the corners of the paper's surface, and correcting the lens and perspective distortion of the projector. All algorithms are implemented as OpenGLES shaders, thus do not introduce noticeable delay.

Further, the mobile phone runs a Wi-Fi hotspot which allows the tablet computer to connect and communicate via an Application Programming Interface (API) to set or get various parameters of the current system state. These include changing pen features, setting or getting the projected image, attaching or deleting meta-information from any place on the projection screen, and storing and loading calibration data.

The client component running on the tablet computer controls the user applications. It leverages the API of the server to trigger the calibration procedure, change pen or paper (such as graph and ruled paper) properties, set or retrieve the projected output. The client is written as an Android fragment view that can be added to any Android application. It adds an expandable control bar at the bottom of the application that provides access to all system features.

6.4 INITIAL USER FEEDBACK

Before the hardware can be studied in a longer field trial (see Conclusion), we conducted an initial usability study with 10 non-professional participants (5 female) of average an age of 29 years (25-33 years). The objective was to understand how users interact with the device when coping with a task requiring parallel usage of touchscreen and pen for creating annotations.

Initially, we demonstrated all interaction features of the Penbook to participants and they had the opportunity to make themselves familiar with the prototype. Then we introduced them to a task, which was to browse a website and search for items (i.e., products) matching specific criteria (e.g., price) while taking notes using the scratchpad application described in Subsection 6.2.2.4. The goal was to decide which is the best available option by the help of the overview and backlink features of the scratchpad.

After the practical part, we interviewed participants about their usage experiences. 9 of 10 participants emphasized that they would buy and use the Penbook if it was commercially available. Many of them stated it would be perfect for taking notes on the slides during a lecture. One stated that he did not like his handwriting and that he was much faster with a hardware keyboard. 5 participants highlighted the interconnection between tablet and projection, i.e. the backlink feature, as useful. 9 participants expressed that they very much appreciate the haptic affordance of real paper for writing (including one person who owns a high-class convertible laptop with a digital pen).

These results show a generally very positive opinion of participants towards the Penbook, and most participants cherished the benefits of the Penbook over traditional tablet computers. As part of the *Hospital Engineering* project, the Penbook is part of a model hospital that allows to gain in-depth insights to how patients and physicians take advantage of multi-modal interaction in a real world setting.

6.5 CONCLUSION

This case study has presented the Penbook, a nomadic multi-display device that supports writing with light on a built-in real sheet of paper without significantly enlarging the device's form factor. It bridges the gap between the digital world and paper-based workflows as it combines the benefits of both characteristics in one device. Based on a domain-specific design process, a first hardware and software solution with example applications has been presented which especially in the hospital domain that still relies heavily on paper-based records may aid the shift towards digital devices. Results of the presented preliminary user study indicate that users highly appreciate the concept and find it easy to understand and use. Further, as the tablet's built-in camera watches the projection area through the second mirror, future work may explore use cases that include object and paper tracking above the projection area.

More generally, the Penbook addresses existing mobile deficiencies in the following ways:

OUTPUT/INPUT SIZE At almost no cost to the size of the device, the Penbook provides a second display, doubling its in- and output area. This is mainly achieved because the projection is able to leverage an existing surface—the lid of the case—which the majority of tablets bring anyway.

MULTI-TASKING The Penbook supports using different applications (the note taking application allowed creating and accessing notes while surfing the web) and multiple parts of the same application (such as forms and organ schemata in the medical applications) simultaneously.

ENVIRONMENT The augmentation of real paper provides for an alternative input modality currently not possible with other display technologies. Further augmentations may include existing paper forms or the own hands to be tracked by the tablet's camera and be integrated into the interaction. This seems promising as the projector could augment the hands with context-aware control options (such as a pen color on each finger) and feedback on gesture recognition. Regarding research question R2 we can add pen-based input on in reality occurring surfaces as well as hand augmentations to the list of enabled input modalities. > *R2 page 8*

Smartphones even support nomadicity considerably better than tablets and the next chapter will explore how advantages of the Penbook can be transferred to smartphones and advanced to support multi-user interaction as well.

7

CASE STUDY ON NOMADIC MULTIUSER EVERYWHERE TABLETOP INTERACTION

Deficiencies addressed by this chapter

***Output/input size** (D1)*

***Multi-tasking** (D2)*

***Collaboration & Privacy** (D3)*

***Environment** (D4)*

Related video

The previous chapter explored increasing display and input size/modalities of a tablet device. The setup of the device supported specific single-user scenarios but not yet collaboration between multiple users. In Subsection 1.1.1.3 we have already seen how current nomadic devices, especially smartphones, lack support for (privacy-respectful) collaboration. This chapter will thus extend the previous idea to support ad-hoc collaboration (>D3) using projected interfaces. To support the mobility of the user even further, this time a smartphone instead of a tablet will be equipped with a projector. Instead of the support for pen-based input, the projections of individual devices allow to be merged to larger shared interactive surfaces, allowing to leverage furniture in the environment (>D4) for an AR experience impossible with screen-based displays. Nonetheless, as with the Penbook device, the projected interface is also useful for the single user, for instance to multi-task showing the active browser tab on the projected display and other open tabs on the phone.

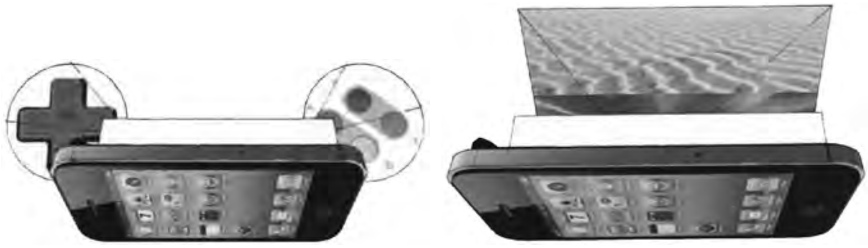
The next sections will introduce the concept of the SurfacePhone and fully functional prototypes together with their evaluations in user studies.

This chapter is based on the previously published refereed conference paper

[W5] **Winkler, C.**, Löchtefeld, M., Dobbelstein, D., Krüger, A., Rukzio, E., "SurfacePhone: A Mobile Projection Device for Single- and Multiuser Everywhere Tabletop Interaction." In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14. New York, NY, USA: ACM, 2014, pp. 3513–3522

and extends this in particular by providing more detailed results of the first user study, more iterations and an explanation of the tracking pipeline of the (high-fidelity) technical prototype, and the implications of the case study on Nomadic Projection Within Reach.
In addition, the following partially related thesis was supervised by the author:

- "Entwicklung von Anwendungsszenarien und kamerabasierter Fingererkennung für ein iPhone mit integrierter Pico-Projektion". Pascal Spengler. Bachelor's thesis. 2012



(a) External controls **(b)** Picture presentation

Figure 7.1: Initial concept sketches showing externalization of game controls to address the fat-finger problem and privacy-respectful picture presentation.

7.1 INTRODUCTION

Personal computing more and more transitions to mobile devices. Sometimes, this involves more complex tasks involving multiple or multi-windowed applications for which mobile devices do not offer enough screen space. But users should not be forced to use (and own) completely different devices for these scenarios that demand bigger displays. An increasing number of devices tries to address this through multi-display solutions. Besides devices based on multiple screens such as the Nintendo DS or Hinckley's Codex [112], projector phones set out to enable the exploration of large-scale content and support collaboration in a mobile setting.

MMDEs like the aforementioned ones that consist of multiple screens, still are limited by the maximum size of the device. Hence, their possible increase in display estate is comparably low, usually not more than twice the single display size (cf. Nintendo DS, Kyocera Echo). In contrast, MMDEs including a projected display allow for a much higher increase in display estate, especially as compared to the small display size of a smartphone, while still keeping its small form-factor. However the display setup of current projector

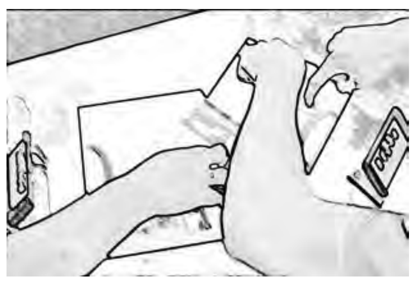


Figure 7.2: Two users merge their projections for a larger shared surface.

phones such as the Samsung Galaxy Beam is often characterized by two displays being visually separated, i.e. not in the same field of view (e.g., one on the phone the other one on the wall). This setup precludes many of the prevalent sharing and collaboration techniques that are well known and investigated for example in multi-device interaction on today's multi-touch tabletop systems (cf. [229]).

In contrast, the setup of the SurfacePhone allows to recreate such tabletop-like interactions in mobile scenarios with the added benefit of providing a private and a public display. Similar to the Penbook—only this time for a smartphone—the SurfacePhone can be placed on a surface to augment it with a second (or multiple) projected display(s), only this time right *behind* itself (see Figure 7.1). Because both displays are in the same field of view—different to existing projector phones or other works on MMDEs like [133]—they are well suited for all kinds of extended single-user scenarios as well as sharing and collaboration between multiple-users. The projected display is not only touch- and gesture-enabled, but additionally orientation aware, which allows multiple projected displays to be optically combined and merged to a single shared display (see Figure 7.2).

After reviewing some related work specific to MMDEs, the design process of the SurfacePhone will be presented. Starting with the considerations for such a system and the envisioned usage concepts, two prototypes will be presented that have been evaluated in user studies. The initial concept prototype allowed for easy evaluation of the concepts and ideas. The technical smartphone case prototype was developed to show and evaluate the technical feasibility of the SurfacePhone concept. The chapter concludes with design guidelines derived from the results of the studies and their implications on Nomadic Projection Within Reach in general.

7.2 SPECIFIC RELATED WORK

Some aspects of MMDE devices have already been explored in research [72, 112, 133, 168]. With Codex, Hinckley et al. [112] created a dual-display device that allows for re-orientation of two physical hinged displays. Besides, they also explored various application scenarios for different configurations as well as novel interaction techniques. Nevertheless the display arrangement of the here presented SurfacePhone has not been mentioned nor explored in the related work, probably because the arrangement would not provide much merit with multiple small screen displays but only with projected interfaces.

Kane et al. explored with Bonfire a laptop equipped with a pico-projector that allows to create a secondary display right next to the notebook

[133]. Even though the SurfacePhone also makes use of a similar configuration, one of its main aspects is to explore collaboration on a potentially mutual interactive surface with multiple devices. This requires new sharing techniques between displays and devices and dynamic merging of their projections, which will be presented and enable new public/private interaction scenarios. Furthermore, the SurfacePhone improves on touch accuracy and portability and evaluates touch and gesture recognition which could not have been carried out before.

Merging multiple displays to a single one has been shown before using screens [114, 136, 161]. However, the thick bezels between the devices hinder the experience of a single display, both output and input-wise. Further on, the brightness of the image is optimized towards orthogonal view-angles and quickly declines towards lateral ones. Most of all, supporting a configuration like the SurfacePhone that provides a private view for each user and a large combined one to the group would require every user to carry a smartphone *and* a tablet device, whereas the SurfacePhone requires a projector phone only.

Finally, with the PlayAnywhere system, Wilson [270] had presented the idea of a mobile tabletop system that uses projection. The SurfacePhone extends this idea by providing a truly mobile and nomadic device that further provides an additional personal display to facilitate private content management and decision making as well exemplified by Schmidt et al. [229] or Seifert et al. [233].

7.3 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 application scenarios for *single device/single user* (SDSU), *single device/multi user* (SDMU), and *multi device/multi user* (MDMU). All of these will be discussed in the following.

7.3.1 Position and size of projection

Hinckley et al. [112] 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 or Penbook posture, thus a rather private setup. This is because the projection is mainly visible to the user facing the device. A projection to either sides of the

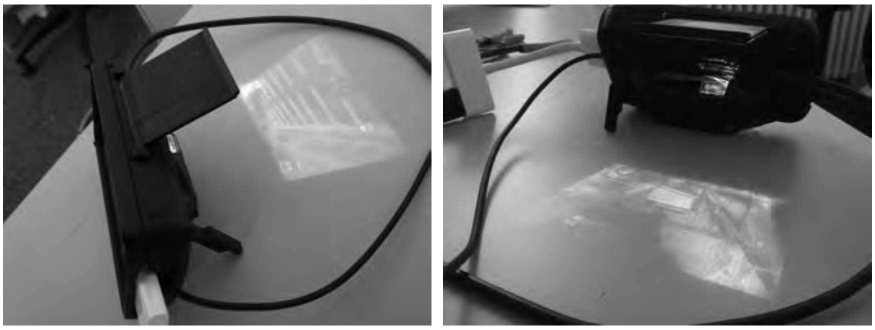


Figure 7.3: First implementation of the real prototype for an iPhone 4. Improvable is the rather small projection size (only 4"), the not very robust stand, and the device requiring external power and an external video connection between iPhone and projector.

phone would imitate the setup of Bonfire [133], where the projected surface is still within easy reach of the user, but more public than in the laptop scenario. Still they do not necessarily invite other people to interact with the projection.

Whereas these two configurations have been explored intensively, a projection behind an upright standing phone has been neglected so far. The latter—the configuration of the SurfacePhone—consists of a public projected display and a private display (as can be seen in Figure 7.4b) and presents a more collaboration-oriented setup. To some extent, it resembles the Battleship setup of Codex [112], albeit the difference that the primary user (usually the owner of the device) 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 by the primary user who has to circumvent the phone to touch the projection.

Additionally, this MMDE setup follows the recommendations of [68] to diminish visual separation between displays. 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.

Details of our technical prototype can be found in Section 7.5. But to give an idea of the size and position of the projection early-on: through several iterations of the prototype (cf. Figure 7.3) an optimal (undistorted) projection behind the phone was found to measure around 17 cm × 14, cm in size (8.7" diagonal) and to lie 14 cm behind the phone and 4 cm to the left of the center of the device. The latter accounts for the offset of the phone camera sitting at the very side of the phone. The projection, thus, adds a more than four times bigger display to the 4" screen of the iPhone 5 for which the latest iteration was designed.

7.3.2 Configurations

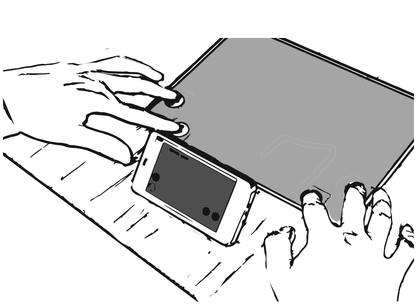
The SurfacePhone can be used alone (single device / single user – SDSU), or by multiple users using one (single device / multiple users – SDMU) or multiple SurfacePhones (multi-device / multi-user – MDMU). This section describes the application and interaction space of these configurations.

7.3.2.1 *Single-device, single-user interaction (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 7.4a) to the projection or showing the main view of the game on the projection. Apart from that, the projected display could be used as general secondary display, for instance, showing a task manager or notifications of applications currently running on the device. Finally, phone screens are very useful for augmenting the reality of the user, but cannot serve publicly visible augmentation. 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.

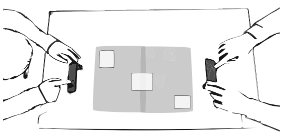
7.3.2.2 *Single-device, multi-user interaction (SDMU)*

Leveraging the inherent differences in publicity of the displays, the SurfacePhone can be used for several sharing tasks in small groups (compare Figure 7.4b). 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 give out their phone to other people; that the content can be presented to all people simultaneously; and that only specific pictures or slides for presentation 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 touch interaction.

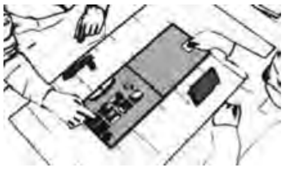


(a) SDSU

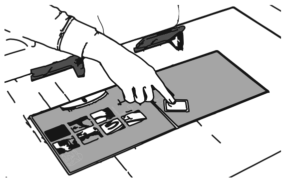
(b) SDMU



(c) MDMU facing front



(d) MDMU facing side



(e) MDMU same side

Figure 7.4: (a) SDSU: a single user increases their display estate fivefold to support multitasking or externalization of controls. (b) SDMU: One user is presenting pictures to another user. MDMU: (c) Two users sitting face to face with the projections merged at the long side. (d) Two users sitting face to face with the projection merged at the short side. (e) Sitting next to each other on the same side with the projections merged at the short side.

7.3.2.3 *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 7.4e) 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 can first explore offers on their personal devices, then share it to the surface. Being able to also see other users' phone screens may significantly improve communication in collaborative planning.

On the opposite, sitting FACE TO FACE (Figure 7.4c) 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 7.4d) 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 > 2 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.

7.3.3 Interaction Techniques

In the following required interaction techniques for the SurfacePhone will be discussed that suit aforementioned application and usage scenarios. Here we draw from users' experience and familiarity with smartphones and tabletop systems to find intuitive and still technically feasible interaction techniques. The technical feasibility will be discussed in the implementation section of the technical prototype (Section 7.4).

With today's prevalence of multi-touch interaction, users would expect to be able to interact with the projected content using *direct touch* which is in line with the Nomadic Projection Within Reach concept anyway. 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. As the phone camera is watching the scene from above, mid-air gestures above the projection could also be considered.

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.

Concept-wise, it seems interesting to explore if and how the invisible space in the center of such a ring of projections—dead space by its own—can be leveraged for interaction. During a card game, for instance, cards played by the players could be animated towards the invisible center and eventually from there to the player receiving the trick. If each played card was visualized on each player's projection, seeing the trick in-between would not be necessary. Still the space would be leveraged to mimic spatial relations of real game play. When playing roulette with many users, the play field could run across the ring and additionally a portion of the roulette wheel with everything inside beyond the number compartments be visually hidden (see Figure 7.5). Findings from [W10] further suggest that users embrace hidden parts in an application that are not revealed by the projection as challenging part of the game.

This could, for instance, come in handy when many SurfacePhones merge so that they form a ring rather than a central space. Then, 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 & Taylor [236] 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 in normal lighting environments. By letting each device blink a unique pattern every now and then, devices could possibly be uniquely and spatially detected. However, this requires further research and evaluation.

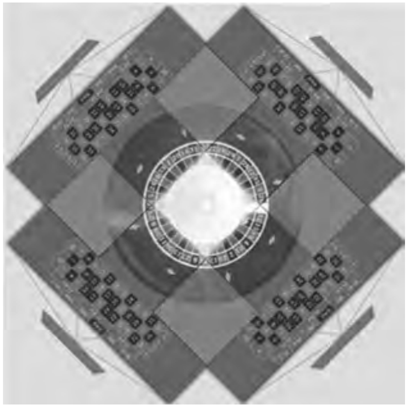
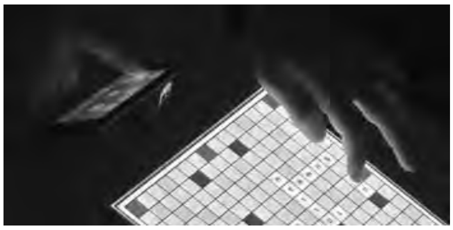


Figure 7.5: Illustration of 4 users merging their projections to play roulette, utilizing the projection-lacking center as wheel. The phone display would show the current credit in form of chips that can be placed using a transfer technique, e.g. *swipe*.

As the SurfacePhone is a mobile device, movement of the device 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 (paragraph Spotlight metaphor, Subsection 2.5.4). Any table could thus become the personal virtual desktop that is explorable by moving the SurfacePhone across the table, from notes to documents, to reminders, to pictures of beloved ones, etc.

7.3.3.1 *Transfer Techniques*

A regularly occurring task when using the SurfacePhone is to transfer content from the screen to the projection and vice-versa. Following on [34] we can distinguish between three main categories of transfer-techniques that can be supported: *direct*, *binned*, and *mediated* transfer.



(a) Direct transfer: Placing a scrabble piece at a precise position on the board through simultaneous touch.

(b) Binned transfer: Elements from the bin element (here the bench) on the phone are placed on the projection using *touch-swipe*.

(c) Mediated transfer: The presenter drags another picture on the proxy element at the top of the phone.

Figure 7.6: The explored techniques for content transfer between displays.

- *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 interaction category, *Human Link* is proposed to be used. The body of the user is conceptually used as a medium to transfer the content between the two displays (cf. [271]). 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 7.6a).
- *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 7.6b). 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 an object it is simply dragged &

dropped on the *proxy* to appear close to the proxy element on the other display (Figure 7.6c).

7.4 CONCEPT PROTOTYPE

To explore and evaluate the SurfacePhone concept, a concept prototype was built 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 (see Figure 7.7). This allows to test users' experiences providing a more robust, responsive, and clearer multitouch surface than would have been possible through developing a technical prototype (which is presented later on) in the same time. The following will discuss the implementation of the concept prototype, its applications, and a qualitative user study using the concept prototype.

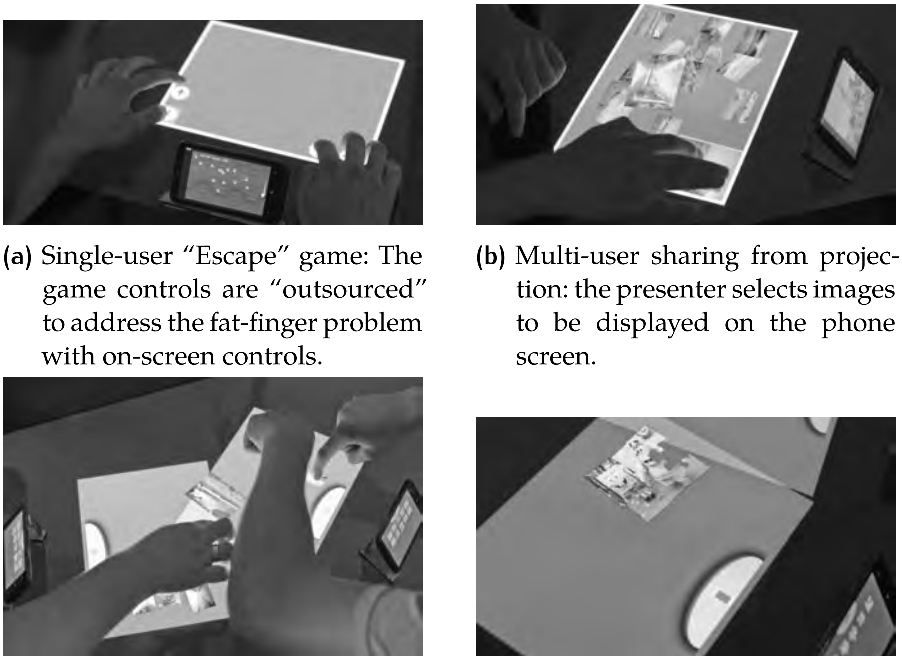
7.4.1 Implementation

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× 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 are able to build devices that support projections of these dimensions by using short-throw lenses or curved mirrors (e.g. LG PF1000U), it can be assumed that 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 (as in Figure 7.7c), or something different like a shared playboard within the concave hull of the projections' corners (as in Figure 7.6a).

7.4.2 Applications

7.4.2.1 *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



(a) Single-user "Escape" game: The game controls are "outsourced" to address the fat-finger problem with on-screen controls.

(b) Multi-user sharing from projection: the presenter selects images to be displayed on the phone screen.

(c) Multi-device: The users merge their projections to share images from and to their private display and with the other parties.

(d) Multi-device: Both users have pieces for a collaborative puzzle they are supposed to solve on the merged projection.

Figure 7.7: Four of the five apps used in the concept study in different configurations (see Figure 7.6 for the remaining scrabble app).

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, presumably users will perform better in the projected mode (Figure 7.7a).

7.4.2.2 *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 7.6c). The other possibility is that users browse their content on the projection and present items fullscreen on the phone by double tapping them (Figure 7.7b). Different to the first way, the user gives up their privacy for the benefit of having a larger space themselves that can be explored more quickly.

7.4.2.3 *Multi-device picture sharing (MDMU)*

The exemplary picture sharing—which would similarly work with other content types—is very similar 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 merged 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.

7.4.2.4 *Multi-device scrabble game (MDMU)*

The scrabble application (Figures 7.6a and 7.6b) particularly emphasizes the private display on the mobile phones. It shows a standard scrabble playboard on the merged projections. The phone screens show the letters available to the users and a virtual bench on the bottom where words can be arranged with the letters using drag&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 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.

7.4.3 User Study

With the first user study the quality in terms of usefulness, applicability, and usability of the overall SurfacePhone concept and its several components is assessed using the presented 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.

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

7.4.3.1 *Procedure*

First we explained the concept of the SurfacePhone by showing them a concept design (similar to Figure 7.10a) of the technical prototype and to convince them that similar 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 (Figure 7.7d).

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 [126]). The study was video captured and later qualitatively analyzed. The textual answers were later analyzed using Grounded Theory and axial coding.

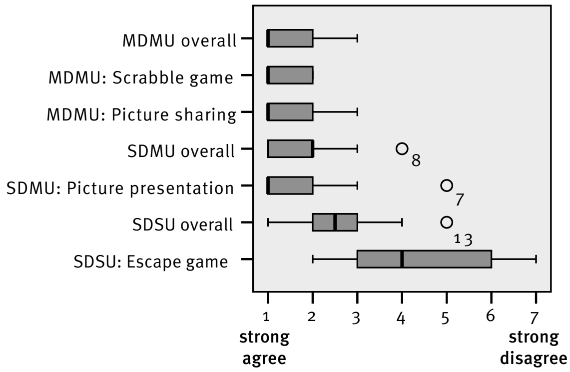


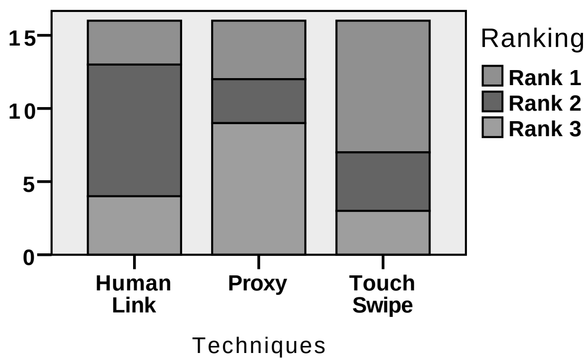
Figure 7.8: Participants' ratings on the usefulness of the different configuration and application scenarios in nomadic everyday computing scenarios after the first study.

7.4.3.2 *Results*

In terms of preferred interaction technique, to transfer information between the two screens, most participants favored *touch-swipe* (see Figure 7.9). 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 two thought they would be better with either *Human Link* or *proxy*. Their comments revealed that for the majority (9 out of 16) when using Human Link they faced the problem that when they touched the phone's screen and tried to touch the projection at the same time the phone would slide away. The absence of bi-manual control was especially seen as an advantage 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,

Figure 7.9: Participants' final ranking of the transfer techniques at the end of the first study.



allowing both participants to view each other's phone displays, helping them to identify the correct pictures before putting them on the surface. Three other couples divided the work between in 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 such a device 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 fulfillment of hardware constrains such as reasonable size and battery life to be mandatory.

Asked openly for advantages of the SurfacePhone, all but two participants brought up the advantage of the private display on their own. In the words of P15: "Everybody owns their private area and everybody can push something towards the shared table center. Everyone can interact with their phone and the table area (nobody must watch only) and moving data between devices is easy"; or similarly in the words of P1: "shared and public space, everyone decides what to share and has no control [over] or access [to the private area] unless I grant it". As a follow-up, we asked users about which of their data they have privacy concerns and if they see these addressed by the SurfacePhone. All users mentioned pictures as their main privacy concern, but some also mentioned documents and games. Several people stressed that not only the type of content but also the people you are sharing with are a key factor and that the SurfacePhone should provide default configurations like "family sharing" or "work sharing" where copying of pictures was by default allowed or disallowed, respectively. The existing means of copy protection by quickly withdrawing with the projection was regarded as useful but not forceful enough to cover every situation. Overall, participants seemed very aware of the *collaboration* and *privacy* deficiencies of current nomadic computing and saw these for the most addressed by the SurfacePhone. Ten participants further stated that they liked that they do not need to pass their device around when presenting to groups. Regarding privacy and the conditions in the study that compared two privacy-preserving picture presentation techniques—one where the picture was selected on the phone screen and the projection used to present it and the other the other way round—the former seemed more reasonable to the majority of participants.

The preferred combination of devices and users was MDMU (Figure 7.8 depicts participants' preferences regarding application scenarios). Participants found the possibilities that arise from having a mobile device that is able to create ad-hoc complex mobile multi-display environments 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 investigate 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% as a mean of system usefulness (87.3%) and interface quality (81.9%)[1]. 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.

Negative comments included that for the merging functionality, a sufficient market penetration is required to make the feature applicable and useful in nomadic scenarios. While this is true and shared with screen stitching approaches that require the same software to run on all devices, it is noteworthy that the other scenario groups (SDSU, SDMU) work on their own and with only one device.

7.5 TECHNICAL PROTOTYPE

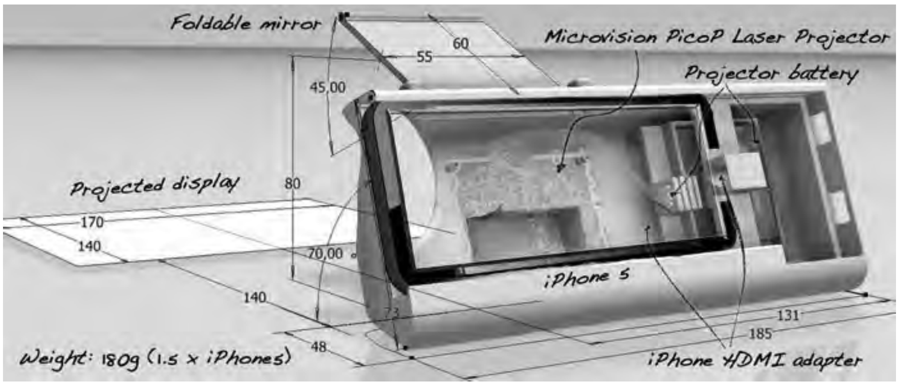
The positive results of the first study were motivation enough to build a technical SurfacePhone prototype that can support aforementioned interactions. The aim with this prototype is twofold: Firstly, the technical requirements and challenges for such a device shall be investigated and corresponding solutions found. Secondly, it shall serve as prototype for another user study that delivers quantitative measures how well touches and gestures can be performed by users and detected by the system.

7.5.1 Hardware Design

As no similar device configuration has been presented so far, the author started from scratch. The iPhone platform was chosen, since it was the only mobile platform that allowed two different outputs on screen and projection at the time of investigation. After several different projection engines (e.g. TI DLP 2 or Microvision PicoP) attached to the backside of the device had been tested, it became obvious that without a fitting short-throw lens that was not available, the size of the projection would become too small. The problem could be solved by attaching a mirror to the top of the phone and the projector 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.

The first explorations resulted in a custom-made case for an iPhone 4 that already demonstrated the concept well but which provided only

1 Documentation quality was not applicable and thus left out.



(a) The technical prototype design (dimensions in mm).

(b) The implemented prototype that tracks finger touches using in-built camera and accelerometer. The red border (for illustration) 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).

Figure 7.10: Design and implementation of the prototype

a 4" projected display, was not standing robustly, and was not working standalone, yet (the inclined reader might take another look at Figure 7.3).

After further iterations that tried to include a separate battery, maximize the distance between projection and mirror, and steepen the projection angle further, the final design resulted in the projector case as depicted in Figure 7.10a. Besides the changes to the projector's light path, it uses an iPhone 5 for better performance. 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. In the camera image that is sampled at 640× 480 the projection appears between $P_{TopLeft}\{183, 238\}$, $P_{BottomRight}\{590, 316\}$, thus takes up 407px in X and 78px in Y direction (Figure 7.10b). Obviously, specifically the resolution in Y direction is quite small and the projection is not centered in the image. Nonetheless, this is the best compromise that was found between maximizing the size of the projection and still completely seeing the projection in the camera and also regarding overall performance of the standalone system.

7.5.2 Implementation

Following our previous design considerations, the prototype should support direct touch on the projection, 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. First, intrinsic and extrinsic camera parameters need to be calibrated using printed chessboard and projected chessboard patterns respectively. Having these parameters we can map the projected area from object space to an interpolated orthogonal view of the projected region (see Figure 7.11a) and use this for tracking. In a final sturdy SurfacePhone device 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. This step will be automated in the future when the phone API allowed to measure focus and exposure at the center of the projected display, which, for instance, modern Android-based phones already do.

7.5.2.1 *Finger Position*

In step 4, we capture a still frame (Figure 7.12c) for subsequent background subtraction. As the background of interaction can be arbitrary we use background subtraction to separate moving fingers from the background (Figure 7.12d). 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.

Our following tracking pipeline (see Figure 7.12 on page 121) runs at 22 FPS. To eliminate shadows as best as possible we first convert the camera image to the HSV space and then work on the saturation channel (Figure 7.12e). The literature recommends working on the hue channel to eliminate shadows, but we found that table 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 (Figure 7.12g). Because of the steep camera angle in our setup, blob area sizes can range from a few pixels to sizes that fill half of the image. This makes the classification of correct blobs more difficult. Further, we cannot rely on standard CV techniques like convexity defects for finding fingertips as often there is only one finger plus parts of the thumb in our image which do not provide the defects information (Figure 7.12h). Instead, the algorithm we developed first computes the convex hull of the contour and its normalized approximation. Then, for each point on this new contour that has a tangential slope of less than 15deg with its surrounding points, it calculates a probability that this point is the fingertip by minimizing equation:

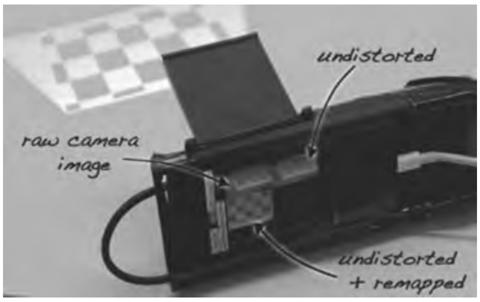
$$P(Fingertip) = \frac{(SCD * W_{SCD} - DCP * W_{DCP} + SA)}{AREA} \quad (7.1)$$

- where SCD (Figure 7.12k) is the second closest distance of the point to the corners of the bounding box. A finger should create a very rectangular bounding box where the fingertip lies almost at the center of the smaller side of the rectangle yielding small distances to the two closest corner points;
- where DCP (Figure 7.12i) is the summed distances between the point and corner points of the bounding box that lie on or outside of the edge of the camera frame. A correct fingertip of a pointing finger will always have maximum distance to the hand center. As we do not see the hand the corner point is only an average guess;
- where SA (Figure 7.12j) is the estimated area of the fingertip above the current point. This is calculated as the sum of distances between up to 15 surrounding points on the contour to both sides. Correct fingertips should yield smaller results than arbitrarily shaped blobs with peak endpoints;
- where $AREA$ is the size of the blob;
- and where W_{SCD} and W_{DCP} are weights found by experiment set to 10 and 5 respectively.

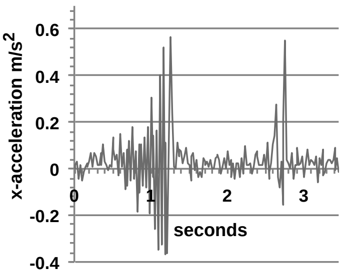
Finally we have to decide which blob represents the primary finger, which one is a possible second finger and which ones are not of interest. Our blob sorting and filtering algorithm favors blobs with fingers, high finger probabilities, less circular shape (to filter out hand areas) and lower Y position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

7.5.2.2 *Finger Touch and Touch Modalities*

To support different *touch modalities* we cannot rely on the 2D camera image as small changes in depth are indistinguishable from small changes in height for touch recognition. Kane et al. in their Bonfire system used a combination of position tracking using the camera and touch recognition using an accelerometer that measures the touch vibration on the table [133]. This approach seems the most promising as the hardware is readily available in most (projector) phones—in contrast to e.g. Harrison et al. [104] and Wilson [270]. However, their cam-



(a) Calibration proof mode: The projector projects a chessboard, the screen shows different proof views.

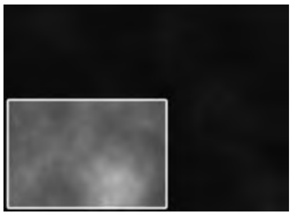


(b) First a strong, then a light touch, logged on the SurfacePhone.

Figure 7.11: Implementation details

(a) Scene

(b) Camera image



(c) BG capture

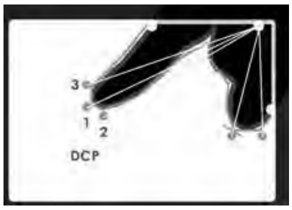
(d) BG subtraction

(e) Shadow removal

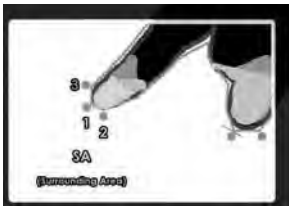
(f) Hand segmentation

(g) Contour finding

(h) Possible fingertips

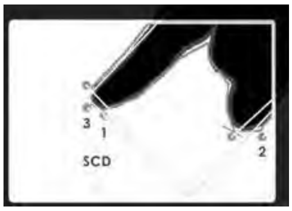


(i) Heuristic DCP



(j) Heuristic SA

(k) Heuristic SCD



(l) Same as (k) but with non-optimal guess

Figure 7.12: Computer vision pipeline to recognize the user's fingertip. As only very little of the user's hand is visible to the recognition algorithm, standard algorithms like convexity defects do not yield reliable results (h) so that several heuristics (i-k) are required to decide between possible fingertips (h).

era angle was almost orthogonal which simplified finger tracking. Further, they used the accelerometer of a laptop that shares a much larger space with the surface than our prototype and therefore allowed to work with simple thresholds. Correctly classifying touches with the SurfacePhone seems like a bigger challenge.

Furthermore, when relying on surface vibrations for touch detection we cannot distinguish between touch down and up events, for example, to classify a touch as "long touch". Similarly, double taps cannot be recognized as successive events since the second vibration could overlay the first one. We can, however, recognize the intensity of the touch quite well since light and strong touches create distinct vibration patterns (see Figure 7.11b). These two modalities, light and strong, can, for instance, be used to start dragging of an item using a strong touch.

For the touch detection we measure the current acceleration in X direction every 10ms. Then we compute the touch vibration as the difference between the averaged sum of the absolute amplitudes of the recent 150ms (15 values) and the previously calibrated sensor noise. Based on thresholds we then decide whether the measured vibration corresponds to a strong, a light, or no touch at all. The default strong threshold is twice the default light threshold. As not all surfaces transport vibrations equally, we also implemented a detection procedure that vibrates the SurfacePhone with a constant pulse and measures the resulting phone vibration. Based on our tests with different tables, lightweight tables will be good mediums resulting in low phone vibrations (down to $0.1m/s^2$) and good touch recognition whereas strong tables will not pass on the vibration very well, resulting in phone vibrations up to $0.5m/s^2$. Through this procedure we can adjust the default thresholds to increase touch accuracy on different tables.

7.5.2.3 *Gesture Recognition*

Since in our setup we only see small parts of the user's hand, gesture support of hand postures does not make much sense. However, we can well recognize gestures that are based on a trajectory of movement such as directional gestures (left, right, top, down swipes) or more complex gestures like a circle. 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. [278], for instance, to recognize a circle gesture.

(a) Touch task (b) Gesture task

Figure 7.13: Tasks performed in the second user study. (a) Touching small (shown in figure) and large targets, with light and strong touches, at different positions. (b) performing one of five gestures above the projection.

7.5.2.4 *Detection of other SurfacePhones*

Although not fully integrated into the prototype yet, we evaluated the use of Qualcomm's Vuforia on the SurfacePhone using *projected* frame- and image markers. Our interest was to see how the steep camera angle and the much lower resolution of the projected image compared to printed markers affects the recognition algorithm. Fortunately, the recognition worked better than expected. Frame markers of a size of 1/8 of the projection size are perfectly recognized as soon as they are completely visible in the scene. Image targets are recognized even up to a 1/50 of the projection (see Figure 7.14).

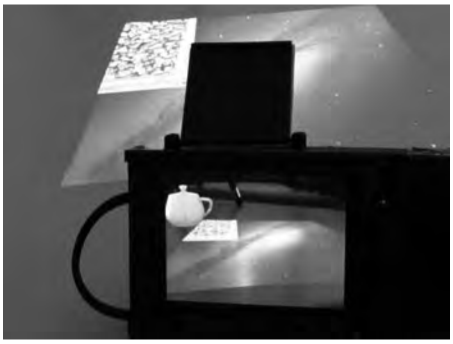


Figure 7.14: Test using Qualcomm's Vuforia to recognize image markers (up to 1/50 of the projection's size) on the projection of another surface phone.

7.5.3 Technical evaluation

For the technical evaluation we recruited 15 untrained right-handed participants (5 female) of an average age of 27 years (ranging from 22 to 65 years).

7.5.3.1 *Evaluation procedure*

In the first part of the study, we assess multi-modal finger touch. The projection shows circles the user has to touch. There are three different sets of circle radius (30, 50, 70 px respectively 1.36, 2.2, 3.16 cm physical diameter) that each are arranged in a grid of 3× 4 circles (Figure 7.16 shows the grid and target sizes in relation) that span the projection area. Additionally, each target exists once for light and once for strong touch, marked with a big L or S (see Figure 7.13a). Thus, in total there are 3× 12× 2 = 72 different targets split across three circle groups. After a test round, each participant performs 2 successive study rounds (resulting in 2160 touches overall). The order of circle size sets is counterbalanced and the display order of targets randomized. Participants can take as much time as they need to perform the touch as the focus of the study lies on gathering a best case estimate of touch accuracy.

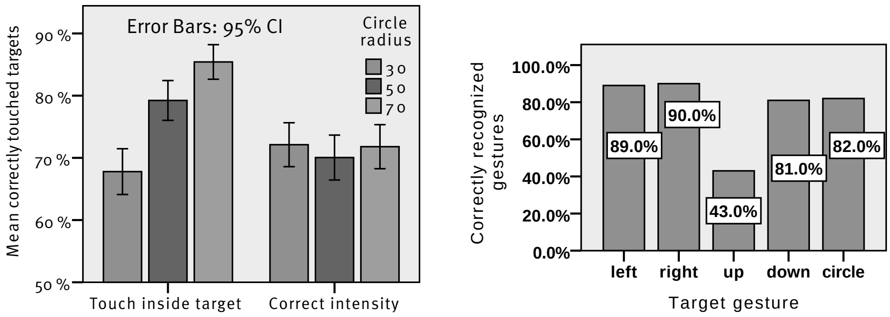
In the second part of the study participants performed the gesture (circle, swipe left, right, top, down) that was written on the projection (see Figure 7.13b). Again, users performed one test and two study rounds of 4× 5 gestures in random order (resulting in 500 recorded gestures overall).

7.5.3.2 *Results and Discussion*

Overall, 93% of touches were recognized (7% have been performed too light to be recognized), 71% of these were hit with the right intensity and 77% of targets at the right position (see Figure 7.15a). Furthermore, we measured that clearly misclassified fingertips ($\delta > 300px$ off the target center) have been responsible for about 12% of false position recognition.

Factorial repeated-measures ANOVA on the touch data reported significant main effects of *circle radius*, *target position in Y direction*, but not *target touch intensity* at the $p < .05$ level (Greenhouse-Geisser corrected) on positional accuracy. Post-hoc analysis using Bonferroni corrected pairwise comparison of means revealed significant differences ($p < .05$) between small and middle and small and large sized circles as well as target heights 140 px / 340 px and 240 px / 340 px. Thus, the larger targets have been and the further they have been away from the device, the better they have been hit in terms of position. Touch intensity recognition is statistically independent of both target position and circle radius.

Left and right gestures yielded recognition rates around 90%, down and circle gestures around 80%. Only the up gesture performed significantly worse than all others with only 43% (Figure 7.15b). The reason for this is that the tracker confused 44% of up gestures with the circle gesture. This may be due to the fact that after performing the correct up



(a) Touch performance. **(b)** Gesture performance.

Figure 7.15: Results of the second user study.

gesture participants moved their hand down and to the right to their default position, which the recognizer that evaluates the gesture after the hand has left the frame may have misinterpreted.

We also wanted to know if users would show a similar over- and undershooting behavior as they do on touchscreens (cf. [110]). Figure 7.16 depicts all performed touches except for the far outliers ($\delta > 300px$). From the touch distribution and their mean marked by the crossing of the fit lines (based on least mean squares deviation) we can conclude that the target position indeed has an effect on over-/undershooting in both X and Y direction. Targets to the lower-right are more overshot than targets to the upper left. However, overshooting is only compensated for more distant targets without transforming into obvious undershooting as on touchscreens [110]. We assume the reason for this is that due to the steep viewing angle of the user on the projection the fat-finger problem only exists close to the device and decreases with increasing distance from the device. Also, perspective misjudgment may counterbalance overshooting. The issue of overshooting may thus also explain the significant effect of Y direction on touch accuracy mentioned before.

Regarding personal experiences, all participants thought that the device is already usable in many scenarios but maybe not for tasks like text entry (3 participants). Similarly, 10 participants stated that the difference between light and strong touches was difficult to learn, maintain (especially after performing the same intensity multiple times before switching), or to perform. For three female users the threshold for strong touch was set too high, for two male users rather too low. Overall, light touches have been slightly but significantly better ($p < .05$) recognized than strong touches (75% vs. 68%).

Overall, the results of the exploration and study of the second prototype reveal that a working SurfacePhone featuring touch, drag&drop (with the help of strong touches), and gesture interaction as well as merging of projections can be built with today's mobile phone hardware. At the same time, there is room for the improvement of the sys-

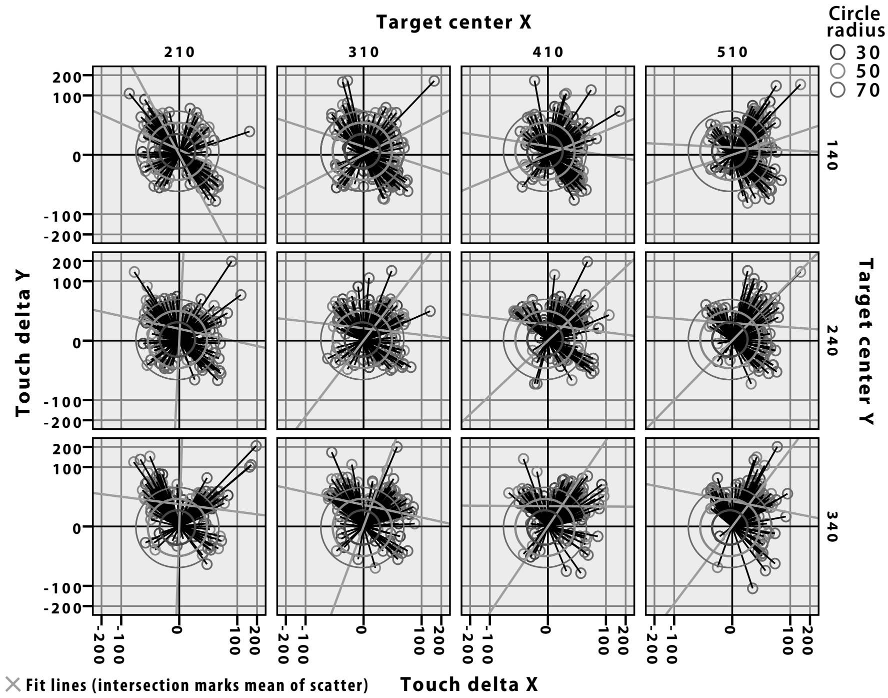


Figure 7.16: Participants' touches visualized by target position and target radius. The orange fit lines (calculated by least means squares regression) and their intersection indicate general under-/overshooting.

tem's accuracy. The following guidelines that are either specific to the SurfacePhone or related to Nomadic Projection Within Reach in general have the potential to improve the accuracy of the system in many ways.

7.6 LESSONS LEARNED AND GUIDELINES

From the explorations and studies conducted with the SurfacePhone concrete guidelines can be derived for the design and usage of devices following the SurfacePhone concept. Further on we gained new insights about Nomadic Projection Within Reach in general. Guidelines specific to the SurfacePhone include the following:

1. As no haptic *system* feedback is available, users should receive a visual feedback about their recognized touch intensity (e.g., a color meter around their touch) to support their mental model of touch intensity. Further, touch thresholds should be personally adaptable to account for anatomic differences.
2. Specific to our implementation (size and angle of the projection and tracking algorithm), interactive elements should have a ra-

dius of at least the size of a large fingertip (radius of 50 pixels or 11 mm respectively) to ensure an accuracy above 80%.

3. Multi-modal touch decreases the accuracy of touch recognition as sometimes users touch too light while they try to keep the intensity below the threshold for strong touches. Thus multi-modal touch should be disabled whenever the interface gets by with single touch plus gestures to increase the accuracy of touch detection to at least 93%.

4. The proposed automatic calibration of touch thresholds is only meaningful 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 (e.g. stone) should be avoided.

Guidelines that extend to Nomadic Projection Within Reach in general, which all happen to be related to R2, are the following:

> *R2 page 8*

1. We have seen that, in theory, new bi-manual interaction techniques like *Human Link* (Subsection 7.3.3.1) are enabled because both displays are simultaneously within reach and the hands of the user are free because the device is put down for interaction. In practice, however, the user study revealed that the bi-manual interaction leads to accidental device movement, because the user has no hand available to keep the phone in place. Thus interaction techniques should be used that do not require simultaneous interaction on both displays or a very robust standing of the device must be ensured.

2. Different to out-of-reach interaction, the device and the projection share a common surface which enables new types of around-device-interaction techniques. Subsection 7.3.3 already assessed many future opportunities (crossing edge-interaction and a new way of spotlight interaction) and with the merging of projections (and withdrawing to maintain privacy) we have seen two implemented example interactions.

3. Regarding transfer techniques (cf. Subsection 7.3.3.1) between the displays, the *proxy* was seen as an advantage in the SDMU case where the screens are divided between the users and nobody wanted to intervene on the display of the other. Nonetheless, in multi-device scenarios with merged and thus larger displays, transfer techniques that supported precise placement (like *touch swipe* and *Human Link*) were favored. Hence, interaction techniques for Nomadic Projection Within Reach should consider the "intimacy" of users with the projected display which, with projections, seems to depend more on size and position (the spatial re-

lation to the user) than on the matter of property like with device screens.

4. Interacting with touchable displays usually involves under- and overshooting because humans have a dynamic (inaccurate) mental model of their finger contact area which depends on many factors like approach angle, approach distance, view on the finger, etc. (see [117] for "understanding touch"). Henze et al. [110] had shown that a correction function which takes shooting behavior into account can significantly improve touch accuracy and mobile operating systems make already use of that. For instance, we can apply a simple offset function to the resulting touch points of the second SurfacePhone study: by moving all touches per target by the vector between the best fit of these touches (interaction of the orange fit lines in Figure 7.16) to the center of the target, we already achieve an improvement of 3.7% over all target sizes, and 7.5% for the smallest targets. As most systems supporting touch interaction on projections do this by applying computer vision, they are usually able to infer the direction of interaction by looking at the intersection of fingers, hands, or arms with the camera frame. Because of that, they will also be able to adapt the offset function to the direction of the interacting user or even multiple ones at once.
5. Following on the previous point, if the projection is oriented towards the public, as it is in the case of the SurfacePhone, interaction may happen from different sides. As the pointing analysis depicted by Figure 7.16 has revealed, users pointed very precisely when they had a good view below their finger (top left target) and scored significantly worse when they were forced to a top-down view inducing the fat-finger problem (bottom right target). Application designers should thus think about the directions from which the interaction will most likely occur, force users through the orientation of typography and other material to take certain positions, or make targets big enough that they can be accurately hit despite occurring under- or overshooting (in our scenario increase the radius by at least 50 px).

7.7 CONCLUSION

This chapter presented the SurfacePhone, a novel configuration of a physical display and a projector that are aligned to allow ad-hoc tabletop interaction on almost any horizontal surface found in nomadic environments. We explored its design space and identified new single- as well as multi-user application scenarios with tailored interaction techniques.

The evaluation of the concept prototype (Section 7.4) indicates that the device is suited for a variety of nomadic everyday scenarios, ranging from personal and collaborative information management to personal and collaborative gaming. The results of the first study further revealed that users are very aware of the *collaboration* and *privacy* deficiencies of current nomadic computing where "only one user can interact at a time" (P1), "the device has to be handed off" (P7), "the device is too small for multiple people to see it" (P16) and no private display exists. All participants acknowledged that these deficiencies are well mitigated by the SurfacePhone. Regarding the mobile deficiencies, the first study revealed:

OUTPUT/INPUT SIZE The projection that is more than 4 times the size of the phone display allows content like websites to be viewed at a glance. Furthermore, regarding R3, we have seen that large shared displays can be created by merging projections in a unique AR experience that mitigates typical drawbacks of stitching approaches [114, 136, 161]. These include (1) usually different screen sizes, which complicate their alignment, (2) the fact that even in the stitched display each device remains the property of only one user, which may influence the manner of interaction of non-owners, and (3) that no private display remains available unless each party brings two devices to the table. All of these drawbacks are solved by projections, which do not inhibit the same notion of property.

> *R3 page 8*

MULTI-TASKING Different to the Penbook, the displays of the SurfacePhone have very dissimilar sizes. This can make them less suited for multi-tasking between different applications. However, they seem very suitable for multi-tasking *within* the same application like showing the main content on the projection and an overview, e.g. of other open documents or websites, on the phone. Several transfer techniques between the displays and user's preference towards binning techniques that support precise placement (such as the *Swipe* technique did) have been shown.

The notion of primary and secondary display can also be switched to aid awareness (see upcoming paragraph about awareness).

COLLABORATION & PRIVACY The SDMU, MDMU setups and their configurations, i.e. the purpose of the private/small and public/large displays in the SDSU and SDMU setups and merging of long and short sides in the MDMU setup, have shown the breadth of new opportunities for collaboration enabled by Nomadic Projection Within Reach in form of the SurfacePhone.

Regarding R4, all privacy concerns—besides some minor suggested

> *R4 page 8*

improvements—of users were found to be mitigated by the SurfacePhone's capabilities. Whilst the main contributor to privacy naturally seems to be the MMDE configuration, smaller features have been mentioned by study participants as well. One such is the rich support for interaction on the public display that renders many interactions on the private phone screen unnecessary. Another is the support for arbitrary sitting configurations that allow very variable levels of trust and intimacy, exceeding the fixed set of setups provided by solutions based on multiple screens such as the Codex [112]. Lastly, this enables device movement as a physical handle for maintaining privacy while entering or leaving sharing that is more quicker and more intuitive than maintaining software privacy settings.

ENVIRONMENT As it was before nomadic computing, tables serve as perfect places for sharing information. Current nomadic devices do not leverage their space and properties at all. Through the merging of multiple projections on tables, large existing spaces for collaboration can be leveraged and included into the interaction between multiple users. In single-user scenarios, the purpose of the displays can also be switched to use the space behind the phone as ambient display in the user's periphery (e.g. to subtly alert to new notifications) while focusing on a task on the phone display.

This chapter later presented a fully functional prototype that demonstrated how the SurfacePhone can be built with only today's commodity phone hardware and the help of a specialized case and customized algorithms based on state-of-the art techniques for finger tracking and multi-modal touch recognition (Section 7.5). Results of a quantitative user study on touch and gesture tracking accuracy revealed that the present prototype would already be applicable to many single- and multi-user scenarios and how it could be further improved, for instance, by counterfeiting typical overshooting behavior and personal adaptation of touch intensity. To spark further research on Nomadic Projection Within Reach-devices with MMDEs, the components of the technical prototype, i.e. the SurfacePhone software, STL print files of the hardware, and assembly instructions, have been made available for download at `http://uulm.de?SurfacePhone`.

The case study on the SurfacePhone focused on colocated collaboration. However, one of the most important qualities of mobile devices in nomadic environments is their ability to connect *remotely* located people. The next case study will investigate possible collaboration and privacy support of Nomadic Projection Within Reach exactly for the case of remotely connected people, namely while being in a phone call with each other.

8

CASE STUDY ON NOMADIC COLLABORATION DURING PHONE CALLS

Deficiencies ***addressed*** *by this chapter*

Output/input size *(D1)*

Multi-tasking *(D2)*

Collaboration & Privacy *(D3)*

Environment *(D4)*

Related video

While the previous case study investigated *colocated* collaboration support through Nomadic Projection Within Reach, many nomadic scenarios involve *remote* communication (phone calls, messaging) and should support remote *collaboration* all the same. Mobile phones currently only offer some support for asynchronous collaboration through sharing and editing files or online documents, but no support for *synchronous* remote collaboration. The SurfacePhone could come to aid by projecting a space shared between the remote users that on the sides of both parties allows to copy and share content between the shared space and the own device screen. However, this would require the phone to be operated in loudspeaker mode or through a headset, which in nomadic scenarios is oftentimes not socially acceptable or available. This constraint is shared with current mobile phones, which held at the ear during a call do not provide any access to their data.

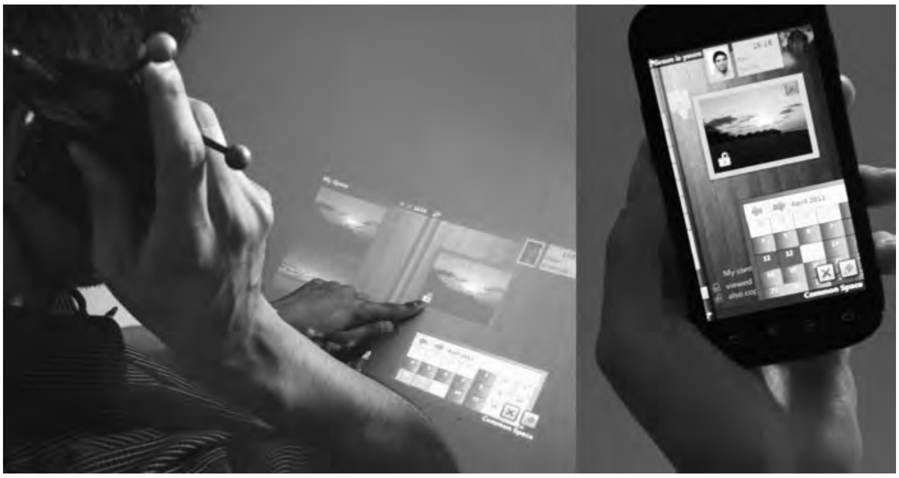


Figure 8.1: The idea of the interactive phone call (IPC). A projector at the bottom of the phone projects a touch-enabled desktop-like interface while the phone is held at the ear during a call (*IPC Projection* mode on the left). When no surface for projection is available, the interface can be used on the phone with loudspeaker mode enabled, more limited by display space though (*IPC Screen* mode on the right)

This case study therefore investigates how by applying Nomadic Projection Within Reach and with the phone held at the user's ear, unhindered phone access and collaboration with the calling party can be enabled during a call (>D2). Starting from the concepts of the SurfacePhone, the hardware design and the interaction techniques for sharing and copying content as well as to maintain privacy have to be adapted (>D3). An important issue in such remote collaboration further is to create awareness (>D4) for each other's actions which does not come naturally as it does in colocated collaboration. The following describes the design considerations, the implementation of the Interactive Phone Call (IPC) prototype and an evaluation study with 14 participants. The study reveals the very positive impact of the large projected display (>D1) on awareness and privacy deficiencies that are apparent without the projected display.

This chapter is based on the previously published refereed conference paper

[W11] **Winkler, C.**, Reinartz, C., Nowacka, D., Rukzio, E., "Interactive phone call: synchronous remote collaboration and projected interactive surfaces." In: *Proceedings of the 2011 ACM international conference on Interactive tabletops and surfaces*. ITS '11. New York, NY, USA: ACM, 2011, pp. 61–70

and extends this by relating it to the SurfacePhone of the previous chapter, to Nomadic Projection Within Reach more generally and providing more details in many sections.
In addition, the following related thesis was supervised by the author:

- "Interactive Phone Call—Synchroner Datenaustauch während eines Telefongesprächs mit der Hilfe von Projektionshandys". Christian Reinartz. Master's thesis. 2011

8.1 INTRODUCTION

Mobile phones are nowadays used as pervasive interaction devices supporting a large variety of communication means, services and applications. Surprisingly, the original function of mobile phones, voice communication, did not benefit from the services and features added to those devices in the last decade. We make frequent phone calls but while doing so it is difficult to use other applications available on mobile phones or to collaborate with the other party. As Gunaratne et al. [98] pointed out, there are many situations, when synchronous collaboration is desired during the synchronous voice conversation, such as sharing pictures or directions and scheduling appointments. Currently, the only available means for remote collaboration are the usage of asynchronous file exchange protocols (e.g., sending of pictures via WhatsApp or e-mail) or through the usage of central servers (e.g., using Facebook for sharing pictures). Moreover, it has been shown that people feel more comfortable sharing private information during a phone call than when sharing to the public, e.g. on Facebook, because of the innate intimacy and limited time span (or lifetime) of a phone call [98].

In this chapter we are going to explore the design space of synchronous remote collaboration between two parties during a phone call. Different from previous research on that topic [98], the user is not required to use additional hardware like a computer or headset or to enable loudspeaker mode which are oftentimes not available or socially not acceptable. Instead of placing the device on a surface (like in previous case studies and chapters), the phone is held at the ear as usual and a projector at the bottom projects a touch- and gesture enabled display on a nearby surface within reach as on the left side of Figure 8.1 (this *can* be a table but all the same any horizontal or vertical surface of sufficient size and reflectance). Three differences to the SurfacePhone phone follow from this configuration: firstly, only one display remains available for interaction (although with a diagonal of 21-25" it is still way bigger (≈25 times) than typical phone screens). Thus the separation of shared and private display must be artificially created on a single display. Secondly, the projection device cannot be moved explicitly but is coupled to the user's implicit head movements which presumably are rather disadvantageous to the interaction and must be compensated for in software. Lastly, only one hand remains free for interaction on the projection.

To support two or more calling parties in synchronous ad-hoc collaboration and data exchange during a phone call, the IPC must support some of the application types that the SurfacePhone supported, such as sharing pictures and websites. Moreover, it also has to support use cases more oriented towards remotely located parties like sharing locations, directions, presentations, and scheduling appointments. For reasons that will be explained in the design considerations, IPC uses a desktop metaphor split up into two adjacent spaces (the inclined reader might want to take a look at Figure 8.2), which can be resized in favor of one or the other, in order to leverage the larger space to provide a more pleasant collaboration experience. The left side of the projection is used for a private view displaying personal data and applications available on the mobile phone. The right side is used for displaying a shared space, which is synchronized in real-time between both users. This allows instant sharing and interactive discussion of files and applications by moving files and applications from the private to the shared space. As all interaction happens synchronously, any annoying meta-conversations regarding the state of sharing become unnecessary.

Apart from the (*IPC Projection* mode, Figure 8.1 left), the system also supports a screen mode without projection that can be used on any conventional smartphone (*IPC Screen* mode, Figure 8.1 right). This seems reasonable to support for the moment when the user is forced to move and the projection surface becomes unavailable during a sharing session that should not be forced to end as long as the call continues. Due to the smaller available display size in *IPC Screen* mode the user only sees the private *or* the shared space at a time, though, and moves between spaces with a horizontal swipe gesture. Besides, the underlying approach of the private and shared space of IPC is extended, among others, through concepts like *color based ownership coding*, showing which object belongs to which user, and *copy permission control* defining whether the other person can only see or can also copy and manipulate data from shared space.

After discussing specific related work, the IPC concept and underlying design considerations will be presented in more detail as well as their implementation in a prototype. Later on, the two different IPC modes (*IPC Projection*, *IPC Screen*) will be contrasted with another mobile in-call collaboration tool (*Screen Sharing*) that supports mobile screen sharing between calling parties with native phone software but without support for projection. Through a comparative study between *IPC Projection*, *IPC Screen*, and *Screen Sharing*, the configurations with/without projection and with/without IPC concept will be compared and insights on Nomadic Projection Within Reach for remote collaboration derived.

8.2 SPECIFIC RELATED WORK

Being an Nomadic Projection Within Reach concept, the IPC applies direct touch interaction and therefore relates to works already described in Subsection 2.5.3. Apart from that, IPC relates to works on synchronous remote collaboration, collaboration via mobile devices, and privacy management during sharing.

8.2.1 Synchronous remote collaboration

Synchronous remote collaboration that goes beyond phone calls has already been investigated in the 1970s' by Chapanis et al. and it has been shown that visual collaboration improves task completion measurably [70]. Since then we have seen a very large body of research and commercial products in the area that often involves the usage of an audio / video link and live sharing of applications, documents and the desktop. Nowadays a multitude of applications such as Microsoft Lync, Windows Live Messenger, Adobe Connect, or Skype (with screen sharing) are commonly used.

8.2.2 Collaboration via mobile devices

Most research concerning collaboration with mobile devices focuses on colocated collaboration. Here several users interact directly with each other, e.g., via a short-range network connecting their mobile devices to exchange files [19], by using a public display as a mediator for the collaboration [178], or by sharing their location information with others [40]. Of course, the SurfacePhone has also been a representative of this category.

There exists relatively little research on synchronous remote collaboration between two users calling each other with their mobile phones, though. The PlayByPlay system supports collaborative browsing between two users whereby one is using a mobile device (e.g. calling and asking for directions) and the other one is using a desktop PC [273], which both are synchronized. The Newport system also supports collaboration between calling parties of which at least one person is close to a computer [98]. The users are able to send each other maps, photos, or notes that can be annotated but no live screen sharing is supported unless both persons sit in front of a computer. The commercial system Thrutu[1] enabled in-call collaboration between smartphones but was designed for being used with loudspeaker mode on the mobile display. In contrast, both Newport and IPC support collaboration during a phone call on a large shared display, but only IPC supports completely nomadic usage as no desktop computer is required.

8.2.3 Privacy while Sharing

Mobile phones are considered as very private devices as they often contain information about personal communication (e.g. phone calls, SMS or email) and store private media. This has e.g. been addressed by the work of Garriss et al., which showed the importance of user privacy during mobile interactions with public kiosks [91]. MobShare is a photo sharing system for mobile devices, which considers privacy aspects carefully as it allows users to define explicitly with whom which pictures should be shared [222]. Ahern et al. [19] confirm that people are in particular concerned about the pictures stored on their mobile phones when considering personal security and social disclosure. IPC addresses this aspect via the private and public space. If the user wants to share a file then the user has to move it explicitly from the private into the public space.

8.3 Interactive Phone Call (IPC)

The IPC concept enables users to browse, share, and copy personal data and collaborate in real-time during phone calls. We added a syn-

1 `https://play.google.com/store/apps/details?id=com.thrutu.client` (Thrutu website (http://thrutu.com) has become unavailable)

chronous collaboration channel to the voice communication to resemble colocated collaboration as closely as possible. In this context, the three most important qualities of colocated collaboration seem to be that collaboration happens *synchronously*, i.e. actions of one person have an immediate effect on the perception of other nearby persons, *bidirectionally*, i.e. all persons can manipulate the same objects at (almost) the same time, and happens in the *periphery*, i.e. even actions that are outside the area of interest of the user can be subtly perceived through *peripheral awareness*.

When the research was conducted in 2011, no solely mobile bidirectional collaboration system existed. Therefore existing systems that support remote collaboration and data sharing with the help of desktop computers, such as Newport [98], Skype, Windows Meeting Space, Windows Communicator, Adobe Connect, and Cisco or VNC products were investigated as a starting point. These systems were found to build upon quite similar user interface concepts for sharing. The most prominent of these is that users have to choose between sharing their desktop / application and watching another user's desktop / application, sometimes with the option to take input control of the remote desktop. Another commonly found concept is sharing files asynchronously, peer to peer by means of Drag & Drop of iconic file representations. We were surprised how few existing systems supported *synchronous* and *bidirectional* communication at the same time. Yet, surely enough the synchronous nature of calls demands for a fitting synchronous sharing experience. Otherwise the phone conversation would likely be cluttered with phrases like "Have you already sent the file? I didn't receive it.", "Have you already opened my file and seen ...", or "I have the file here, come and watch my screen". While this may be tolerable in traditional remote collaboration, mobile phone calls are often likely to happen on the go and last a relatively short time. Therefore, communication and collaboration has to happen as efficiently and effectively as possible.

Apart from how sharing is supported, in-call collaboration on mobile devices entails some more challenges specific to mobile phones. Among those are the stored very personal data that is likely to raise privacy concerns during collaboration; and that mobile phones do not have a desktop interface like PCs but follow the single-application-focus design pattern. The IPC should as well support user's mobility during phone calls, like when being at home, on the go, being able to project or being able to activate the loudspeaker, respectively.

The following will present the IPC and its concepts, each starting with underlying design considerations and how they got reflected in the implementation. The last subsection will then present some apparent use cases supported by the specific implementation of the IPC.

8.3.1 IPC Concepts

8.3.1.1 *Surface vs. Phone Metaphor*

Computer and phone user interfaces developed quite differently in the past according to their diverse usage requirements and the available screen space. While the WIMP (Windows, Icons, Menu, Pointer) metaphor, including Drag & Drop, became very widespread on computer operating systems, modern mobile OS at least abandoned the Windows, Pointer, and sometimes as well the Drag & Drop concepts. This holds true for projected interfaces of available projector phones at the time the research was conducted, which just mirrored the mobile interface and which is still widespread today. As the mobile phone offers more and more desktop PC functionalities, concepts for data, file, and object manipulation regain importance. An example was the Webtop framework on the Motorola ATRIX™ phone that resembled a standard WIMP desktop interface when connected to a bigger HDMI display [273].

Since users are more familiar with concepts for data sharing stemming from traditional computer operating systems, it was decided to build on these by using a desktop-like interface that shows a status bar, application icons, and title-less windows for every opened object or application (see Figure 8.2). Every window can be moved, scaled, or rotated with Drag & Drop, Pinch to Zoom or two-finger rotate gestures. Content inside windows is mostly manipulated with single finger touches. Thus the interface builds on modified WIMP concepts as they are also used in current applications multi-touch tabletop computers. Different from existing tabletop applications, the IPC further allows the user to interact with the surface by just hovering over it. This is for example used to display a close button on windows or hints on certain elements only when the user's finger is close to it (see Figure 8.2 right middle).

8.3.1.2 *Share and Copy between Private and Shared Space*

The desktop space is further divided into a private and a shared space, which can be resized in favor of one or the other space with the divider in the middle (Figure 8.3 left) to account for changing space requirements. The shared space is seen by all other call participants and synced in real-time, including window movement and content manipulation. Each participant can share windows by means of Drag & Drop from the private to the shared space, and copy windows in the opposite direction, if permitted by the window's owner (Figure 8.3 right). The original window returns to its former place after it was shared or copied and an identical copy is created at the place where it was dragged.

Figure 8.2: IPC GUI: Ken and Lisa share pictures, maps, and appointments synchronously during the IPC.

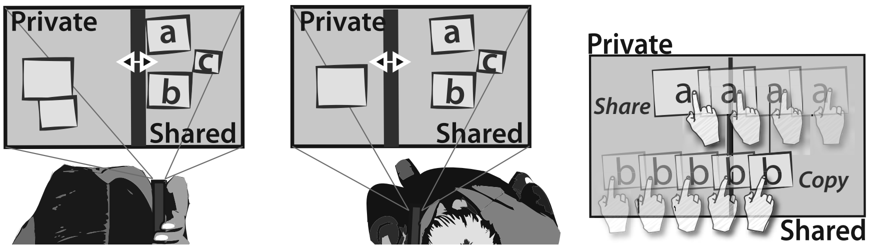


Figure 8.3: Resizable private & mirrored shared space. Sharing and copying is performed via drag & drop.

The shared space on the right, mirrored between all call participants, makes it obvious, which windows are currently shared and how, even at which size-level, all participants currently view them. If permitted by the owner, content can be manipulated collaboratively in shared space and copied to the private space at any time at everybody's discretion.

The real-time synchronization further enables people to add visual communication like gestures to their spoken words as they would do with physical objects when colocated. If someone talks about a certain picture for example, they might point to it, grab and move it, or even point to certain positions in the image. Consequently, the concept of private and mirrored shared space avoids all asynchronous interaction during the synchronous phone call.

8.3.1.3 *Ownership color coding*

Giving the user feedback about which of the shown information belongs to them, the other calling party, or both is critical in sharing private data.

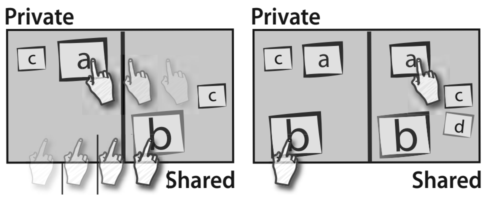


Figure 8.4: Ownership color coding

Different window border colors are used to show the origin and ownership of windows/objects. The own color is always the same, i.e. dark blue, whereas calling parties have different colors. Each window border can thus carry a number of different colors up to the total of call participants. When windows are dragged from private to shared space they receive a border in the owner's color (Figure 8.4). When windows are copied from shared to private space, the border color of the person copying the window is added to the shared object, giving feedback to the former object owner(s) that the item has been copied (Figure 8.4 right).

8.3.1.4 *Supporting Copy Rights and Privacy*

Mobile phones store a lot of very personal data, e.g., e-mails, SMS, pictures, contacts, appointments. Making this data available for sharing is likely to raise privacy concerns. Thus giving users control over copy permissions and the granularity of sharing is important.

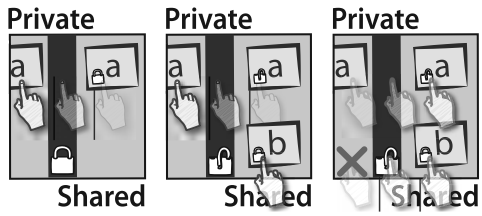


Figure 8.5: Default and individual copy permissions

For IPC we differentiate mainly between three different types of sharing: *not shared*, *view-only* and *manipulate & copy*. The concept of private and shared space already serves the not shared type, as nothing is shared that is not explicitly dragged into the shared space by the user. The user can further toggle between *view-only* and *manipulate & copy* by opening or closing a lock button at the window's corner. The lock button is added to any window that may present private information when the window is dragged from private to shared space across the divider, inheriting the state of the default lock button that sits on the divider as initial state (Figure 8.5). As the names suggest, the window can only be moved and changed in size and rotation as long as the lock is closed and freely manipulated, including content, and copied when the lock is open.

Unfortunately, the *spaces* and the *lock* concepts cannot serve every situation adequately. A calendar for example has to show many individual appointments in one window at the same time, which means they can only be shared all together or not at all. Presumably, these cases are best to be solved individually from one content type to another. A calendar window, for instance, can have an additional details button that can be toggled in private space to influence shared space. When the button is off, only free/busy is shown in shared space without any further details about the appointments (see Figure 8.2 left again). In fact, a *details* button is able to solve a lot of privacy issues, but probably not all of them.

8.3.1.5 *Supporting Freehand Annotations (The Pencil)*

Despite support for real-time mirroring of the shared space, some information can be much better visualized with the help of freehand-sketched annotations. Therefore, all call participants can activate a pencil to draw on top of the whole shared space in their respective color (bottom right in Figure 8.2). This way content such as maps and pictures can be annotated or freehand drawings can be painted on the shared space.

8.3.1.6 *Phone and Call Status*

The IPC software supports features similar to the standard call application. It shows the phone's status bar (battery, signal strength, etc.) at the top of the private space and all call participants by gender-aware icons, names and contact pictures. Starting calls from the contact list within the application is supported as well as ending calls by clicking on the red receiver icon on top of the other calling party, which is revealed as soon as the finger hovers over it (Figure 8.6).

Figure 8.6: The call can be ended by first hovering and then touching the call status widget

8.3.1.7 *Stateful space*

Mobile Phone calls are inherently fragile. The connection can drop when one party moves, loses network coverage, or a phone runs out of battery. Social circumstances can disallow continuing the conversation. During a conversation supported by IPC, lots of data and annotations can be created in the shared space that must not be lost by accident. Therefore, if a dropped session is reopened, the participants are asked whether they want to continue their old sharing session or start a new one. Without another connection however, the shared space cannot be accessed not to undermine the privacy rules of other call participants, who intended showing their content only for the time of the call. In consequence, content users want to store persistently has to be copied to the private space before the end of the call.

8.3.2 Switching IPC Modes and States

An IPC can be in one of several states (sharing or not sharing), modes (*IPC Projection*, *IPC Screen*, or *Non-IPC*) and configurations, which result from multiple users collaborating in different modes.

8.3.2.1 *IPC States*

Apart from call support, the IPC software can also be used without an ongoing call to use the larger projected touch surface to interact with personal content. When a call is coming in (Figure 8.7, 1), it is visualized on the projection and can be answered from there (2). Once answered, the iconic representation of the other calling party is placed on the shared space of both users. However, the shared space is grayed out since sharing content may not be the intention of any call participant. To enable content sharing, one calling party must click the large "Start common space" button that covers the shared space (3). When clicked, other parties receive a unique notification (vibration or audio) (4) to indicate that they must take action to either accept or deny the sharing request that is presented on the shared space; this might require switching to *IPC Projection* or *IPC Screen* mode (4a). This initial handshake about sharing prevents unknown callers from presenting inadequate content without prior consent. Once the request is accepted (5) the shared space is established (6) and participants can share content until the end of the call.

8.3.2.2 *IPC modes*

Today, most people still telephone by holding the phone close to their ear. Despite alternative options as loudspeaker mode or the usage of hands-free accessories, the advantage of the original telephone behavior is that it does not require additional hardware like a headset, which may not be available at the moment or has run out of battery. Further it is much more unobtrusive than loudspeaker mode and can be used in relatively noisy environments. We think a projection can serve many circumstances where loudspeaker mode is no considerable alternative. However, there may be situations where this is vice versa. Therefore, IPC supports both *Projection* and *Screen* mode and seamless switching in between. Since the current mode affects the available space for sharing, these state changes have to be accounted for during calls as well (see next section).

In total IPC knows three different modes it can operate in as depicted in Figure 8.8. The user can cycle through modes by pressing the projector hard button on the side of the phone, which current projector phones such as the Samsung Beam offer.

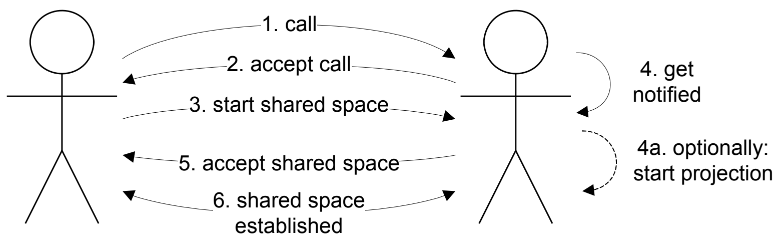


Figure 8.7: IPC states. Handshake before sharing starts.

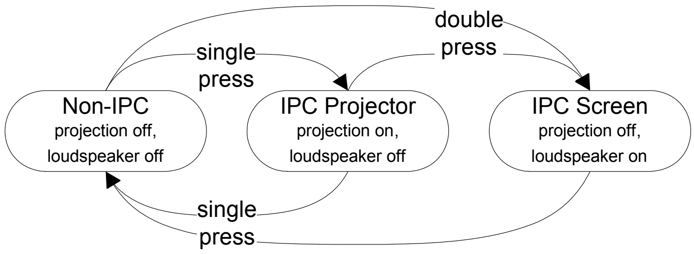


Figure 8.8: IPC modes, changed by different presses of the phone's projection button.

In *Non-IPC* mode, the IPC software runs in background, only, and listens for button presses and sharing requests. The *IPC Projection* mode shows the IPC user interface described earlier on the projected display. The user holds the phone to their ear and interacts with the projection with their free hand. For the *IPC Screen* version, where the phone is held in front of the user, we considered whether we wanted to use the very same desktop UI on the much smaller phone screen or develop an alternative UI that resembles more the UI of standard phone software. Since we did not want to have completely different UIs for the same application not to overstrain users, we only slightly adapted the projected version to the mobile screen. Due to the smaller space on the mobile screen, we decided that private and shared spaces are exclusively visible, only. The user can move between spaces manually by performing a horizontal swipe gesture on the screen or automatically by dragging objects between space borders. Since views are separate, the mobile version does not allow resizing spaces.

8.3.2.3 *IPC size configurations*

As mentioned earlier, our goal is to support seamless switching between *IPC Projection* and *IPC Screen* mode. When a user switches from *Projector* to *Screen* mode, the size of their private and shared spaces shrink considerably. If the other calling party is in projected mode, the shared space sizes that are synced 1:1 are not equal any more. To account for that, the space available on the mobile screen is highlighted on the shared space of the user in *Projector* mode and remaining space is grayed out.

8.3.3 Group sharing and collaboration

During the design phase of IPC care was taken not to limit the concepts to a two-person setup, but find solutions that would scale to conference calls with multiple users. The presented concepts, in particular the shared space with iconic representations of all participating users, color coding, lock, and pencil work equally well for a conference call with multiple participants.

8.3.4 IPC Use Cases

8.3.4.1 *Calendar*

The calendar application allows for dragging the own calendar to the shared space and stacking calendars of different persons on top of each other to merge them and see all appointments together with corresponding color coding for each day and appointment (Figure 8.9). If new appointments are added in shared space, they are added automatically as shared events for all call participants that own the merged calendar, i.e. all persons that dragged their private calendar on top of the shared instance before. The calendar widget shows all information when being displayed in the private area but only information about free and occupied slots is visualized when shown in the public area. This helps to preserve private information. If needed all details can be revealed in the shared space by toggling the "details on" button in one's private space.

Figure 8.9: Merge calendars by stacking them in the shared space to quickly see free spots to meet.

8.3.4.2 *Maps*

Map windows include a Google Maps browser window, a search field and zoom and navigation controls. They can be used to bring a certain location into view and then share it, or to collaboratively explore the map in shared space. Annotations can be used to mark users' current locations, as well as spots to meet, or park the car (Figure 8.10). This widget is in particular beneficial when discussing places to see or visit, when discussing routes or when planning a trip.

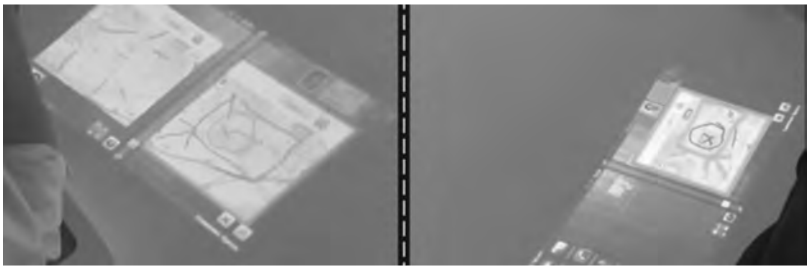


Figure 8.10: The maps application allows sharing of a map selection, that can be further panned and zoomed in the shared space and annotated using the pencil tool.

Figure 8.11: Sharing pictures and videos with the other party by moving them across the border. Multi-touch gestures allow for scaling and rotation.

8.3.4.3 *Pictures and Videos*

Similar to the native phone gallery software, media files can be opened from a list of available gallery albums and their windows can be resized, rotated, or shared. This allows collaborative discussion and sharing of pictures and videos which we envision as one of the central usages of IPC (see Figure 8.11).

8.3.4.4 *Presentations*

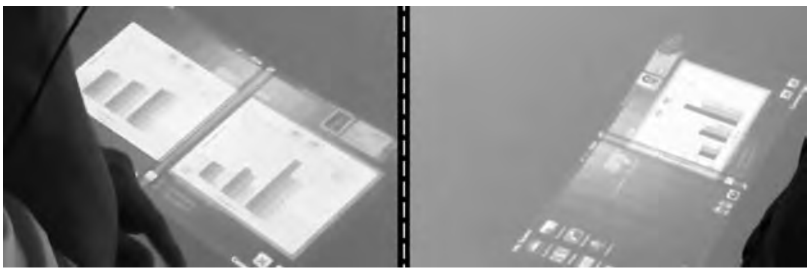


Figure 8.12: Slides can be presented (or a talk given). Controls in the shared space allow (both parties) to switch between slides.

Sharing a presentation can be used to discuss revisions before the actual presentation is conducted. Furthermore, if the intended audience is not colocated, the presentation can be moved to and conducted in the shared space during a conference call. The IPC software supports sharing presentations, moving between slides in the shared space and annotations by means of the pencil (see Figure 8.12).

8.3.4.5 *Live Camera Image*

A live stream from the camera capturing the area in front of the user can be shared if the mobile phone's camera is located on the same side as the projector. Anything in the view of the user that is relevant to the conversation, such as hand gestures, a sight of a beautiful landscape, items of interest while shopping or pictures on traditional paper, can easily be brought into the collaboration session.

8.4 PROTOTYPE

The following presents the system's setup, which was designed to fit the targets of the subsequent user study. A brief summary of these targets is that participants should be able to engage in a real phone call while performing some collaboration tasks on the projected and the screen interface and experience the system as fully functional.

Thus a "projectorphone" prototype was to be developed that projects a rectangular surface in front of the user whilst held at the ear; further a system that tracks the user's fingers and maps it to the projected surface, independent of the surface's position; the IPC software with its underlying concepts and support for some content types (pictures, calendars, maps, presentations); finally the integration of the aforementioned to achieve the *IPC Projection* and *IPC Screen* modes. Moreover, the system setup had to be doubled and both systems connected to each other in order to achieve a realistic call scenario between two persons.

The resulting overall system setup for the IPC is depicted in Figure 8.13. For reasons described later, the IPC software runs on computers that are via LAN connected to each other, that are connected to the system that manages finger tracking, and additionally to respective projector phones via VNC. The audio connection between calling parties is over standard cellular line from one phone to the other.

8.4.1 Projector Phone Prototype

Because no suitable projector phone was on sale, a projector phone prototype was built that consists of a Samsung Nexus S Android phone attached to a Microvision SHOWWX+ laser pico-projector (Figure 8.14).

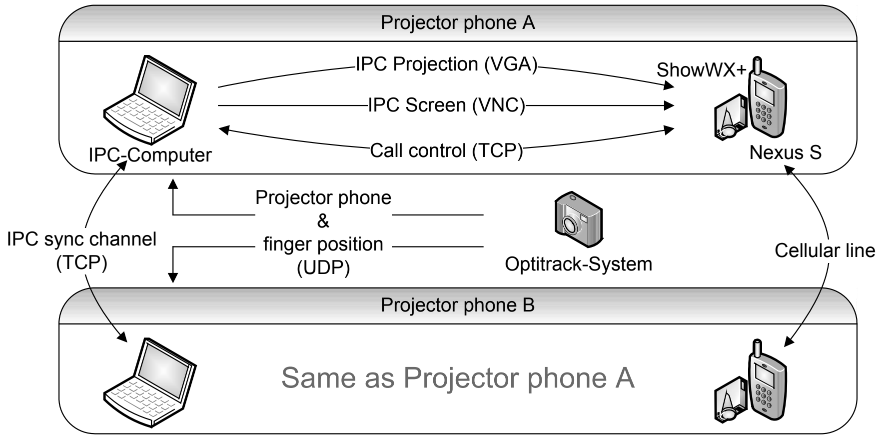


Figure 8.13: System setup. Two synchronized PCs running the IPC software receive input events from the tracking system and deliver their output via VGA to phone projectors and via VNC to phone screens.

In order that the projector phone can be held as usual, i.e. to the ear and parallel to the face, the devices had to be attached orthogonally, to project a landscape image in front of the user. Additionally, retro reflective markers were attached to the projector, in order to track it with the 6DOF infrared-based tracking system OptiTrack from NaturalPoint. The usage of a laser projector further allows the user to change height and angle of the projection without the image losing focus (cf. Subsection 2.3.1). At a typical distance of about 50 cm between ear and center of the projection on the table, the SHOWWX+, thanks to its throw ratio, projects a bright 46 × 26 cm-sized image with a resolution of 848 × 480 pixels illuminating the surface with 125.8 lx. The projection distance is farther away than in previous systems and therefore the illuminated surface less bright, but it is still short and bright enough that even small text (> 11pt) could be easily read in the slightly dimmed room during the user study. Another advantage of the projector-at-ear setup is that any jitter of the projection is almost unnoticeable to the user since head and projector move simultaneously. In theory, jitter might only become a problem when the user tried to touch the surface, but was hardly an issue in our tests and studies.

Figure 8.14: The IPC hardware prototype. A SHOWWX+ Pico Projector is equipped with retro reflective markers and by two aluminum angles and Velcro tape orthogonally and with height offset (for better handling and larger projection) attached to a Nexus S phone.

Obviously, phone and projector are only attached but not directly connected to each other. Nevertheless, the good integration between the components made the separation unnoticeable to study participants. For instance, the phone was bidirectionally connected to the PC such that calls were visualized on the phone and the projection and toggling the on/off button on the side of the Nexus S, cycled through different IPC modes as desired. The connection between the phones ran over real cellular line and could be started and ended from either displays. Furthermore, geometric compensation of the projection was implemented and display-fixed finger tracking in order that users can independently move, rotate, or change the height of the projector during a call to support their comfortableness. In this case, the 3D support of the Windows Presentation Foundation (WPF) has been used to render the projected image from the position of the projector that was made available by the OptiTrack system to project a counter-distorted image (see Subsection 2.4.3 for the mathematical background). However, users were bound to the calibrated frame of the tracking system which does not let them stand up or move away from the table in front of them.

8.4.2 IPC Software

The IPC software (Figure 8.2 on page 138) was built with C# and Microsoft's Windows Presentation Foundation (WPF) and Windows Touch frameworks, because current mobile phone UI frameworks are neither designed nor suited for tabletop interaction and the much larger display space. The projected version is therefore served from a PC via VGA to our projector phone.

The *IPC Screen* version runs on Android OS and builds upon Android VNC Viewer to display and control the same IPC software via VNC, showing one space at a time as described earlier. The VNC connection introduces a small latency, but which is almost unnoticeable from a user standpoint. Furthermore, the Android application runs a service in the background that communicates phone state (calling, ringing, off the hook) to the PC and receives control commands like "start call to x" or "end call with x" from the PC. The PC uses RealVNC for transmitting the IPC output to the mobile phone.

Our chosen setup has the advantage that it presents the user with the same user interface on the projection and the screen, while still retaining most affordances of the phone like hard buttons and the touch screen.

8.4.3 Real World Deployment Considerations

A real deployment of the system poses two major challenges. First, the rotation of the projector against the surface must be known to be able to project a counter distorted image. Second, finger touches must be recognized through some kind of optical tracking system that is on-board the projector phone. If the system further was to support unplanar projection surfaces as well, these must be detected through optical surface estimation as well.

Although there is no standard, ready-to-use method available, recent research and products showed how such a system could be realized. The orientation of the device can be sensed with the help of inertial sensors like accelerometers, gyroscopes, and magnetometers already present on smartphones (cf. Subsection 2.4.3). If the phone featured a camera looking in the same direction of the projection, the camera could capture small visual markers in the corners of the projection to estimate spatial relation and distance between projector and surface without calibration. Of course, a depth camera would further simplify finger tracking. Available approaches have been discussed in Subsection 2.5.3.

In regard to the real-time synchronization, recent advances in mobile data networks (4G and later), and the already available support for mobile video-teleconferencing indicate that the realization of IPC's shared space is feasible today.

8.5 EVALUATION

An initial evaluation was conducted to receive user feedback on the IPC system. This was to analyze and distinguish (a) the effect of projecting the interface during the call instead of using the mobile screen and (b) the extend to which the IPC concepts enhance in-call collaboration compared to standard phone software. For the sake of (a) we treated the *IPC Projection* (C1) and *IPC Screen* (C2) modes as two separate study configurations, which only differed in using the projection or screen. For the sake of (b) we introduced another mobile in-call collaboration system, the *Screen Sharing* system, as third study configuration (C3), which will be described by the next subsection.

8.5.1 Screen Sharing Prototype

Configuration 3 (Screen Sharing) adds support for screen sharing and remote pointing to standard smartphones. The software consists of a

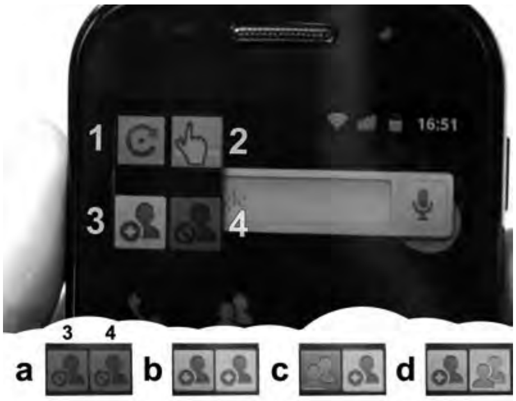


Figure 8.15: The Screen Sharing widget that runs as background service and lies on top of all other applications. It allows to be placed in one of four corners (1), supports non-verbal pointing gestures (2), sharing the own screen (3) and watching the other party's screen (4). More detailed explanations are found in the text.

small widget of four buttons (Figure 8.15) that lie on the topmost layer of the Android view stack and remain therefore always visible. With the lower buttons (3 and 4) the user can control screen sharing. At the same time they serve as indicator for the current state. (a) shows the initial state where the user has not shared their own space (left icon) and neither has the other call participant (right icon). With a long press on (3) the user can allow (or prohibit again) sharing their own screen. (b) shows the state after both participants allowed screen sharing. The user can invite the other participant to start watching their screen with a normal press on button (3). This may result in state (c), which indicates to the user that the other participant is currently watching the own screen. By pressing (4) while the other participant allows sharing, the user starts watching the other participant's screen (d). Pressing (4) in (d) stops watching again. Moreover, (4) can be pressed in (c) or (3) in (d) to directly switch between watching jointly the own or the other's screen. With (1) the whole widget can be moved to the next screen corner to give free sight on the area below. With the pointer hand (2) both users can bidirectionally visualize their screen touches with separate colors to each other while viewing the same screen.

The support for screen sharing and pointing was added to build a system that only uses standard phone software and at the same time allows solving the tasks of the user study. Content can be shared view-only through screen sharing and annotated by means of the pointing hand. But content can only be copied/shared by e-mail, MMS or social networks, as is the current state of the art.

8.5.2 Study Participants and Setup

For the user study 14 students (6 female, age 20-25) were recruited who all were experienced smartphone users. Participants were explained all configurations at the beginning and participants could explore the three configurations on their own.

During the study one experimenter stayed with the participant to give instructions on the tasks to perform and to assist if a participant would get stuck in solving a task. Another experimenter in another room acted like a close friend of the participant and engaged with participants in a real conversation over the phone. We chose this setup to mimic a real scenario while at the same time assuring that all conversations and actions took almost the same course. We employed a within-subjects design to compare *IPC Projection*, *IPC Screen*, and *Screen Sharing*. The order in that participants were exposed to the configurations was counterbalanced; the order of the tasks was always the same. Moreover, both experimenters followed a detailed script to synchronize their interaction and to ensure that all interaction with participants was very similar across all study sessions.

After each configuration participants were asked 12 questions, which could be answered from "strong disagree" to "strong agree" (5-point Likert scale). Further they were asked about perceived advantages and disadvantages of the system. After finishing the third configuration and questionnaire participants were asked to compare and rate the configurations in terms of performance and personal liking and to tell when, where, and why they would use such systems. One study session lasted approximately 80 minutes.

8.5.3 Study Procedure

Participants had to fulfill a series of tasks with each configuration. The first was to call the experimenter and to establish the collaboration with the means provided by the present configuration. Similarly, the last task was to end the call. In between, users had to perform the following four tasks:

1. Open the gallery application, select and share own pictures and copy pictures from the other party. One own picture was considered private and therefore the participant had to ensure that the other calling party was not allowed to copy it.
2. Recommend a place to meet for coffee to the other calling party, which pretended not to know the place. Here the user had to open the Maps application, look up the address by entering a search phrase, optionally pan and zoom, and then share the map centered on the destination. Further it was to be annotated (with the pencil) based on questions asked by the experimenter on the phone.
3. Participants had to open the calendar application, merge their calendar in shared space with the calendar shared from the other party, and find a free slot for an appointment in the merged calendar. The other party added the appointment to the merged cal-

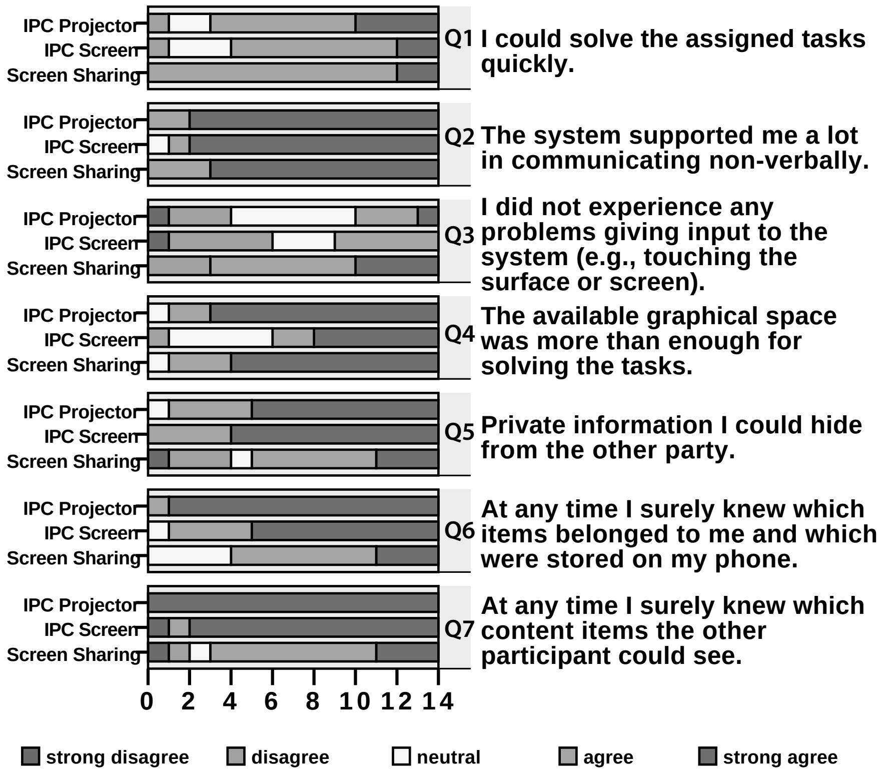


Figure 8.16: Selected questions/answers from the questionnaire participants filled in after each configuration.

endar and instructed the participant to check the appointment in their own calendar in private space.

4. Finally, participants were requested to open and share a lecture presentation they had on their phone in order that the other party could look up a specific slide.

During the *Screen Sharing* configuration, pictures, maps, and presentations were presented through screen sharing. Pictures and presentations were exchanged via e-mail. Calendar entries were viewed separately with turn taking screen watching and the appointment was added by both calling parties separately.

8.6 RESULTS

Overall, participants gave very positive feedback about all three systems. In regard to the question if and where they would use such systems they reported they would use them in almost every location and situation. In the case of *IPC Projection* answers included "I would use it only in private areas, like at home or at work", but more feedback was like "everywhere I am where I have a large projection surface" and even

"I would use it in public transport if space allowed". One participant highlighted the fact that not projection *or* loudspeaker must be used, but that projection and loudspeaker can be used *together* for an easy to set up teleconferencing system between remotely located groups (of which at least one per group has an IPC-enabled phone to create the projection shared by everybody else in the group).

8.6.1 Input and Output

In the final comparison of the configurations, 12 of 14 participants mentioned the bigger space as advantage of the projected interface on their own, and similarly 5 that the phone can still be held at the ear. Further answers from the questionnaire are depicted in Figure 8.16. Unfortunately, the IPC configurations could not provide the very same input experience as the native Android applications did in the *Screen Sharing* configuration (Figure 8.16, Q3). In *IPC Projection* this was due to the tracking system sometimes not recognizing a user's finger accurately for clicking or the click was recognized with a slight offset when the projector was tilted beyond a certain angle. No user complained about the projection jittering. In *IPC Screen* there was a small noticeable lag in interacting with objects due to constraints of the VNC connection. Responses to Q4 further show the desktop metaphor was the right decision for the projected interface but it does not perform so well when used on a mobile screen.

8.6.2 Collaboration and Privacy

Regarding R3, answers to Q1 and Q2 indicate that the three configurations facilitated collaboration and that all tasks could be solved in a reasonable time. With more robust input in IPC configurations, we expect IPC approaches to perform at least as fast as *Screen Sharing*. One problem that participants had with the *IPC Screen* and all the more with the *Screen Sharing* configuration was that because they could only see one space (private or shared) at a time, they sometimes missed an action the other participant performed in shared space. As such they were not as aware of the other party (R5) as in the mode utilizing projection.

> *R3 page 8*

> *R5 page 8*

The biggest difference however, we found in the perceived support of users' privacy R4). Answers to Q5, Q6, Q7 indicate that IPC configurations performed better in supporting the user's privacy. In the *Screen Sharing* configuration, users could only share their entire screen or nothing at all whereas the IPC configurations allowed a much more fine-grained control. Because the given tasks required several switches between the own and the other calling party's screen, participants told

> *R4 page 8*

us they were not always too sure whose content they were currently looking during *Screen Sharing*. Further they felt uneasy with the fact, that they could not be sure when exactly the other person started to watch their screen (again)—at least if they had permitted access in general before. Although the permission could be revoked at any time, users obviously did not feel the same control as with the concept of private and shared space. Moreover, users did not know which content belonged to whom and which they had already shared (Q6). Interestingly enough, the smaller space available in *IPC Screen* mode has had a measurable negative impact on privacy (differences in Q5-Q7) as it did not provide for the same awareness (R5) although beyond the size difference both systems were identical.

> *R5 page 8*

Finally users were asked to rank the three tested configurations depending on which configuration provided the fastest experience in solving the tasks and which configuration they preferred overall. The received answers are interesting since although most users (6) felt the fastest with the Screen Sharing configuration (compared to 4 and 4), 11 participants in contrast favored one of the IPC configurations (6 *IPC Screen* and 5 *IPC Projection*). This can speak for the aforementioned advantages of the IPC concept.

Lastly, the positive user feedback across *all* three configurations shows that three equally sophisticated systems with similar functionalities but opposite qualities have been compared—a prerequisite for any comparative study. The results further indicate that the IPC improves in-call remote collaboration in a variety of use cases, even better than remote collaboration could be implemented on top of standard phone software. The evaluation also revealed that IPC concepts like the desktop interface and the private and shared space reveal their full potential only with the projected interface, which also has the advantages that it does not require activating the loudspeaker and may be used to integrate nearby persons in the call with help of the projection.

8.7 CONCLUSION

This chapter presented the IPC and its concepts that facilitate synchronous collaboration during a phone call. Previous research proposed the use of additional hardware like PCs [98] while this case study explored the possibilities and requirements for a system that solely relies on mobile phones. The presented system supports projector phones as well as conventional phones through the *IPC Projection* and *IPC Screen* modes, which can be seamlessly switched to serve different mobile situations. The evaluation of the IPC, also against another likewise novel system called *Screen Sharing*, indicates that the IPC concepts, e.g., private and shared spaces, color coding, copy permissions, and the projected interface highly improve the user experience in synchronous remote collaboration in terms of visual communication and user control over sharing and privacy.

More importantly, the study has shown that these concepts adequately address current mobile deficiencies, starting from the

OUTPUT/INPUT SIZE deficiency, to which regard IPC allows a single user to explore phone content on a large projected display of the size of a desktop monitor. Presumably, it would be considered awkward to hold a projector in the air to create this display or to hold a projector at one's ear—but for a projector phone this seems perfectly fine and no user mentioned any awkwardness during the evaluation.

COLLABORATION & PRIVACY More importantly, in case of the IPC we have seen that it is not an MMDE that enables collaboration and privacy as has been in previous case studies, but the large size of the projected display that allows for a private and adjacent shared space on a *single* display.

ENVIRONMENT As the other party is not part of the real environment as in colocated sharing, awareness has to be created artificially. Therefore the concept allows either side of the split display to serve as view in the periphery that aids awareness about the own information as well as actions undertaken by the other participating party. The evaluation revealed that this awareness was missing without the projected display.

This chapter concludes the part on Nomadic Projection Within Reach. The case studies presented so far enabled multi-modal single- and multi-user interaction and collaboration in nomadic scenarios but shared a single requirement: a suitable horizontal[2] surface, usually a table with one or multiple chairs. On the other hand, all applied Nomadic Projection Within Reach in a narrow sense and thus shared the advantage of a short projection distance that yielded bright projected displays and devices that could be commercialized today.

In contrast, the next part will take a look at an even higher level of nomadicity. So far we have only looked at the "moving from one place to another" part of nomadicity as it was defined in Chapter 1, excluding the time of "moving". The next part will explicitly concentrate on these on-the-go scenarios and how their deficiencies can be addressed through an adapted or extended form of Nomadic Projection Within Reach.

2 or also vertical in the case of IPC

Part III

NOMADIC PROJECTION WITHIN *EXTENDED* REACH FOR CROSS-DISTANCE INTERACTION ON-THE-GO

9

CASE STUDY ON *CONTINUOUS* NOMADIC INFORMATION MANAGEMENT ON THE FLOOR WHILE ON-THE-GO

INTRODUCTION TO NOMADIC PROJECTION WITHIN EXTENDED REACH

On-the-go scenarios pose quite a challenge for projected interfaces since only very few surfaces—actually, only the floor and the user's hands[1]—are constantly available for projection and interaction. At the same time, projected interfaces entail a huge opportunity to increase the user's awareness (R5, >D4) and decrease the lead time to interaction (R6) in nomadic on-the-go scenarios, because of their unique advantage of being able to create persistent displays in the visual periphery.

> R5
> R6
page 8

Beginning from that, this and the subsequent chapter will explore the design space of applying Nomadic Projection Within Reach (NPWR) to floor and hand surfaces while moving. The NPWR concept needs to be slightly extended, though, since the floor is rather "out of" than "within" reach as demanded by NPWR. It *would* be within reach for foot and toe interaction as has been presented in paragraph Hands and Feet (Subsection 2.5.3), but does not seem suitable since it would preclude simultaneous movement and interaction which is our primary focus. When aiming for interaction with arms, hands, and fingers, we are faced with a cross-distance interaction space ranging from the out-of-reach floor to the within-reach hand. The cross-distance interaction space was already mentioned in Subsection 3.3.3 and the taxonomy on page 44. As we remember, the cross-distance space is not just a definition of a certain distance in the middle of out-of and within-reach distances, but describes an interaction space that reaches across various interaction distances and adapts to those. As the taxonomy depicted, almost no works so far have dealt with varying interaction distances in nomadic projection scenarios. When we take the within-reach distance a little more figuratively, we can extend NPWR to allow out-of-reach objects to cross the distance and be brought into reach. In the following AMP-D case study, for instance, objects on the floor can be pointed at to be brought into reach on the user's hand for NPWR interaction, to eventually be put back out of reach again. Analogously, the extended frame-

1 except for gloves in winter but which could be chosen white. Other body parts do not provide the same angle and distance to a worn projector and are therefore less suitable.

work is called Nomadic Projection Within Extended Reach (NPWER). As Subsection 2.4.1 has explained, though, a disadvantage of current projector technology regarding NPWER is that it cannot support a good contrast for out of reach projections. However, it can be argued that it will (now or in the near future) still be bright enough to create the required awareness in the user's periphery (e.g., through attention grabbing animations) and provide a minimal lead time to interaction as the display is always available. As soon as objects are brought into within-reach distance for active interaction, they provide the required contrast as previous case studies did.

The following two case studies both try to support typical information management tasks that occur on-the-go without the help of other mobile devices. These tasks include notification management, reading and writing messages, navigation, to name but a few. The first system (AMP-D) presented in this chapter is more advanced and provides a persistent pervasive display in front of the user but is also more cumbersome to wear, yet. The system presented in the next chapter (SpiderLight) is much smaller and comfortable to wear, but sacrifices some of the functionalities of the former.

Related video

The rest of this chapter is based on the previously published refereed conference paper:

[W4] **Winkler, C.**, Seifert, J., Dobbelstein, D., Rukzio, E., "Pervasive Information Through Constant Personal Projection: The Ambient Mobile Pervasive Display (AMP-D)." In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14, **Honorable Mention Award**. New York, NY, USA: ACM, 2014, pp. 4117–4126

In addition, the following partially related thesis was supervised by the author:

- "NaviBeam—mobiles Assistenzsystem mit persönlicher Projektion". Markus Broscheit. Bachelor's thesis. 2011

9.1 INTRODUCTION

Deficiencies addressed by this chapter

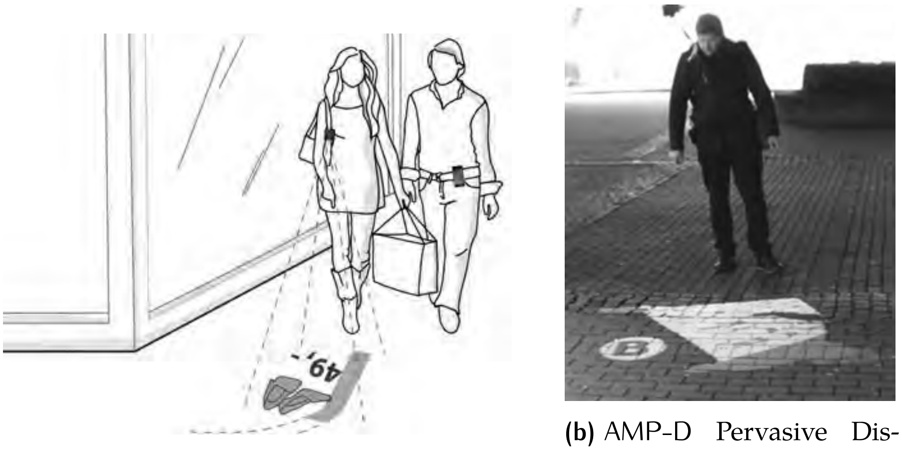
***Output/input size** (D1)*

***Multi-tasking** (D2)*

***Collaboration & Privacy** (D3)*

***Environment** (D4)*

Today, we observe an ever increasing interest of mobile users towards pervasive information access and serendipitously discovering new information. This is currently achieved by means of smartphones and public displays in the environment. Public displays will likely never be widespread enough to fulfill this desire alone. Inversely, smartphones can only contribute to this vision when they are held in hand and are actively operated. This becomes a challenge especially when the user is on the go as the device distracts the user's focus and connection to the environment (>D4). The idea of wearing a location-aware pervasive display that alerts to new relevant information and provides quick



(a) Vision

(b) AMP-D Pervasive Display

Figure 9.1: The vision (a): small wearable projectors reveal serendipitous information and its implementation (b) in the AMP-D prototype.

access to deal with it, is very compelling. It could provide access to a great variety of information, ranging from *public* content (e.g. stationary pictures from Flickr) to *personalized* content (e.g. personalized advertisements in front of shopping windows (Figure 9.1)), and *personal* content (e.g. notifications about new text messages). Previous works on mobile speech interfaces, head-mounted displays, public-display networks, and mobile projectors attended to this vision one way or the other. One crucial requirement for such a wearable pervasive display is that the display is always available. Another is that the display is located in the user's periphery and uses ambient presentation to minimize the risk of annoying other people and distracting them from their primary tasks.

This chapter investigates the use of constant personal projection as a novel display technology for personal pervasive displays that supports the aforementioned use cases (pro and contra of a peripheral projection compared to other display technologies, e.g. HWDs, have already been discussed in Section 2.2). The Ambient Mobile Pervasive Display (AMP-D) integrates a wearable projected display that accompanies users on the floor in front of them with a projected hand display and a smartphone to a continuous interaction space. The floor display is a constant pervasive window into the user's digital world, lying in the user's visual periphery (Figure 9.1b), meant to integrate digital information into the environment to increase the user's awareness (>D4). As such it allows for subtle user alerts without occupying the user's hands or visual field and can be consumed and interacted with during other tasks (>D2) such as walking. The hand display allows to deal with information instantly without having to reach to a physical device. The smartphone supports exploring and sharing content from and to the virtual world. Users interact with AMP-D entirely through hand gestures. For

(a) Floor interaction

(b) Hand interaction

Figure 9.2: The user has received a text message and picks it up from the floor (a) and reads the scrolling message text in the own hand.

instance, when a new text message is received, a message notification box rolls into the user's view on the floor (Figure 9.2a). Users can then pick the box up and read the message instantly in their hand (Figure 9.2b). Optionally, they can answer the message using their smartphone.

In addition to the wearable mobile multi-display environment (MMDE), AMP-D uses a consistent information space for typical public and personal mobile content that augments users' virtual world through Spatial Augmented Reality (SAR), giving users a natural means of discovering new information.

After discussing specific related work, the AMP-D concept of constant personal projection and its interaction techniques will be described. Further on, their implementation in a fully functional prototype and various implemented application examples that explore and highlight the applicability of AMP-D to typical mobile scenarios. The chapter concludes with the lessons learned from a small user evaluation and its implications on Nomadic Projection Within (Extended) Reach.

9.2 SPECIFIC RELATED WORK

The ideas of using the floor or hand for projection are not new as Chapter 3 has presented. However, using constant personal projection on the floor or combining floor and hand display to a continuous interaction space are. Constant projection had so far only been used by Leung et al. [157] who demonstrated a system that constantly projects a representation of the wearer's online social identity to the ceiling above. As its intended audience are spectators it is not designed to be interactive, though, and limited to indoor scenarios by design.

Further related work is in the domain of mobile or head-mounted peripheral displays which try to increase awareness of the mobile user as AMP-D does. Wearable augmented reality displays date back to the works of Feiner et al. [86, 87] who developed head-mounted see-through displays for spatial augmented reality. This display type constantly overlays the user's foveal view making it less suitable for everyday scenarios. Apart from that, head-mounted displays to date entail unsolved challenges such as perceptual issues (e.g. different focus planes and narrow FOV); security issues (e.g. they may distract from or obstruct an approaching danger), and social issues (e.g. collocutors having no means of knowing what I am looking at).

A mobile *peripheral* display is the eye-q system [75] that uses a low-resolution LED display embedded in a pair of eyeglasses to provide a small set of notifications. An advanced display version is provided by Google Glass [93] (and the like), whose display lies slightly to the top of the foveal field of view, also qualifying it for ambient display. Unfortunately, at the same time, its position and small size make it less suitable for complex visual output, augmented reality, or direct interaction.

Works on mobile projected displays so far dealt with within-reach and out-of-reach projections separately whereas AMP-D presents a continuous interaction and information space *across* these display types. Similar continuous interaction spaces have only been presented in static smart-space setups. AMP-D aims to bring this compelling vision to the mobile user in everyday use cases. In these nomadic scenarios, ambient display properties are much more important, which so far have been neglected in works on mobile projections as well.

9.3 CONCEPT AND DESIGN OF AMP-D

The AMP-D is a wearable multi-display system that provides a pervasive window into the user's virtual world on the floor. Unlike smartphones which have to be taken out to be operated, the AMP-D display is constantly available. Therefore it is suited for ambient alerting to many kinds of public or personal information that is available via the user's connected smartphone. Among others, these information types include location-aware notifications, communication, reminders, and navigational instructions. Additionally, information is not only visualized, but can be handled through gestures in the user's hand which is used as on-demand secondary display.

The concept of AMP-D will be illustrated by first discussing each of its basic concepts. Following on that various use-cases will be presented that highlight the applicability of AMP-D. All of these use cases have been implemented in the AMP-D prototype which is presented later.

9.3.1 Basic Concepts

9.3.1.1 *Pervasive and Ambient Floor Display*

To provide an always-available display, the floor is well suited since it is the only space that is always existent in our current life. Further, it "must" be looked at regularly, at least while moving which is a benefit for notifications that must not be missed for too long. Besides, it is easy to glance at quickly. Thus AMP-D projects the permanently available display on the floor, yet content is only displayed when required. The floor display lies outside the foveal visual field of the user, therefore it is qualified for peripheral display. Research on peripheral vision and cognitive psychology offers evidence that peripheral vision supports a separate cognitive channel, thereby reducing overall cognitive load [252]. More importantly, the effect of tunnel vision supports users in effectively blending out unnecessary information in the periphery when their cognitive load is high [75]. Inversely, when users' cognitive load is low, the display supports the serendipitous discovery of new information.

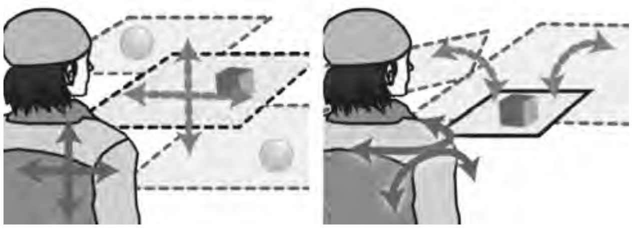


Figure 9.3: The virtual window follows the user's (implicit) body movement (left) and orientation (right)[2].

As the peripheral vision is much more sensitive to movement than to color or detail [75], we adapt the degree of animation on the display to the priority of the shown content and let the user's cognition outbalance load and priority. This functioning was tested with AMP-D in a pilot study and the effect can be described as similar to the sudden recognition of a nearby small animal such as a mouse on the ground that is only detected when it starts moving, even though it was present before. Pousman et al. provide a thorough definition of key characteristics of ambient systems [202]. Based on this definition, to make AMP-D more environmentally appropriate, it "focus[es] on tangible representations in the environment" [202] by refraining from including any typical GUI elements such as windows or buttons on the display. Instead, the projection only shows a projected window into the user's virtual world, i.e. invariably, all projected content is clearly located in the worldwide coordinate system. This concept builds on the Spotlight metaphor (Sub-

2 All concept images depict user movement through red arrows and system animation via blue arrows. Concept images in this section are courtesy of Julian Seifert.

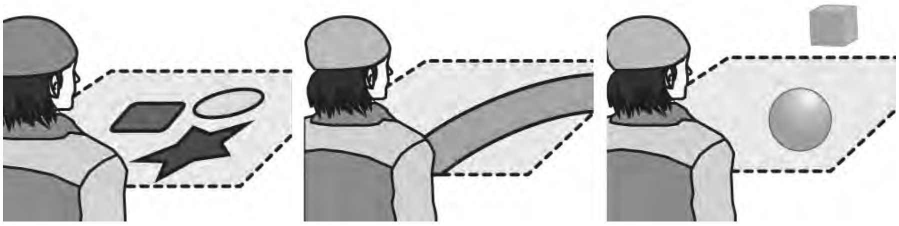


Figure 9.4: 2D World Graffiti (left), navigation paths (middle), and spheres and boxes as interactive elements (right)

section 2.5.4), SAR [48] and world-fixed presentation [86], as opposed to the standard display-fixed presentation. As in case of AMP-D the spotlight is worn, a wearable lantern makes for a better metaphor, though. The system tracks users' movement and orientation to provide the corresponding illusion (Figure 9.3). As all content is only revealed on a fixed location on top of the real world, the projection blends with the environment, for the user as well as for spectators. The publicity of the projection might also lead to interesting new behaviors. For instance, seeing a person uncover public content such as a sign or advertisement with the projection may lead spectators to explore the item themselves with their own AMP-D (or similar AR) devices. Thus the public floor projection also provides a new way of blending real and virtual worlds between users and spectators.

9.3.1.2 *Information Space: World Graffiti, Boxes, and Spheres*

The virtual world of AMP-D consists of only two distinct types of visualizations: two-dimensional *World Graffiti* and two three-dimensional objects: *boxes* and *spheres* (Figure 9.4).

The two-dimensional graffiti is a stationary texture on the ground, such as a navigation path or personalized advertisement. Its flatness indicates that it is not meant to be interacted with. In contrast, the three-dimensional box and sphere items indicate that they are supposed to be interacted with. We choose and limit the system to these two shapes, as it enforces consistency for the user who can approach items from arbitrary angles and they still look familiar. Of course, both visualizations can be combined to create interactive World Graffiti by placing virtual items on top of it.

Spheres are always rolling, accompanying the user wherever they go until they interact with it, or until the sphere is no longer required. For instance, an incoming phone call is represented as sphere item that accompanies the user until the call is taken, declined, or eventually missed. As opposed to spheres, boxes typically lie at static places. However if they are new and supposed to have an ambient alerting impact on the user (e.g. a notification), they roll into the user's field of view.

Figure 9.5: The information space concept splits applications into urgent (spheres) and non-urgent information (boxes). The latter is presented to the user differently depending on its nature (explained in the text). The type of information can be quickly recognized at a glance and also whether it is new, indicated by a yellow border.

If the user is currently moving, they further accompany the user for several seconds before coming to rest.

Boxes and spheres have defined content types which the user can quickly recognize from their different textures as seen in Figure 9.5. Additionally, new boxes the user has not yet interacted with, carry a yellow border. In this manner, unlike with the use of ambient vibration alerts in smartphones, the user can quickly discern the type and novelty of new notifications by just glancing at the projection. To further interact with the box or sphere, users use their bare hands which are tracked by the system. By reaching out with their splayed hand towards the object, a green selection disk appears in the projection. It acts as hand extension that can be moved beneath the object of interest that is *out of reach*. As long as the selection disk remains beneath the object, it performs a subtle bouncing animation to indicate its pre-selection. By closing their fingers, the user selects the object (picks it up) and the object performs a jump animation into the user's hand (see Figure 9.6), thereby transitions to *within reach* interaction. To make the selection of a new object that rolls into view even easier, the rolling object can be picked up instantly by directly moving the closed hand into the projection path, skipping the selection step for the rolling object. This further simplifies interacting with new content while walking.

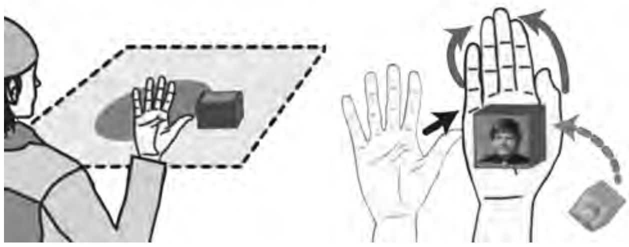


Figure 9.6: Object selection (left) and pick-up of objects by moving the fingers of the hand together (right).

Initially a lot of different hand gestures have been experimented with, especially for picking up objects and releasing them again. Up and down gestures come to mind quickly, but, as other commonly used gestures, do not work well because they inhibit movement themselves.

As the user is moving anyway, we found that gestures based on hand postures work best, followed by gestures that only inhibit horizontal movement.

9.3.1.3 *Private Hand Display*

Previous works [104, 106, 240] have demonstrated that various interactions can be performed in people's hands. The hand display perfectly fits our envisioned interaction scenarios, as many actions can be performed without a separate device. In contrast to the floor display, AMP-D's hand display supports two-dimensional, display-fixed content.

As soon as content has been picked up to the user's hand, the focus of the projection changes to follow the user's hand, showing a focused image within the hand's boundaries. Consequently, the floor projection becomes blurry. This provides the user with a very suitable means of knowing which interaction zone is currently active.

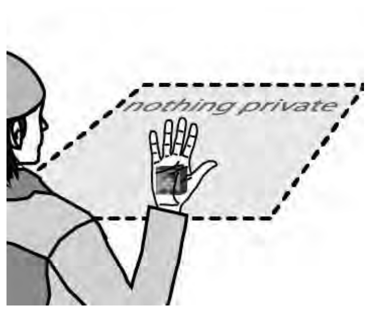


Figure 9.7: Private content is only revealed in the user's hand or on the user's phone.

The hand provides a more private display than the public floor projection (see Figure 9.7), comparable to a phone display. When picked up, many objects can disclose more sensitive information. Message boxes, for example, can show a picture of the sender of the message. Hand gestures allow the user to interact further with the content. By turning the hand 90 degrees towards the center of the body, the user switches to *reading mode*. The box held in the hand shrinks and a preview of the content is displayed. For instance, a text message or the subject of an email as scrolling text, or a preview of a picture can be displayed (see Figure 9.8). Of course, touch interactions similar to [104] could be supported as well, but required bi-manual input which was not the focus of this work.

Figure 9.8: By turning the hand 90 degrees a preview (scrolling text, thumbnail, etc.) of the content in hand is displayed.

Instead, binary decisions (e.g. accept a call with or without loudspeakers activated) can be supported by uni-manual interaction. Users toggle between binary options by flipping their hands so that the thumb points inwards or outwards and select the option by performing a click gesture by briefly moving the thumb to the palm and back (see Figure 9.9).

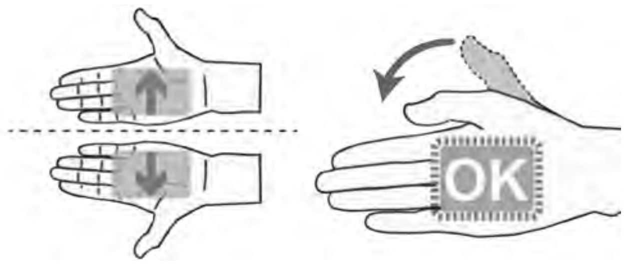


Figure 9.9: Binary decision gestures.

As long as the user holds the box in hand, it moves with them. This way, users can reposition any three-dimensional item in their virtual world. Finally, users have two options how to proceed with the object: By splaying out their fingers and/or moving their hand outside of the projected area, the item falls down back to the floor in front of them. Or, by performing a gesture as if to throw the item over one's own shoulder, the item is removed from the virtual world (see Figure 9.10). The meaning of these gestures depends on the content type (use case) and is explained later.

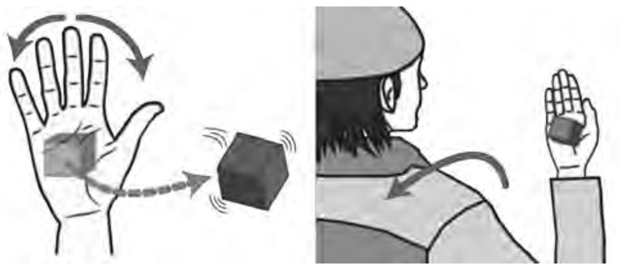


Figure 9.10: By unfolding the fingers the object falls back to the ground (for notifications this is equivalent to "snoozing" the alert). By throwing the object over the shoulder it is removed from the virtual world.

9.3.1.4 *Continuous Interaction Space: Floor / Hand / Smartphone*

In situations when the hand display is not sufficient to adequately deal with content (e.g. to answer a text message), interaction can continue on the smartphone. When a user picks up an object, the smartphone is notified of the selected item. For as long as the user holds the item in hand and a short grace period after that, the smartphone is automatically unlocked and presents the full representation (e.g. the full email view) either immediately, or as soon as it is taken out from where it was stowed (see Figure 9.11). When users are in a private environment or to support collaboration, they may also want to show full content representations on the large floor display. AMP-D could easily support this through an additional gesture. Also more complex interactions such as text-entry could be supported on the floor projection but are outside the scope of the presented concept for information visualization and management.

Another usage of the smartphone is to add content from the smartphone to the virtual world. Despite auto-generated items, the user may want to share content at specific locations. For a personal purpose, for example, a reminder such as "don't forget to put the bins out" can be placed on the threshold. Moreover, pictures, for instance, can be dropped to the world where they were taken to share them with friends or the public (explained in a moment). The smartphone provides an always available "share service" that allows supported content on the smartphone to be dropped into the environment as virtual boxes of the corresponding type (see Figure 9.11).

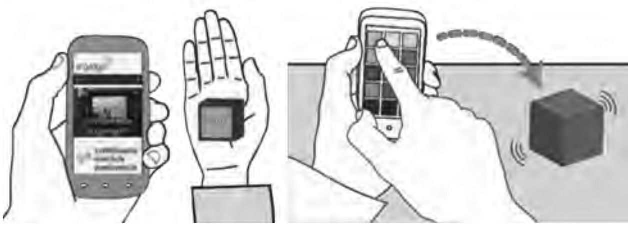


Figure 9.11: Content in hand can be further explored in detail on the phone (left) and arbitrary content from the phone (notes, pictures, etc.) can be placed as box item in the virtual world (right).

Thus, the three interconnected displays form a continuous interaction space providing different levels of details to the user's requirements at that time.

9.3.1.5 *Public Content And Collaboration*

Besides personal content of the user, the constant availability of the projection invites friends, advertisers, or even the local administration to create public content in the virtual world similar to physical signs, banners, etc. The virtual content has the advantage that it is much cheaper to create and can be personalized for the user. Its disadvantage is that the projection is smaller than a person's visual field, therefore it may not reach the same audience. The intrinsic publicity of the projection also invites many colocated multi-user scenarios. For instance, colocated AMP-D users can overlap and "merge" their floor projections very similar to the SurfacePhone (Chapter 7) to drop content from the smartphone to the shared floor display, to then be picked up on the other side by another user.

9.3.1.6 *Privacy*

When using AMP-D, users neither exclusively select the projected content, nor do they monitor the projection all the time as is the case with existing projected displays. Furthermore, the surrounding audience of the projection is rather random. Projected content may be appropriate in the current context, but not in another. Thus, users require effective means to protect sensitive information and ensure that only appropriate information is disclosed. A first means is already given through the concept of SAR. When a user wants to hide information on the floor display quickly and for a short moment only, a small movement or rotation along any axis is often enough to move the window in order to hide previously disclosed items. If this is not sufficient, AMP-D supports a simple horizontal swipe gesture in front of the projector to enable/disable the projection entirely (see Figure 9.12).

9.3.1.7 *History and Overview*

The SAR concept entails another advantage within the context of AMP-D. The implicit revealing or hiding of information using body motion can also be used to look up upcoming content or to revisit past content. For instance, when a user recognizes content on the floor projection too late, walking a few steps back or just turning around will bring the item back into the projected window. Similarly, when users share their foot trails as World Graffiti, they can revisit them later, e.g. to find their way back to their car. As opposed to that, for instance, tilting the projection far ahead during navigation tasks allows users to preview directions further ahead (the inclined reader may also look at NaviBeam [W12] where this was investigated first). Results from a study by Billinghurst et al. [48] indicate that people can easily navigate and relocate spatially augmented information as they are used to the interaction from real life. Therefore, the SAR (or lantern/spotlight) concept should be able to provide a natural interaction with the information space of AMP-D.

Despite AMP-D's support in changing the FOV in all directions, no contemporary display technology can compete with the overview (visual field and resolution) of a person's real view into the distance. Thus, searching for virtual content on the ground can require significantly more effort than searching for real objects. Therefore, the system provides vertical swipe gestures to change between AMP-D's standard view and an elevated third-person perspective. This acts like a map on a scale of 1:10 to spot nearby objects of interest without having to actually go there or search for them (see Figure 9.12).

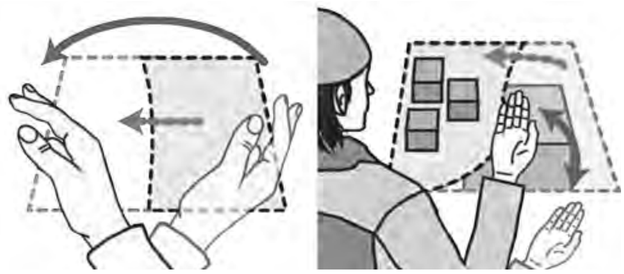


Figure 9.12: Enable/disable the projection entirely with a horizontal gesture (left). Change from a real to an elevated perspective to get an overview of content around (right).

9.3.2 AMP-D Use Cases

AMP-D is meant as a general mobile companion and the following implemented use cases highlight how AMP-D supports typical mobile scenarios.

9.3.2.1 *Context- and Location Awareness*

Context- and especially location-aware information such as friends being nearby, or interesting offers in the vicinity are increasingly available to users. With AMP-D being constantly active and capable of displaying visual context information, it is well-suited to provide such information to users on the go. People on a shopping stroll, for instance, see additional offers as World Graffiti and box items (textured with a t-shirt icon) on the ground in front of the shopping windows they pass. By picking up the box and reading its contents, a personalized advertisement appears in the user's hand (Figure 9.13a).

The system further supports persuasive hints. They have been shown to be able to motivate users to change their behavior in a positive way [250]. For instance, when users walk close to an elevator, the system reminds them of their activated fitness plan by showing a red line leading to the elevator and a green line leading to the staircase as World Graffiti beneath the users' feet (Figure 9.13b).

(a) Personalized Advertisements

(b) Persuasive Computing

Figure 9.13: Two location-aware use cases

9.3.2.2 *Data sharing*

Data sharing via the smartphone is used to create new virtual items in the virtual world. Currently, we support two data sharing applications on the smartphone. The first allows the user to create note boxes with text content at world-fixed positions. This can be used to place location-dependent reminders or messages for oneself, or, for example, a colleague or family member in the own environment who will literally stumble over the information (Figure 9.14a), and can read the contained message in their hand. The second application supports the sharing of pictures from the smartphone's gallery. The boxes are created right in front of the user (Figure 9.14b) and are textured with the contained image (in a gray inset, identifying them as boxes of the type "picture"). Given the small size of the boxes, it is not possible to recognize the actual image on the floor projection, but it is already useful to distinguish between different items in a pile of pictures. Once users pick up an image box they are interested in, the image is revealed in the user's hand when entering reading mode. This presentation already delivers a much better impression of the picture than the floor projection. As with other content types, the picture can further be viewed on the phone.

Once boxes are created, they can also be easily repositioned by taking them by hand, moving with the box to the desired location, and releasing them again.

(a) Node Sharing **(b)** Picture Sharing

Figure 9.14: a) A note is found in front of an office door ("Meet me at coffee place"). b) A picture box transfered from phone to virtual world.

9.3.2.3 *Notifications: Communication, Reminders, News Feeds*

The most frequent tasks performed on smartphones - especially while the user is on the go - are related to communication, and management of notifications. Calendar and task reminders, for instance, have become very popular on smartphones. The most important aspect is to actually read them, be reminded at regular intervals if the notification was missed in the first place, and perhaps perform some simple interaction such as snoozing or dismissing the notification. For the user on the go reading the notification on a smartphone often involves a significant overhead. The user must stop or at least slow down, take out the device, possibly unlock it, only to read a few words of text.

AMP-D uses its box items to visualize new notifications regarding text messages, emails, calendar reminders, and news feed updates. As described earlier, the user can quickly discern the type of the notification from their appearance prior to any interaction (see figures 9.5 and 9.2a). When a message box is picked up, a picture of the sender is displayed on the upper face of the box. The first 160 characters of the item's content are displayed as scrolling text in the user's hand when turned to reading mode (see Figure Figure 9.2b). Otherwise, only a teaser is displayed, or the subject in case of an email, and the whole message can then be read, for instance, on the smartphone. Similarly, news feed updates appear as feed boxes that show their source (publisher) as a texture on the box, reveal their subject in the user's hand, and can be continued to be read on the user's smartphone. They particularly demonstrate the usefulness of serendipitously stumbling over new information when the cognitive load is low.

The visualization of dynamic notifications using the world-fixed SAR concept is not straightforward as the information is time- and context-dependent instead of location-dependent. A solution is to multiplex the information in the time and location domain. For instance, when users receive a new notification, it is created at the user's current location and rolled into their projected window. Shortly after the user passes by the notification without picking it up, it is removed from the old position and inserted in the same animated way at the user's current location. Once the notification box has been picked up (thus noticed), users decide whether they want to return the box to their world to either "snooze" (cf. Figure 9.10 left) the notification or dismiss it by throwing it over their shoulder (cf. Figure 9.10 right). In the former case, the box will continue to regularly appear across the user's way but without any type of animation (these intervals can be defined on the corresponding smartphone app, Figure 9.19b). In the latter case, the notification—not the content—is deleted.

Incoming calls, in contrast, are presented as a sphere that accompanies the user for the time of the call. It can be picked up to show the caller's picture—if available—as texture on the sphere and to reveal the name or phone number in the reading mode. In this scenario, taking out the smartphone after having picked up the sphere will automatically unlock the smartphone and accept the call; releasing it to the world will keep it ringing; and throwing the sphere over the shoulder will decline the call.

(a) Path navigation **(b)** Signs

Figure 9.15: Path navigation with additional turn-by-turn instructions and signs that help ahead-way planning

9.3.2.4 *Navigation*

AMP-D supports augmented reality navigation where the user follows a path as a virtual line overlaying the ground. The path approach fits the concept of World Graffiti on the floor display, since it gets by without time-based instructions the user could miss. Instead, whenever the user pays attention to the projected navigation, directions can be grasped at a glance (Figure 9.15a). Additionally, to provide the necessary context for users to plan their way upfront, turn-by-turn navigational instructions (e.g. turn left in 50 meters) and location-dependent help (e.g. take the stairs up, not down) are overlaid right next to the path similarly to road signs (Figure 9.15b).

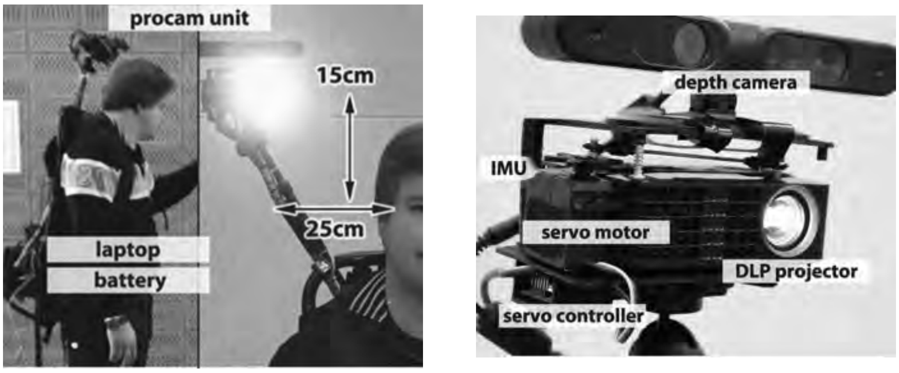
9.4 PROTOTYPE

The AMP-D concepts were implemented in a prototype in order to investigate the technical challenges involved in building such a system and explore possible interaction designs for aforementioned use cases.

9.4.1 Hardware design

The AMP-D prototype (see Figure 9.16) consists of a ProCamS unit, a backpack, a laptop, and two batteries (one for the projector, one to increase battery time of the laptop). Part of the overall system is also a wirelessly connected Android smartphone running a corresponding service software.

The ProCamS unit (Figure 9.16b) is attached to the backpack (Figure 9.16a) that positions it approximately 25 cm to the side and 15 cm to the top away from the user's eyes (no offset in the forward direction). We found this position to yield a good trade-off between the large size of the pro-



(a) System setup (b) ProCam unit

Figure 9.16: The AMP-D prototype. a) A backpack holds the ProCam unit, laptop, and battery (left). The ProCam unit appears as worn on the shoulder (right). b) Close-up of the ProCam unit.

jection on the ground ($\approx 2\,m \times 1.25\,m$ (≈93" diagonal), 2 meters away) and the positioning of the hand in order for the user to comfortably reach into the projection for the selection of objects and the hand display.

The projector is a Dell M110 providing 300 Lumens and a resolution of 1280 px × 800 px. On top of the projector sits an ASUS Xtion Pro Live depth and RGB camera (640 × 480 px @ 30 FPS), which was chosen for its very small size, low power consumption, and well-suited depth range (0.5 m to 3.5 m). Finally, an inertial measurement unit (IMU) by x-io technologies is part of the ProCamS unit and delivers precise and robust 9DOF orientation and acceleration data of the user. The system is controlled by a Dell M11x laptop with i7 CPU, 1.7 GHz running Windows 7 and the prototype software that performs all computations at a frame rate of 60 Hz.

The projector and the IMU are powered by batteries and the rest of the components are powered by the laptop. The system's power lasts for 5 hours of continuous usage. The system can be worn to both sides, depending on the primary hand of the user, which should be on the same side as the projector to lie within the projection path.

9.4.2 Software

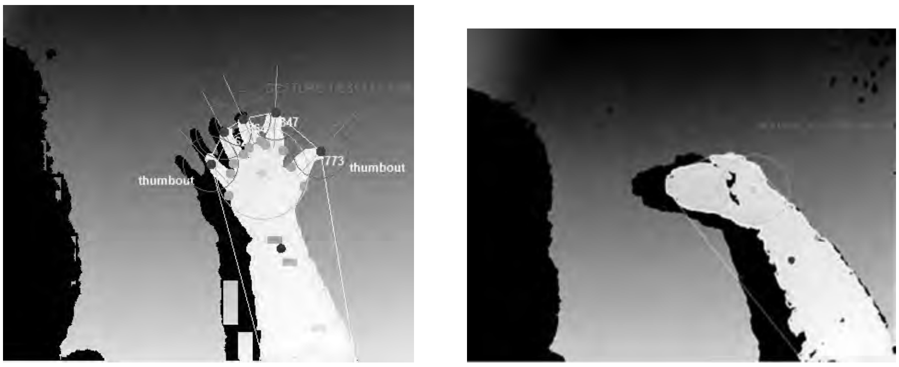
The software components (except smartphone service software) are responsible for computing the graphics and physics of the 3D world augmentation, the user's movement, and the hand and gesture recognition. The software is written in C# and runs on the laptop.

9.4.2.1 *3D World Augmentation*

The virtual user and all projected content are modeled in a 3D scene using Microsoft's XNA, the JigLibX [127] physics engine, and a graphics framework [205]. The rendering of this scene delivers the 3D world augmentation to the projector.

The 3D scene includes World Graffiti as 2D floor textures, and 3D boxes and spheres. The skeleton of the virtual user, who moves through this world, consists of a lower and an upper body entity. The correct perspective projection is achieved by attaching the virtual camera to the user's torso entity with the exact offset that it has in reality. The engine will then compute the correct off-axis perspective homography that lets the projection appear as perceived through the user's virtual eyes. Moreover, it lets the virtual camera turn around the center of the user's body instead of turning around itself. In addition, the virtual field of view has to be inversely matched to the actual field of view provided by the projector. Currently, lens distortion of the projector or the small offset between projector and depth camera are not accounted for, both of which would further improve registration accuracy.

As we use a fixed physical orientation for the projector independent of the user's height, we can automatically calculate this height that is required by the system, based on the floor distance we receive from the depth sensor. Thus the system does not require manual calibration. The accuracy of the optical illusion during tilting or rolling of the torso can be further improved, though, by providing the exact length of the torso to the system in order to accurately determine the center of the body.



(a) Finger tracking (b) Posture tracking (here "reading")

Figure 9.17: Software tracks hands, fingertips, and posture for gesture-based interaction.

9.4.2.2 *Floor/Hand Tracking and Focus Adjustment*

Floor and hand tracking is computed on the depth image from the camera. On every cycle, the algorithm first decides whether the user's hand is present in the depth image:

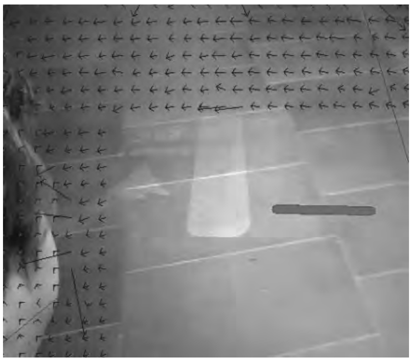
We use computer vision to recognize hand contour, finger gaps and tips, fingertip direction, the direction of the hand, and the centers of palm and hand (see Figure 9.17a). The recognition builds on three-dimensional segmentation of hand-sized clusters and simple heuristics based on sizes, distances, and changes in the derivation of the contour. Our particular shoulder-worn setup allows some assumptions that further simplify the recognition procedure: valid hands must not be further away than 1.5 m (depth culling); must not span a depth range larger than 0.5 m; and the user's arm (the cluster) must always reach into the camera frame from the bottom and/or right edge (for right-handed users). The recognition is fast and accurate in various environments. When more than one finger is recognized, we detect the *unselected* state that allows the user to steer the green selection disk (cf. Figure 9.6) for object selection. When one or no fingers have been recognized, we detect the *selected* state. Further, we recognize the user's thumb and compute its relation to the center of the hand to distinguish between the two states of binary decisions. Comparing hand positions, hand directions, and finger counts over multiple frames allows us to recognize the remaining gestures such as *reading mode* (cf. Figure 9.8 and see Figure 9.17b), click gesture, and the horizontal and vertical *swipe gestures* (cf. Figure 9.12).

When the user's hand is *not* detected, the surface in front of the user is analyzed to decide whether it is suitable for showing the floor projection. The depth image is sampled at several grid-based points across the image and first averaged individually for each row, then for the whole image. Based on the depth gradient from individual rows we can decide whether the projection falls on a floor-like (vertical), mostly plain (depth deviation) surface. Additionally, based on the overall depth average, we can then adjust the projector's focus to show a sharp image on the floor.

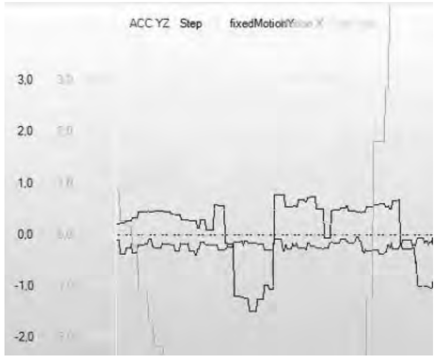
9.4.2.3 *Tracking of Orientation and Movement*

In parallel, inertial sensor data is received from the IMU. It is used to compute the orientation of the user's torso in all three dimensions to adjust the virtual user and the attached virtual camera in the 3D world accordingly.

Additionally, we use the acceleration data from the IMU for step detection. As absolute positioning systems are not always available, particularly indoors, AMP-D needs a way of detecting the user's movement



(a) Optical flow



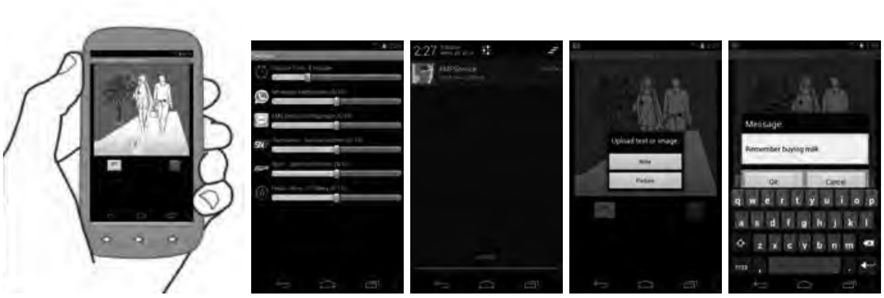
(b) Fused with inertial measurements for step detection

Figure 9.18: Orientation and step detection

based on dead reckoning. Naturally, this will only work for a short time due to measurement errors and must be regularly corrected by reliable data from an absolute positioning system (e.g. GPS or indoor localization systems increasingly becoming deployed). For the sake of testing the AMP-D concept, we only require short movements for which the dead reckoning approach is sufficient. Algorithms for robust pedestrian navigation usually build on a technique known as zero-velocity-update that relies on data from an accelerometer attached to the user's foot (e.g. [149]). Following our initial vision of a palm-sized form factor of the system (Figure 9.1), we want the system to get by without further user instrumentation. With the IMU unit attached to the ProCamS unit, we cannot rely on the zero-velocity-update method. Instead, we detect steps by finding peaks of vertical and forward acceleration, which are homogeneous in human walking behavior. Step length is approximated based on the automatically calculated height of the user. With the IMU unit alone, we could not reliably detect the user's walking *direction*, though.

A working solution which increased the reliability of detecting the step direction was found in computing the optical flow of the camera's RGB image. More precisely, the optical flow is calculated in a 100 px wide border at the top and left side of the RGB image (for right-handed users) wherein the user does not interfere while interacting with the primary hand (see Figure 9.18a). Optical flow towards the user indicates forward movement while optical flow away from the user indicates backward movement. Gyroscope data is used to counterbalance the effect on the optical flow generated by the up and down swings caused by the human gait (Figure 9.18b).

Combining these approaches, our system can detect the user's forward and backward steps 9 out of 10 times, which is sufficient for our investigation but leaves room for improvement. By decreasing the form factor of the prototype, for instance, the system can be brought closer to the



(a) (b) (c) (d) (e)

Figure 9.19: The smartphone application that belongs to AMP-D. It allows the service to be enabled/disabled (a), adjust settings for the frequency of notifications (b), and share supported types of content by sliding down the top bar (c-e).

user's body which will benefit a more accurate step detection. Nonetheless, a general problem with step detection based on inertial measurements would remain: as the algorithm cannot detect a step until it is close to being finished, a small computational delay is induced. This delay counteracts the optical illusion when walking starts or stops and sometimes leads users to walk one step further than they intended.

9.4.2.4 *Smartphone Service*

The Android smartphone is connected to the laptop via Wi-Fi and runs a background service (Figure 9.19a) which starts polling the phone's light sensor whenever the user takes a virtual box into their hand and stops soon after it was released again. Whenever the measured light significantly increases during this time interval—an indication of the phone having been taken out of a pocket or bag—the service wakes the screen, disables the key-guard, and starts the corresponding activity showing the content related to the box. In addition, access to placing information in the world from the smartphone is provided through a notification service. By pulling down Android's notification center from the phone's status bar and selecting the AMP-D service (Figure 9.19c), the user can create notes (Figure 9.19e) and select pictures from the gallery which are then dropped into the world. The app further allows the frequency of notifications to be set per app (Figure 9.19b).

9.4.3 Limitations and Improvements

For most indoor scenarios, the brightness of the displays of the present prototype is already sufficient. For most outdoor scenarios, only the hand display is sufficiently visible as it is very close to the projector and can be shielded against direct sunlight. To the floor projection these reasons do not apply, hence the prototype is currently limited to twilight conditions outdoors. As the brightness of pico-projectors constantly increases (cf. Figure 2.3 on page 24), they may reach sufficient brightness for outdoor conditions someday in the future.

Another limitation is the current size of the system. With pico projectors advancing quickly, and depth cameras of the size of a finger becoming available, the ProCam unit can likely be shrunken considerably in the near future. Additionally, the software should soon be able to run on small on-board systems such as a Raspberry Pi or Intel Compute Stick or mobile phones with built-in depth cameras such as Google's Tango. Power consumption will likely be the most challenging factor for a much smaller version. This could however be mitigated by intelligent power saving which reduces power consumption when no display surface is available.

Finally, the step detection needs to be further improved, e.g., by pointing a second camera towards the user's feet which can immediately decide whether the user starts moving, thereby eliminating the initial detection delay of the current system. Further, GPS can be used outdoors or geo-located image-based pose estimation indoors (cf. [28]) to correct step detection errors.

9.4.4 Initial Evaluation

As Subsection 1.3.1.3 has pointed out, longitudinal studies that would be interesting and required to perform on AMP-D, are often not possible with the current sophistication of projector technology as they are not bright enough or too bulky in size. An initial investigation, though, wanted to find out if the most important features of AMP-D work for untrained users. Thus 6 participants between 25 and 30 years (mean 27 years) were recruited, to identify strengths and weaknesses of the concept or the current implementation. They have been smartphone users for 1.5 years on average (min 0.5, max 2 years) and all used their smartphones at least for messaging, web browsing, calendar, and traffic information. The study lasted between 45 and 60 minutes and was conducted in a public floor (12 x 7 meters) of the university building with regular by-passers (to provide users with experiencing their attitude towards public usage).

9.4.4.1 *Procedure*

First participants were asked (using 5-point Likert-scales from "strong agree" to "strong disagree") about their regular smartphone usage. All showed *strong agreement* that they receive regular updates (notifications) on their mobile phones. There was further *agreement* to check if notifications have been missed, also while on the go, by all participants. Finally, there was *strong agreement* by all participants that they usually react to new messages immediately. These answers show that the participants reflected the right target audience that is addressed by AMP-D.

After that all participants tried out all applications of the prototype. This includes: receiving boxes rolled into their view while walking, looking straight ahead, picking up boxes, reading their contents, moving them, releasing and dismissing them. Further continuing reading the contents of a box on the phone as well as taking a picture with the phone and creating a reminder note on the phone and sharing both to the own virtual world. Finally, they also tried to follow a navigation path that led them a marked route around obstacles we had set up.

After having tried the AMP-D prototype, participants showed a generally very positive attitude towards the AMP-D. Again they were asked using the same 5-point Likert scale. All participants at least *agreed* that they recognized new notification items on the floor without looking at them. Further, all but one assumed the system would not disturb but enrich their daily life if it was available at a smaller size. Further, all participants at least *agreed* that they think they could react to new information more quickly using AMP-D versus a smartphone. Finally, all *agreed* that the prototype worked fine for them, that they enjoyed using it, and that they could handle its complexity. In contrast, users were split in their answers to our questions regarding social acceptance and price/performance ratio—considering AMP-D would double the costs of the smartphone—both resulting in overall *neutral* feedback.

In response to open ended questions participants criticized, for instance, physical fatigue caused by the high number of interactions tested in the user study. Two participants were concerned with performing large, eye-catching gestures in public space. The majority of users questioned whether the floor display would be bright enough outside (which had to be negated at that time). But users also came up with constructive comments regarding the technical challenges like brightness, battery life, and size of the system: one participant, for instance, proposed to show and select between all objects in the vicinity along a virtual string in the hand when the floor display is not bright enough, which would be similar to the SpiderLight system presented in the next chapter. Further on, participants suggested several new application scenarios, among those: using AMP-D for navigation and context-aware instructions for craftsmen on building sites; remotely placing reminder boxes for items to buy across the supermarket at the right locations (like a decentralized shopping list); similarly, using AMP-D as city tour guide with POIs realized as info boxes to stumble over interesting information while keeping connected to the primary interest, the environment.

Naturally, this evaluation is only a first step in evaluating the device. User studies with more users over longer periods and against ordinary smartphone usage, for instance, are required.

9.5 CONCLUSION

This chapter introduced the case study on the Ambient Mobile Pervasive Display (AMP-D). For the first time it investigated the use of *constant* personal projection for a personal pervasive display that accompanies the user in the visual periphery. AMP-D provides a wearable MMDEs that combines an ambient out-of-reach floor display, a private within-reach hand display, and a smartphone into a continuous interaction space for mobile everyday scenarios. By following the Nomadic Projection Within Extended Reach concept, it notifies users about new information on the less bright display out of reach, to allow objects of interest to be easily brought within reach to the user's hand for further inspection and management on a bright display. The demonstrated use cases highlight the applicability of AMP-D in mobile scenarios. These are embedded into a consistent information space concept that uses Spatial Augmented Reality (SAR) together with World Graffiti and virtual boxes and spheres to cover a broad and extensible range of interaction scenarios.

Moreover, the complex prototype demonstrated unique technical solutions to the specific challenges of the AMP-D concept like:

- automatically changing the projector's lens focus to the user's zone of interaction,
- step detection fusing inertial and optical movement measurements,
- and tracking of novel hand gestures in a truly mobile environment.

These components have been integrated to a standalone mobile system that does not require instrumentation of the environment or the user (despite wearing the system), and runs for several hours. The conceptual and technical solutions have been the result of a two-year long evolution, starting from similar ideas that had originally been investigated and tested with users for indoor shopping malls and handheld projection in the NaviBeam project (cf. Winkler et al. [W12]). The positive user feedback during our evolution of AMP-D has been the consequence of this evolution.

Mobile Deficiencies

MULTI-TASKING The concept further presents a new approach to serendipitous access to digital information (by stumbling over it) that can be applied to our physical world, thereby likely reducing the individual's effort to receive and deal with information. In particular, the system can both be perceived as well as operated partly hands-free (through body movement), allowing for other tasks such as walking or working in large spaces (as was recently presented by Audi [29]). the simple select and pick-up gestures (or just pick-up if information is brand new), make for a really short lead-time to interaction compared to smartphones, which first have to be taken out of where they are stowed. Even compared to smart watches, whose display is about 62 times[3] smaller than that of AMP-D, AMP-D is likely able to deliver a quicker information overview during walking. Thus, regarding R6, the large size of the projection and its characteristic of being always ready-to-use, enables a short lead-time to interaction, maybe even allows for peripheral interaction on-the-go (cf. Winkler [W6]). As such, it depicts a future direction towards the original vision of pervasive computing.

> *R6 page 8*

ENVIRONMENT Although the sample size has been smaller in this study than it was in previous case studies, not less than all of the participants indicated that they think they would be both "more aware" of their personal information, as well as quicker than with a smartphone in accessing and managing it. These results at least indicate that regarding R5, the projected floor display increases awareness unlike all carried or worn displays that are not persistently in the user's periphery are able to do (speaking only of visual capabilities).

> *R5 page 8*

Hopefully, further advancements in projection technology will enable larger user studies that assess how the display can blend into the fabric of everyday life and investigate how well AMP-D would perform in crowded places. In western societies, we recently got used to audio pollution—people making phone calls on the street, during public transport, and even sometimes in restaurants has become more or less acceptable. Maybe, the visual pollution created by worn projectors would eventually experience a similar acceptance, especially if the amount of *fixed* pervasive display space and its pollution of the environment could in turn *decrease*.

3 based on the 1.49" display of the Apple Watch. The resolution is only 8.4 times smaller, though.

10

CASE STUDY ON *SELECTIVE* NOMADIC INFORMATION MANAGEMENT IN THE PALM WHILE ON-THE-GO

Deficiencies **addressed by this chapter**

Output/input size *(D1)*

Multi-tasking *(D2)*

Collaboration & Privacy *(D3)*

Environment *(D4)*

This chapter is based on the previously published refereed book chapter:

[W1] Gugenheimer, J., **Winkler, C.**, Wolf, D., Rukzio, E., "Interaction with Adaptive and Ubiquitous User Interfaces." In: *Companion Technology - A Paradigm Shift in Human-Technology Interaction*. Ed. by S. Biundo, A. Wendemuth, and A. Bundy. Red. by J. Carbonell, M. Pinkal, H. Uszkoreit, M. M. Veloso, W. Wahlster, and M. J. Wooldridge. Cognitive Technologies. Springer, 2016, to appear

In addition, the following partially related thesis was supervised by the author:

- "Development and evaluation of a wrist-worn projector-camera system enabling augmented reality". Philipp Schleicher. Master's thesis. 2014

In the previous chapter, we expanded the NPWR to the Nomadic Projection Within *Extended* Reach (NPWER) concept to cover the cross-distance interaction space in on-the-go scenarios. One disadvantage of the AMP-D setup has been its comparably large size, though. This was the result of several factors, but mainly of choosing a shoulder-worn setup, which because of the larger distances to the user's hand (on average $60\,cm$) and to the floor (on average $165\,cm$) will always require larger projectors than previous case studies.

As a consequence, in this chapter we are going to investigate how similarly quick micro-interactions that aid multi-tasking (>D2) can be supported from a device that is closer (again) to the projection surface, and thus provides a bright projection from a smaller prototype size. A wrist-worn ProCamS seems to be able to support this goal, but likely requires changes to the interaction design. In principal, a wrist-worn projector could also be used to create a smaller version of the floor display as presented by the previous chapter. Of course, it would be more limited regarding some of the application scenarios as the floor display would not maintain a fixed distance to the user. Yet more importantly, the ProCamS is able to provide the same bright projection onto the (opposite) hand (Figure 10.1a) from a much smaller device size, leveraging the environment (>D4). As with AMP-D, when projecting on surfaces other than the hand, a selection interaction can be used to "grab" the object for within-reach interaction in the hand.

Nevertheless, the primary focus of this case study lies on interaction happening in the palm. Having said that, it seems crucial to understand how palm projected interfaces should be designed which neither the AMP-D study nor other related works have investigated so far. Consequently, the first part of this chapter presents a user-elicitation study that investigates the types of information, the visual design, and the interaction metaphors that users desire when using a wrist-worn ProCamS for micro-interactions in the palm. Based on these findings, a standalone wrist-worn prototype that supports interaction in the palm and on other surfaces using finger and device gestures will be presented. A second user study evaluates the usability of the device compared to smartphone usage, indicating many situations where participants would want to use a wrist-worn projection device but also highlighting the limitations of using the palm for within-reach interaction.

(a) Pressing middle and ring finger to the ball of the hand quickly (de)activates the projection.

(b) The implementation allows to quickly look up weather information

(c) Apart from the always available palm, any nearby surface can be used for better clarity and single-hand interaction. Finger shadows facilitate button selections.

Figure 10.1: The interaction space of the SpiderLight, which delivers quick access to context-aware information using a wrist-worn projector.

10.1 INTRODUCTION

Since smartphones became a ubiquitous part of our daily life, the urge for being up-to-date and accessing context-dependent information is constantly increasing. By observing smartphone users, we see that oftentimes getting hold of the device consumes more time than the actual interaction. Most of the time, the phone is used for micro-interactions such as looking up the time, the bus schedule, or to control a service like the flashlight or the music player [88]. With the recent emerge of wearable devices, such as smart watches, users can access these kinds of information at all times without having to reach to their pockets. However, most of these wearable devices are merely equipped with a small screen so that only little amount of content can be displayed and the user's finger is occluding most of the display during interaction (fat-finger problem).

A pico ProCamS integrated into a wrist-worn device might be able to deliver a larger display that inhibits the same level of quick accessibility, allowing for interactions using the shadow of the fingers (Figure 10.1c) and movement, especially roll rotation, of the arm. Projecting in the opposite hand (Figure 10.1b) that has an average diagonal of 7.74"[32] (yielding about 27 times the display size of the larger Apple Watch model that has a diagonal of 1.49") or on a nearby wall (Figure 10.1c) would provide a larger information display that is always available at the push of a finger (Figure 10.1a). Existing research on wearable ProCamS usually employed head or shoulder-worn projectors as in the previous chapter. However, wrist-worn devices are socially more acceptable at the moment due to their similarity to wrist-watches and achieve a shorter distance between projector and palm. This comes at the expense of making it less suitable for traditional touch interaction (cf. Section 2.5, esp. Subsection 2.5.3) as both hands are occupied, although direct within-reach interaction remains possible as will be shown later. Interaction with a wrist-worn projector had previously been explored by Blasko et al. [53] using a mockup prototype and focusing on wall projection. In contrast, the SpiderLight system presented in this chapter is the first (real) standalone wrist-worn projector device that supports projection in the user's hand (in addition to nearby surfaces).

Further on, previous systems proposing palm projection [104, 240, 281] have all been designed by experts. As paragraph User-elicitation (Subsection 1.3.1.3) explained, user-elicited approaches often lead to interaction metaphors that are easier to learn and more acceptable to new users. In particular, it seems interesting to elicit which content and interaction metaphors users desire for this new type of device. The next section will mark out the design space of such a system. Subsequently, the user-elicitation study, the derived prototype implementation and its evaluation in another user study will be presented.

10.2 DESIGN SPACE

The purpose of the SpiderLight is to facilitate micro-interactions that are too short to warrant getting hold of and possibly unlock a smartphone. Like smart watches, the SpiderLight is not meant to replace the user's smartphone. Instead, it is meant as an accessory to the user's mobile phone that has more limited in- and output capabilities in favor of a much shorter lead-time to operation. Hence, the user-elicited approach for this new type of device considers the following aspects:

1. the (context-aware) content to be displayed,
2. the interaction design with the displayed content,
3. and the activation gesture for quickly and reliably enabling / disabling the system.

As a starting point we will first assess some typical smartphone use cases (applications) that involve such little content and interaction that using SpiderLight seems an advantage. Similarly, we assess the design space of interactions possible with available hardware.

10.2.1 Applications

The following presents the most common smartphone activities that have the potential to get accomplished more quickly by using SpiderLight. Please note that these are not meant to be exhaustive, but to be extended by participants in subsequent user studies.

LOCATION AWARE OVERVIEW The Location Aware Overview is meant as a counterpart to the lock screen on mobile phones, although with a stronger focus on context-relevant information. As such it can provide basic information like time, weather, bus schedules, currently played music, and notifications about new messages and social network updates. It could show the screen that was last active on the smartphone or even provide a direct interface to simple smartphone functions. An advantage of SpiderLight in this regard is that it does not require unlocking as it is steadily worn and not taken off in public—an advantage shared with smart watches. The more private nature of the own hand further allows projecting more sensitive data than, for instance, in the Sixth Sense scenario [173].

MEDIA PLAYER Smartphones are commonly used as media players, which often require control for a very short time. Headset remotes partly fulfill this requirement quite well, although more complex commands such as skipping or reversing multiple tracks or toggling shuffle or repeat modes often require cumbersome interaction if supported at all. Further on, smartphones are increasingly used as source for home entertainment, streaming to wireless loudspeakers or TVs. In this scenario headset remotes are unavailable and quick control from the SpiderLight may be very convenient.

SHARING MEDIA A huge advantage of the mobile projector is that it can easily be used to create large displays in mobile nomadic (indoor) scenarios to share content with other peers. The SpiderLight could provide a coverflow that contains the most recent pictures taken with the mobile phone. Older pictures could be selected on the phone and then shared to the projection of SpiderLight. The high dexterity of the user's arm wearing SpiderLight allows to target a wide range of surfaces, ranging from the floor to tables, walls and the ceiling (similar to handheld interaction).

10.2.2 Interaction Design

In the next step, suitable interaction metaphors for a wrist-worn ProCamS to sufficiently support aforementioned applications were identified. These include:

1. navigation between (hierarchically structured) screens,
2. scrolling,
3. and moving the device in 6 degrees of freedom.

In consequence, as depicted in Figure 10.2, the hardware is supposed to possibly allow for

1. 6 DoF movement of the projector and shaking of the device;
2. touching and swiping on the projection with the second hand;
3. moving (collapsing and spreading) fingers of the hand of the device (casting shadows);
4. moving (collapsing and spreading) fingers on the hand of the projection;
5. around-device-interaction of the second hand;
6. using speech input;
7. pressing buttons on the device.

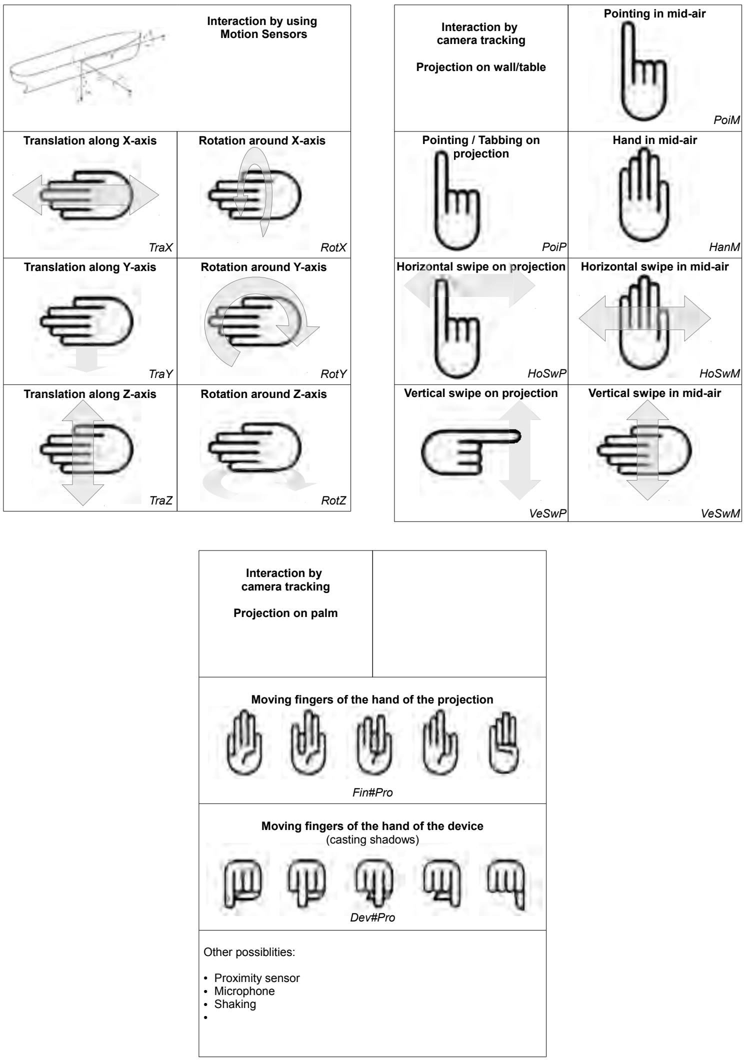


Figure 10.2: Overview of available interactions participants could choose from in the first study

Pros and cons of most of these techniques as related to mobile projection have already been discussed in Section 2.5 and are irrelevant for the user-elicitation. Section 10.4 will later discuss technical considerations based on the outcomes of the user-elicitation study and within the specific context of SpiderLight. The third part of the study considers the activation gesture.

10.2.3 Activation Gesture

Where AMP-D used a large horizontal swipe gesture for activation / deactivation of the projection, the SpiderLight case study focuses more on socially acceptable form factors and interactions. Furthermore, a very small device dictates a smaller battery that will require the user to activate / deactivate the device much more often. An essential aspect of supporting a short lead-time to device and a quick information lookup is thus the design of a quick and easy to perform activation gesture. At the same time, the activation gesture must be robust against being performed accidentally as an enabled projector that is moved unconsciously can be very distracting to the environment.

For instance, it was considered to use a certain sound for activation of the device, such as snapping fingers, but discarded for its lacking robustness. A particular word or phrase like introduced by Google with the "Ok, Google" phrase might work, but could still be socially unacceptable in many situations. The device could further be activated by motion. For instance, a sudden movement, such as quickly turning left and right, could be used for turning the projection on. Another considered approach was to place a button on the back of the wrist that is simple and reliable. On the other hand, it would require a second hand for operation. Considering possibilities for an activation with the same hand, it became apparent that the creators of the Spiderman comic series must have faced a similar question: which one-hand gesture is not performed accidentally, yet easy to perform. Bending back the hand over and pressing a button in the palm with one or two fingers (Figure 10.1a) as used by Spiderman to shoot his webs, is such a gesture and was therefore proposed to participants in the user-elicited study along the previously mentioned alternatives.

10.3 USER-ELICITED STUDY

Nine participants (two female) of an age between 23 and 29 ($\bar{x} = 26$) were invited to our lab to learn about their preferences towards desired content, interaction with and placement of a wrist-worn projector. Three of them were left handed and six were right handed.

10.3.1 Procedure

The user study comprised the following steps: First participants got introduced to the general idea of SpiderLight using application ideas and possible interaction concepts as previously described. Then they created two iterations of palm and wall UI designs for the device using paper prototyping, each time followed by testing the designs using a projector and answering to semi-structured interview questions afterwards. The following will explain these steps in detail.

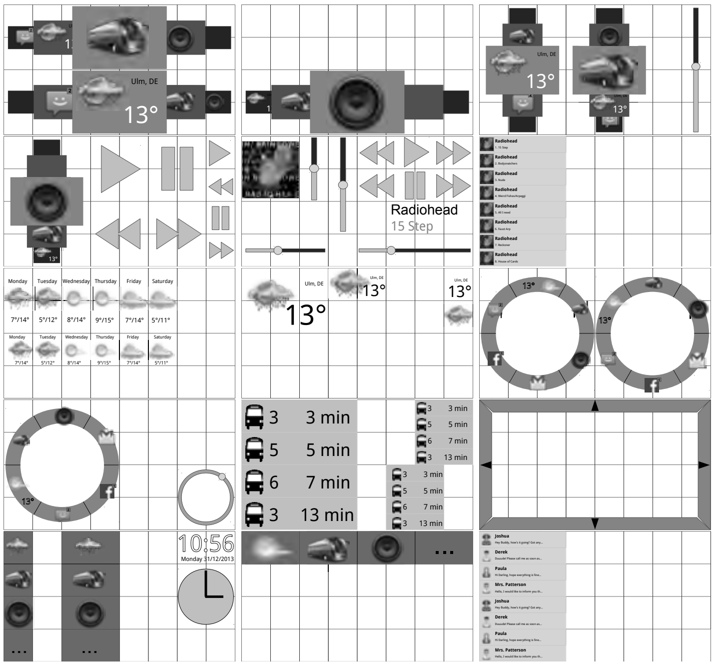


Figure 10.3: Paper widgets participants used to build their UI

10.3.1.1 *Paper Prototyping the SpiderLight UI*

After the introduction, participants used paper prototyping to create UI designs according to their likings. The available screen space was subdivided into a 8× 4 grid to limit the number of elements on the screen. Separate grids for palm and wall/surface projection were created as they induce different constraints regarding size, visibility, and interaction affordances. An overview of the user interface elements the participants could choose from is given by Figure 10.3. This set was inspired by typical UI elements as they appear in mobile interfaces. Thus it included several menu elements such as both horizontal and vertical cover flow, circular shaped menu, side-bar menu, top-bar menu and a menu that has arrows on each edge. The remaining elements were application elements for music player, weather forecast, bus schedule and message reader. Also a clock was provided. The design was to be combined with the interaction concepts depicted by Figure 10.2. Participants were encouraged to put these to creative use and even include custom made elements although none of the participants availed themselves of that offer.

The reason for using paper prototyping was not to provide an already designed interface, because it would have narrowed down the interaction possibilities. For instance, providing a pie menu most probably would have led the participants to use motion as input technique and providing a user interface with a larger number of interaction possibilities would presumably have led them to choose a touch interaction.

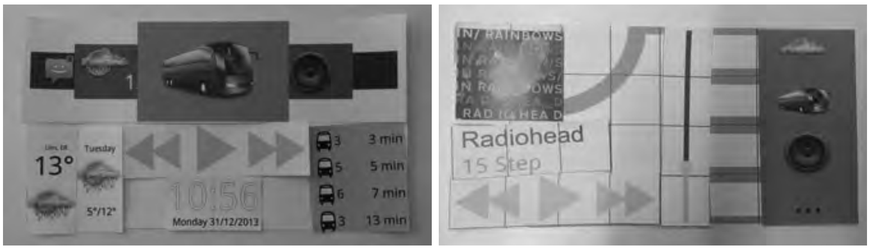


Figure 10.4: Two examples of UI's users finally built for wall (left) and palm (right).

This way, it could not only be evaluated what interaction was preferred but also how much interaction is wanted at all using such a system.

One constraint the participants were given was to use one list element and at least one menu element in their two screens. For instance, if they decided to build a music player they were required to use the play list. This way the study could draw information which interaction method is favored when scrolling.

Participants were instructed to start with the wall design first. After testing their first wall design (see next section), they continued to create a design for their palm, which was then again tested and the whole process repeated as described before. Figure 10.4 shows two examples of UIs created by participants.

10.3.1.2 *Testing and Interview*

After each creation of an UI design, a picture of the design was taken and fed to a Microvision SHOWWX+ HDMI laser projector, velcro-strapped to the participant's wrist. The side on which to wear the projector for the whole study was determined by the arm the user intuitively provided when first handed the device.

Given this approach, participants could directly test both the visibility of their designs as well as reflect on the anticipated usability of the interaction techniques they intended to use. Participants' feedback was recorded using a semi-structured interview after each test round. After the first wall and palm UI concepts had been tested, they performed another round, being able to revise their designs and chosen interactions. After all 4 designs had been tested and individual feedback had been recorded, participants engaged in a final interview regarding their experiences, the differences between palm and wall projection, their preference for an activation gesture, and further use cases they would consider.

10.3.2 Results of User-Elicitation Study

Seven out of nine participants wore the device on their 'strong' hand. The two persons wearing it on the other hand explained that they would like to use the strong hand to interact by tracking the fingers of the other hand. The others mainly stated the reason that they would like to use the fingers on the hand of the device for interaction. Also seven participants chose to wear the device underneath the wrist, while the back of the hand was facing upwards. One participant decided to wear it on top of the back of the hand, similarly to a wristwatch. Another participant wanted to wear it on top of the wrist, but the back of the hand facing downwards. Two participants mentioned that wearing the device sideways would be "more comfortable".

For both wall and palm designs the chosen interaction elements from the designs that we had captured before and the chosen interaction metaphors gathered from the interviews were assessed. The interaction metaphors (cf. Figure 10.2) were divided into scrolling, menu selection, and item selection as atomic interactions required to support earlier mentioned applications.

There was no clear preference between the chosen menu concepts (coverflow, pie menu, top menu), neither for wall nor for palm designs. Regarding wall designs, using the fingers of the primary hand to cast shadows into the projection was the most frequent answer with 3 participants choosing it for menu selection and scrolling and 5 participants selecting it for item selection. What participants liked about this technique was its enabling support for single-handed interaction and its direct feedback. However, for each metaphor one clear competitor became apparent, which in the case of menu selection was rolling the device (3 participants); for scrolling, moving the projector closer or further away from the screen (4 answers); and for item selection direct touching of the projection (3 answers).

Palm designs painted a completely different picture. Here, using the fingers of the secondary hand, on which the primary hand projects, was the most common answer in total, with 5 participants mentioning it for menu selection, 3 participants for scrolling, and 2 for item selection. Again all other answers were mentioned once at most.

When asked, where participants would prefer displaying content, 8 participants preferred output on the wall, next to 5 participants on the hand and only 3 participants on floor or table.

Finally, participants were asked to think about possible problems that might occur when using finger tracking, motion sensors, or a projection in general. Most participants (6) expressed their concern towards masking content with their fingers and the moving (unstable) projection (5) when using motion gestures. Another 3 participants found the finger selection difficult to coordinate and two participants were concerned with random movements that could be falsely interpreted as commands and another two with the interaction becoming exhausting over a longer period of time. Considering projection in general, 5 participants mentioned the low brightness of mobile projectors. When thinking about issues specific to wall or palm projection, almost no problems were reported for wall projection but participants were concerned with the little space available on the palm (7 answers) and the lower legibility of content (5 answers) mainly stemming from the unevenness of the hand (4 answers) and the lower contrast it provides (3 answers).

Regarding the activation gesture, there was not much to learn much beyond the initial ideas, so the Spider unlock gesture as described before was chosen that also inspired the name of the system.

10.4 IMPLEMENTATION

10.4.1 Hardware Considerations

From the previously described study, we drew the following conclusions that motivated the implementation of the SpiderLight:

1. The system should support shadow interaction with fingers of the primary (projecting) hand
2. Roll and translation gestures, which were mentioned second most
3. Both wall and palm projection

Given these requirements, the implemented system must be able to sense finger movements in line of sight of the projection, sense inertial movements, and project preferably with a wide angle not to excessively constrain the minimum distance between projecting hand and opposing palm or wall. In addition, these components were supposed to be part of a single standalone system, with processing power and power supply on-board to support a realistic user experience. As projector it was decided for the Microvision SHOWWX+ HDMI as it was the smallest LBS projector available on the market, providing the widest projection angle, too. The decision for a laser projector seemed inevitably to support quickly changing the projection surface and the projection distance, which would require constant adjustment of the focus using a DLP-based solution—and even then could not provide the dynamic focus range required to project on the uneven human palm.

For the central processing unit different commercially available systemboards like Raspberry Pi, Beaglebone, or Cupieboard and small smartphones that provide video output were considered. However, they all

Figure 10.5: The interior design of the SpiderLight system showing the projector at the bottom, the Android TV stick with the camera mirror on the right, and the battery on the left side. Not visible is the x-IMU which sits behind the projector on the lower side.

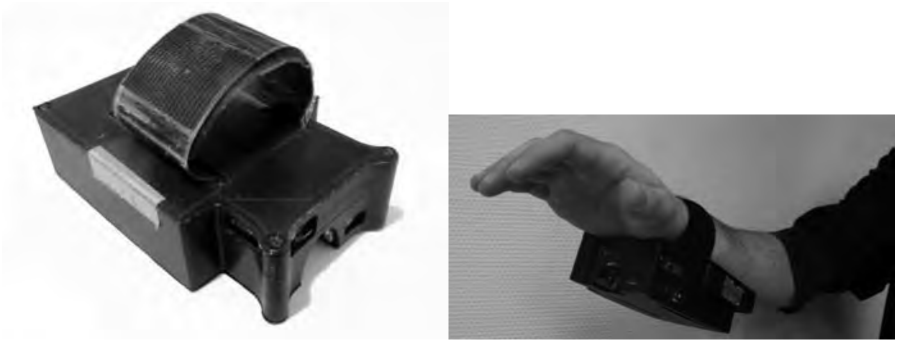


Figure 10.6: The closure of the SpiderLight system (left) and a user wearing it (right).

seemed too bulky by themselves, considering that projector, camera, battery, and potentially additional sensors would all add to the overall size of the system. The decision thus fell on an Android TV stick that would provide the same functionality at a much smaller size. In particular, a system based on the Rockchip GT-S21D was chosen that in addition to HDMI out and USB host—as all TV sticks offer—also provides a camera that is originally meant to be used with teleconferencing. Finding suitable cameras of the desired size that work well together with Android is often a very difficult challenge and by choosing a system that already integrated the camera, the smallest possible footprint of the camera was achieved. However, the decision also implied two consequences: It was decided against a depth camera, which at the time of engineering was not available at the required size and with the required support for mobile platforms like Android. Furthermore, the default placement of the camera required adding a surface mirror to the system to make the camera point in the direction of the projector (more on that in the next section). As the stick did not provide inertial sensors—and inertial sensors of mobile platforms often being not very accurate anyway—we added the x-IMU by x-io Technologies that was already used in AMP-D to the overall setup, which would allow us to accurately measure the device's orientation and translation for pre-warping the projected image against distortion and recognizing rotational device gestures. Finally, a battery supporting two USB ports with at least 1A current output on each port was integrated to power the projector and the TV stick, which in turn powers the x-IMU.

10.4.2 System Integration

All components were fitted into a custom made 3D printed case as shown in figures 10.5 and 10.6. The projector was taken apart to only leave over its PicoP engine and its controller board. This was attached on the bottom side of the case, that is the farthest away from the arm, to leverage the inherent vertical projection angle in contemporary projectors that typically is pointing upwards. As the device is worn on the bottom of the arm—upside down—the projector is mounted the farthest away from the arm and is pointing away from the hand. This way the hand does not occlude the projection but it is still easy for fingers to reach into it. Just behind the projector, the x-IMU is placed and connected to USB power of the TV stick using its GPIO connectors.

The upper side features the TV stick on one and the battery on the other side. The latter can be charged without opening the case. For the camera to point in the right direction, a structure had to be built that allowed to attach the short flexcable of the camera of the Rockchip GT-S21D rotated 90°upwards and counter-clockwise as far as the flexcable allowed. Its image is then again mirrored through a surface-mirror sitting on the opposite side of the camera at an angle of 45°.

Figure 10.6 shows the fully assembled device including the outer shell and carrying band.

10.4.3 Software

The SpiderLight system runs on Android with its UIs created in Java and rendered through OpenGLES. The computer vision algorithms and sensor fusion algorithms are written in C++ and integrated using JNI and Android's NDK interface. Apart from the decisions that were already taken regarding the interaction metaphors, a type of menu interaction had to be chosen from the previous alternatives. Since more users preferred the approach using finger shadows for menu selection, the top menu that was designed with finger shadows in mind and supports absolute pointing (see Figure 10.1c) was selected. Conversely, for scroll selection, rotational device gestures that were answered the most in the user-elicited study were chosen. For item selection, again, selection by finger shadows is employed, whereby the first of four top segments returns to the menu selection and the other 2-3 menu items provide selection commands (cf. Figure 10.1c). The remainder of this section describes the required algorithms for finger detection and gesture sensing.

10.4.3.1 *Finger Detection*

For finger detection, several standard approaches in computer vision were considered, including dynamic background subtraction, skin color segmentation, Haar classifiers, and optical flow detection. Because of the constant jitter of the primary hand dynamic background subtraction was abandoned. Skin color segmentation only works well for light colored skins and is very susceptible to unfavorably lighted environments. Finally, Haar classifiers showed to be not performing very well, maybe due to the limited mobile computing power and the small visible parts of fingers which may not provide enough features. Quite the contrary, optical flow detection on the motion induced by fingers in the image worked reasonably well. The flow is sampled on a grid of 32 × 12 points evenly distributed across the upper half of the region of interest (ROI) as the fingers move vertically starting from the top and would never cross the lower half of the projection.

As optical flow would naturally detect device movements as movement in the image as well, filters were added that would remove optical flows that did not describe vertical finger movements. The first filter subtracts device motion measured by inertial sensors from computed optical flow vectors. The second filtering is done by the criteria that the optical flow vectors require a minimum height which was set to $\frac{1}{6} \cdot ROI_{height}$. Then, by taking the way fingers actually bend towards the palm into consideration and observing bending fingers from behind, a third filter assumes one common vanishing point Q to which all fingers point. Through iterative testing this was eventually be defined to

$$Q_x = ROI_x + \frac{1.5 \cdot ROI_{width}}{3},$$
$$Q_y = ROI_y + 3 \cdot ROI_{height}$$

and all optical flow vectors that do not point to Q—within a certain threshold—are being filtered out. The last filter cancels out any finger interaction during the recognition of rotational gestures (described in the next section).

Finally, the optical flow vectors which are found to be finger movements are grouped by their location into four segments. The segment that contains the most optical flow is selected if at least eight vectors were found in it.

10.4.3.2 *Motion Interaction*

The results of the user-elicited study showed that translational movements were not as much appreciated as, for instance, rotational movements. Furthermore, a rolling rotation keeps the projection at the same place compared to translation movements. Moreover, by counter-rotating the projected image simultaneously, the image can be kept almost stable in the former.

Three applications support roll rotation interaction which are weather forecast, bus schedule, and the music player. Whenever any of these three fragments are active their respective interaction is triggered by a roll rotation that exceeds $\pm 20\,\mathrm{deg}$. In the weather and the bus applications, the rotation controls the scroll view whereas in the music player application a rotation controls the volume.

The geometric correction of the projection follows the procedure described in Subsection 2.4.3 based on the quaternion data received from the x-IMU. That said, it requires a calibration step to define the angle of orthogonal projection which is automatically defined whenever the device is held still for the first time after having been switched on. To further allow the surface to be changed during interaction, another gesture was defined, which by occluding the camera for a short time (e.g. by covering the front with a hand), allows to trigger the calibration manually.

Based on results from the user-elicited study the SpiderLight ideally would implement the spider unlocking technique, bending the hand downwards and pressing two fingers to the ball of the hand. However, the hardware setup was found unsuitable for a reliable introduction of further bending and touch sensors. Instead, shaking—quick rotations in opposite directions—for enabling and disabling the projection was implemented then.

10.4.3.3 *Augmented Reality*

The prototype further supports marker-based augmented reality, for instance to display the nutritional values of food next to its respective markers. By receiving the position and the size of the detected markers, the respective information can be positioned and scaled accordingly to always appear correctly aligned next to the targeted object.

10.5 USABILITY STUDY

To evaluate the performance and usability of SpiderLight a further user study was conducted using the actual prototype. 12 participants (6 female) were recruited which were all right handed (since the prototype was optimized for the right hand) with an average age of 26 (range: 21 to 30). Except for two participants, all have had at least 2 years experience in using a smartphone.

The goal of the study was to compare SpiderLight with a current smartphone in terms of access time, and usability in two applications that depict typical daily activities. Furthermore first impressions of participants using SpiderLight should be collected.

The first task was to look up either the current weather or at what time a certain bus is going to the train station. The second task was to scan an AR code and gather certain information (i.e. nutrition facts). Each task was executed twice with a slight modification but stayed the same in terms of complexity (e.g. only the piece of information to look up changed).

10.5.1 Procedure

The study started with the participants being introduced to SpiderLight. Afterwards they had time to practice and explore the system until they felt comfortable. Participants were encouraged to think aloud and give immediate feedback, which was written down. Participants were instructed to stand in front of a white wall and project onto it but without extending the arm to avoid exhaustion. After the introduction participants were using the smartphone and SpiderLight to finish the three tasks (tasks and systems were both counterbalanced). Every task started with taking the phone out of the pocket and unlocking it respectively enabling the projection of the SpiderLight system. Once all tasks were finished, the users were asked to complete several questionnaires about their experiences using SpiderLight.

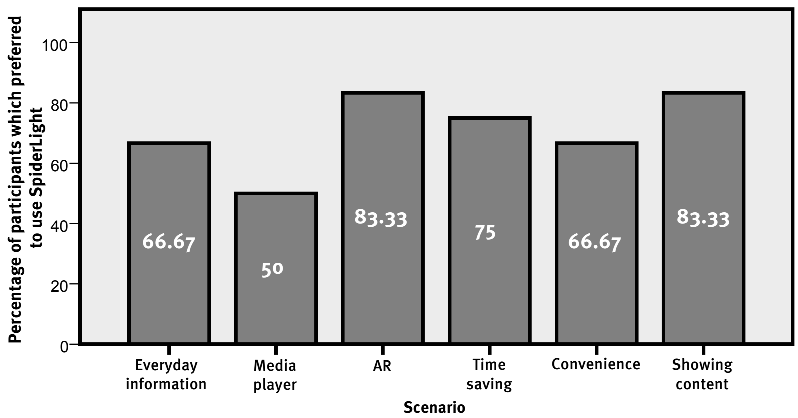


Figure 10.7: Distribution of participants' preference of using SpiderLight for a certain task.

10.5.2 Results

TASK COMPLETION TIME On average it took participants 12.47s ($\sigma = 3.7$) for task one and 19.94s ($\sigma = 9.72$) for task two, using SpiderLight. In comparison to 12.00s ($\sigma = 2.46$) for task one and 14.80s ($\sigma = 3.25$) for task two, using the smartphone. The high standard deviation of SpiderLight in task two seems to be the result of miss detections of input, because the SpiderLight system had sometimes problems in detecting a finger correctly (which was recorded manually during the study). This led to sometimes unusually long interaction using SpiderLight. Nevertheless, looking at results of participants when no miss detection occurred, participants almost exclusively finished the tasks with times below each smartphone time. Therefore, it can be argued that with a more robustly functioning finger detection algorithm, SpiderLight would perform faster compared to smartphones.

QUALITATIVE FEEDBACK In the questionnaires about the usage of SpiderLight, participants reported that rotation interaction was easier to conduct, less physically demanding and had a higher accuracy compared to finger input. This could have partly been influenced by the miss detection of fingers, but also by the fact that using the shadow of a finger to interact with a device was more novel and challenging to participants compared to rolling their arm.

In a last question participants were asked in what scenarios they would prefer to use SpiderLight instead of a smartphone. The answers are depicted by Figure 10.7. Besides the obvious showcasing content to bystanders, which is the central advantage of almost any projector system, also AR support as well as the short lead time to interaction were particularly positively recognized by participants.

10.6 CONCLUSION

This chapter presented a first methodological approach to user interfaces for micro-interactions on palm projected interfaces created by wrist-worn projectors. The presented user-elicitation study led to a set of information types and interaction methods, which users desire for micro-interacting with such a device. As somewhat expected, complex gestures or touch interaction by moving the fingers on the surface side have not been appreciated as much as simple (rotational) gestures and clean UI designs carrying only the most basic information elements. Arguably, if only little information is displayed, the advantage compared to smart watches decreases as they are sufficient in size to show the most basic information at a glance, too. However, four distinct advantages of a projection interface over a smart watch remain:

1. the palm display is significantly larger and has a better resolution;
2. it allows to quickly showcase content to bystanders on a much larger display on a nearby surface;
3. the display cannot be damaged, which is crucial considering that smart watches integrate bio-sensors and are thus regularly used during sports such as playing squash; instead, the projector can be buried within a sturdy closure (does not have to be glass);
4. given better energy performance in the future, it would allow for constant personal projection on the floor, enabling scenarios similar to AMP-D in the previous chapter.

On the side of advantages, smart watches can be looked at using only one hand any time whereas SpiderLight requires a nearby surface or the usage of both hands.

In summary, if the SpiderLight device provided at least the same resolution, brightness, and legibility in the opposite hand as a smart watch, it would constitute a real competitor against smart watches because of its several aforementioned advantages. For people that regularly collaborate on content in nomadic scenarios, it may highly improve convenience without really sacrificing the advantages of a smart watch. Manufacturers of smart watches are currently heavily looking into ways for increasing display size. Lenovo, for instance, recently showed a dual-display smart watch where one display is used like a viewfinder for an even more private display [155]. Instead, a projector combined with a smart watch could provide an additional public display, also enabling some of the collaborations investigated with the SurfacePhone.

The exploration of the prototype and the results of the user studies revealed the requirements of such as system: These include good legibility on the palm, which despite the already applied 2D geometric correction would likely require a correction that takes the complete three-dimensional surface into account [125]. Similarly, radiometric compensation as explained in Subsection 2.4.4.1 would likely benefit the legibility of content. Moreover, a more robust finger tracking is required, be it to recognize finger shadows which were preferred in the user study or touch interaction on the hand. With mobile depth cameras being on the verge of becoming widespread, presumably this will be the least problem in the future as similar functionality has already been commonly shown (cf. Subsection 2.5.3).

The final evaluation could not entirely prove the elimination of all addressed mobile deficiencies by the SpiderLight system. However, we have to take the familiarity of users with smartphones and the described tracking issues during the second user study into account.

MULTI-TASKING In this light, the fact that often users performed much better with SpiderLight than with the smartphone indicates the potential of the SpiderLight concept.

ENVIRONMENT It creates a display that cannot be damaged on a surface that is always available (hand or floor) and where no other display could be attached to providing new input modalities like finger movement on the back or shadow-based interaction (R2) and has the potential to increase the user's awareness (R5) through constant projection on the floor.

> R2
> R5
page 8

This chapter concludes both the part on Nomadic Projection Within Extended Reach for interaction on the go, as well as the case studies presented by this thesis. The remaining part of this thesis will derive concrete guidelines for Mobile Projection Within (Extended) Reach for practitioners in the next chapter and give a thorough prospect on the future role of mobile projection for nomadic interaction before the thesis is concluded.

Part IV

GUIDELINES, CONCLUSION & FUTURE WORK

11 DESIGN GUIDELINES FOR MOBILE PROJECTED INTERFACES

The previous case studies have explored, studied, and proven how projected interfaces, particularly by using Nomadic Projection Within (Extended) Reach, can address existing mobile deficiencies (Subsection 1.1.1) in multiple ways. However, for designers of nomadic computing devices it may not be straightforward to extract the required insights for making the right decisions regarding the question if and how to integrate projected interfaces to their interaction design. Therefore, this chapter will present general guidelines derived from the knowledge gained so far. In particular, these guidelines address the basic questions *when* to apply projection technology at all, *when* to follow the Nomadic Projection Within Reach or Nomadic Projection Within Extended Reach concepts, and *how* to employ these for a given design, including recommendations regarding concept, hard-, and software.

As a note on the side, because these guidelines have to build on the current state of the art in projector technology to be directly applicable, it comes natural that some of these guidelines may lose their relevance as projection technology evolves in the future. Other guidelines describe concepts that are more or less agnostic of the technology and will be applicable to future projection and other display technologies. Despite of that, as other display technologies evolve as well, the underlying considerations of all of these guidelines may hold in the future nonetheless.

Before the presentation of the guidelines it makes sense to briefly recap the four primary deficiencies of current nomadic interaction that these guidelines address and which Subsection 1.1.1 has elaborated in detail:

Output/Input Size Deficiency (D1)

The small *output* size of current nomadic devices hinders overview and quick task completion times. Moreover, the small *input* space causes the fat-finger problem.

Multi-tasking Deficiency (D2)

The small screen space further dictates a single-activity focus to mobile operating systems. This hinders multi-tasking between different apps as well as between several windows within the same app. Another limitation lies in cumbersome switching between real-world tasks (especially those involving the hands) and digital tasks on the device.

Collaboration & Privacy Deficiency (D3)

There is almost no support for synchronous collaboration, whether being co- or remotely located. As such, there is almost no privacy support for sharing either.

Environment Deficiency (D4)

Current nomadic devices are not part of the user's environment but are stowed away. Hence, they cannot easily leverage the user's environment for information display, e.g. by means of AR. Conversely, when they are operated, they disconnect the user of their real environment, leading to security and social issues.

These deficiencies can be addressed by the following derived guidelines. First all guidelines will be named and subsequently discussed in detail. Every detailed guideline will begin with typical questions practitioners are faced with whenever the guideline would apply.

Presented general guidelines for mobile projected interfaces include:

1. When to use projection-based technology
2. When to employ Nomadic Projection Within Reach
3. When to expand this to Nomadic Projection Within Extended Reach

Guidelines pertaining to Nomadic Projection Within Reach are:

4. While maintaining a projection size > 7", prefer a smaller image over a larger one
5. Identify the most common approach angle for the interaction and place the projector on the opposite side to minimize shadow occlusions
6. Position an around-device projection depending upon the purpose of interaction
7. For collaborative use, allow for different physical setups of device(s), users, and spectators
8. Transfer techniques should use animation and involve only one hand
9. Leverage (invisible) optical communication between multiple devices

Guidelines pertaining to Nomadic Projection Within Extended Reach are:

10. Consider privacy and possibly provide privacy-preserving mechanisms
11. Use mediated pointing techniques for object selection
12. Possibly leverage the extended reach for peripheral display

11.1 GENERAL GUIDELINES

11.1.1 When to use projection-based technology (and when not)

- *Which advantages does a projection-based display bring?*
- *Which disadvantages does a projection-based technology impose?*

Let us recall the three traditional main purposes of using projection technology:

1. Create large displays in ad-hoc scenarios where no other form of display is available.
2. Create displays much larger than it would be feasible to build (or transport) using other display technologies.
3. Directly augment the real world and its objects, either manually or by using a projector camera system for tracking and creating a transparent illusion

We have seen that all three arguments apply particularly well to nomadic scenarios where it is not feasible to carry large screens, where large displays cannot be expected to be available in the environment, but where objects and disposable space can be exploited for augmentation to considerably increase display space.

Finally, we have seen that for the nomadic scenario, we can add at least two additional purposes which are

4. Create an *additional* display (e.g. for public/private support) at almost no additional cost for carrying
5. Use optical communication (as has been used for the SurfacePhone and in related works, e.g. [269]) for infrastructure-less communication.

In summary, nomadic computing devices should consider using projection technology when either *additional* or *larger* displays are required. Large displays aid awareness (AMP-D, IPC) and interaction performance (Chapter 4), whereas additional displays typically aid collaboration and privacy (SurfacePhone and IPC) in multi-user scenarios. AR allows to integrate the current environment of the user in ways not possible without projection technology, including paper (Penbook), the floor (AMP-D) or the hand (SpiderLight).

Projection technology should be avoided for nomadic application scenarios or at least pros and cons have to be balanced

1. when the interaction is most likely to happen in direct sunlight, as the amount of Lumens of mobile projectors (100 for pico projectors, around 300 for portable ones) cannot compete with the >30,000 Lumens of direct sunlight indoors or sunny daylight outdoors.
2. when clear and high contrast perception is desired (e.g. a light-table application) because—again due to the constraints on brightness of the image—sufficient requirements to the surface and the environmental lighting cannot be guaranteed or easily modified in nomadic situations.
3. when correct and robust touch recognition is crucial, i.e., touches should neither be missed nor falsely interpreted by the system. If not projecting on a touch-enabled surface, the tight combination of output layer and touch-sensing layer of a screen-based display is superior to ProCamS. The latter usually require computer vision (CV) for the recognition which is less robust due to possible occlusions, shadows induced by the projection, unfavorable changes in lighting conditions, varying skin colors, varying textures of projection surfaces etc. On the other hand, though, it easily supports mid-air gestures which screen displays only achieve by introducing additional cameras.

Provided that the current technological advancement continues at the same rate (cf. Figure 2.3 on page 24), mobile projection technology will not be mature enough for unrestricted indoor usage earlier than 2025 and for realistic outdoor usage (to the same extent as self-illuminated displays can be used outside today) not before 2040. As has been shown, Nomadic Projection Within Reach can be applied to cancel out these disadvantages in certain situations, which leads to the next guideline.

11.1.2 When to employ Nomadic Projection Within Reach (and when not)

- *Can I use Nomadic Projection Within Reach for my design?*
- *Which limitations does Nomadic Projection Within Reach introduce?*

First of all, Nomadic Projection Within Reach especially addresses the requirements of *interactive* projection systems. If content is just to be trivially shown, people know how to cope with it by selecting surface, angle, size, and distance manually and expecting spectators to position themselves accordingly.

All of this is non-trivial for interactive systems. Here, Nomadic Projection Within Reach recommends the flexibility of the additional display over its size and thereby facilitates a bright projection and a familiar touch interaction. Using the concepts of Nomadic Projection Within Reach should be considered when

- the main tasks or a substantial part of the tasks revolve around nomadic information management. The chapter on related works (Chapter 3) has shown a fair amount of works, even for nomadic interaction, that successfully employed other interaction concepts, e.g., handheld usage, to other application domains such as gaming, art, or the military (the bottom/left corner of the taxonomy). In contrast, Nomadic Projection Within Reach is particularly suited to GUI-based information management (as was motivated by Chapter 4 and validated through the corresponding case studies).
- single-user scenarios need more display real estate than the form factor of a mobile device can fit. As free space is often sparse in mobile scenarios, anyway, a small additional projected display with high resolution is more valuable than a large one with low resolution. A smaller display further better preserves the user's privacy as it cannot be seen from farther away and can also better be shielded by the user. As the display is created by the own device, it is safer to use than to rely on display infrastructure existing in the nomadic environment.
- multi-user scenarios need a shared space that is provided by the projection (SurfacePhone, IPC). As projected displays are decoupled from the device and do not inhibit the same notion of ownership, they are more flexible to use in multi-user scenarios.
- multi-user scenarios need a multi-display setup, for instance to support public/private scenarios through a certain setup between the physical and projected displays (SurfacePhone).
- awareness in single- and multi-user setups are to be increased through a larger projection (SurfacePhone, IPC).

Limitations induced by Nomadic Projection Within Reach include that multiple users interacting on a projected display simultaneously, create many shadows which may disturb the experience and which does not scale favorably. Further on, projections are significantly less energy efficient than screens for reasons mentioned before (cf. Subsection 2.4.1).

11.1.3 When to expand this to Nomadic Projection Within Extended Reach

When a short projection distance is not exclusively feasible but Nomadic Projection Within Reach nevertheless promises to be the right interaction concept, the cross-distance space can be covered by allowing more distant objects to be brought into reach and back out of reach again. This setup either requires

- two projectors with different focal lengths to be available;
- one projector with auto-focus and surfaces being available at different distances within the light path (this is the AMP-D setup);
- or one projector for distant objects that can be brought to a nearby display for touch interaction. A smart watch or smart glasses with built-in projector could, for instance, use the projection to indicate objects across a large FOV in the environment to be selected for further interaction on the built-in display.

A suitable interaction technique to select distant objects are pointing gestures either using cursors (AMP-D) or shadows ([78, 176], SpiderLight). The gestures should be robust against accidental movement, though.

Independent of a nearby surface being available or not, it can be intended to position the projected display outside the foveal field of view of the user for peripheral display (AMP-D) and peripheral interaction [W6]. Here, ways of implicit interaction such as the body movement utilized by AMP-D enable peripheral interaction that does not require disconnecting from the current task.

11.2 GUIDELINES FOR NOMADIC PROJECTION WITHIN REACH

When a decision has been made that the interaction should mainly happen *within reach*, the following concrete guidelines solicited from the case studies of this thesis help to achieve a good usability.

11.2.1 While maintaining a projection size > 7", prefer a smaller (brighter) image over a larger one

- *How large should the projection be?*
- *How far should the projection be away? What throw ratio to use?*
- *What is more important, brightness or size?*

Current mobile projectors typically provide not more than a native 720p HD resolution (1280 × 720 pixels). On a 7" screen this leads to a pixel density of roughly 210 pixels per inch (PPI). This already does not meet the quality of color perception of the human eye of about 287 PPI [27] and lies below the screen density of most modern smartphones. Even if resolutions were to climb to native full HD support (1920 × 1080) in the near future, screen sizes like 50" as often advertised by projector manufacturers would still lead to very poor resolutions of not more than 45 PPI. The rapid and wide adoption of retina displays that offer a high resolution instead of a large screen size, demonstrated users' preference for a high pixel density[1]. Considering that larger distances between projector and projection surface do not only result in lower pixel densities (at the benefit of a larger display) but also in a darker image (coming at no benefit) it is desirable to prefer smaller projected images that are brighter and provide a higher PPI. Still, the size should not fall much below 7 inches for current projector generations and 10 inches for future full HD generations as such small projections would lead to PPIs above 287 that cannot be leveraged by the human eye. Furthermore, such small projections are unfamiliar and might be considered awkward by their spectators. The size of the projection should further provide that the user can reach the entire projection area comfortably at arm's length (within a distance of 60 cm) to not withstand direct interaction. Depending on the relation between projector, user, and projection, different throw-ratios are required for the projector or its lens, respectively. For instance, the case studies on Nomadic Projection Within Reach presented before have used projectors with a throw-ratio of 1.10.

With projection sizes of only about 7"-9" diagonal, the Penbook, SurfacePhone, and SpiderLight case studies have shown the value of a second projected display even if of small size. Particularly, not a single of the over 50 participants who took part in the corresponding user studies had complained about a too small projection size—affirming the applicability of this guideline.

11.2.2 Identify the most common approach angle for the interaction and place the projector on the opposite side to minimize shadow occlusions.

- *Where to position the projector in a new device concept?*
- *How to avoid shadow interference?*

During touchscreen interaction, the finger occlusion problem—so-called fat-finger problem—caused by fingers overlaying the display is a well

1 It must be acknowledged that people are very used to consuming media at low PPI, for instance watching full HD video on 50" TV displays, comprising a PPI ratio of not more than 30. However, with the TV market being on the verge of adopting the 4K standard, and considering the fact that TVs are rarely used as replacements for PC monitors (because of too low PPI), PPI ratio seems to matter to users nonetheless

known issue that leads to a measurable decrease in performance and accuracy [39]. Interactions with mobile projections lead to a different occlusion problem when users' interactions cross the light path of the projector, leading to shadows occurring on the projected interface. In distant interaction, this occurs more seldom (or might even be avoided completely) but when it happens shadows are typically very large as they occur close to the projector. In Nomadic Projection Within Reach shadows occur regularly whenever the user interacts with the projection and they occur *in addition* to the finger occlusion problem. As they do not occur close to the projector but instead close to the projection surface, they are comparably small, though. Moreover, shadows are a natural phenomenon that humans are very used to from everyday life and which are already disturbing (and compensated for) in analogue activities such as when writing on paper in direct sunlight or below a desktop lamp.

Based on these experiences, when the projector is placed in a way that the shadow typically occurs below the user's hand, it is not surprising that shadow effects have hardly been recognized, let alone problematized within the case studies. When using the Penbook, for instance, shadows occur typically below the user's hand during writing. Similarly, when using the SurfacePhone, interactions performed by *collaborators*, i.e. users sitting opposed to the owner of the device, cast shadows below the collaborator's hand[2], too. In contrast, interactions by the owner of the device happen from the "wrong" side. However, because of the steep projection angle of the device shadows occur only shortly before the actual touch happens. More importantly, because the user is interacting from a greater distance he or she is not aiming from above their fingertip as in typical scenarios but from below. Therefore, the shadow pointing in the "wrong" direction is actually beneficial. That said, in most scenarios the projector should be placed on the opposite side of the approach angle for interaction. Only if the user is for some reason forced to interact at an arm's length with the projection (as in SurfacePhone), the projector should be *aligned* with the approach angle of interaction if feasible.

11.2.3 Position an around-device projection depending upon the purpose of interaction, whether it is personal, peripheral or collaborative.

- *How to choose position and distance of a projection?*
- *Where to position the projector in a new device concept?*

The projection should be placed in such a relation to the projecting device that it is in accordance to the main purpose of the projection. This

2 if the collaborator interacts on the projection of the other user and users are facing each other

can be one of a personal purpose for active interaction, or a personal purpose for peripheral perception, or for collaboration between multiple users.

Personal interaction requires the projection to be within easy reach of the main user. This typically means that the interaction should be supported directly in front of the user as has been demonstrated in the Penbook and IPC case studies.

Another personal use case is *peripheral interaction*. Here, the focus lies rather on perception than interaction. Not to disturb the user in their main routines the display should be placed behind (as in the SurfacePhone) or to the sides of the device (cf. [133]). The prototype by Qin et al. [204] can also be regarded as a low-resolution projector and augments the device with an aura of dynamic ambient lighting directly *below* the device. The ceiling projection by Leung et al. [157] wants the projection to work as peripheral display for bystanders, for which—indoors—the ceiling seems most appropriate.

Finally, if the anticipated main purpose of the projection is to support collaboration, the projection should be placed in a way that it is most convenient to operate by *other* users. For the SurfacePhone this meant that the projection was to be placed *behind* the device. Although this position makes the operation more inconvenient for the main user, other users are naturally invited to perceive and operate the projection as it is physically closer to them than to the main user. This should go as far as to align and rotate the content upside down to appear correct for bystanders rather than the main user. In most cases, the main user will know the content anyway and will only require the perception of the projection for augmenting verbal communication by non-verbal reference. Further on, the projection behind the device still allows for a very good perception by the main user but does not interfere with interacting with the projecting device for controlling the content on the projection. It was also shown that the AMP-D prototype can similarly be extended to multi-user interaction by allowing to merge the floor projections. Analogously, content on the projection should be aligned to face the collaborating user.

As the study on touch locations has shown (see Figure 7.5.3.2), wherever the projection is placed, depending on the distance and approach angle to the projection, users will tend to undershoot or overshoot the target, respectively. Therefore, if the system knows about the user's posture towards the projection (by some form of inside-out tracking for example), a corresponding offset function should be applied to the user's touch. If such knowledge is unavailable, the most common posture of interaction should be applied.

11.2.4 For collaborative use, allow for different physical setups of device(s), users, and spectators to support different levels of privacy and intimacy.

- *Will users know by themselves how to use and place the device(s)?*
- *Which collaborative use cases do specific projection setups support?*

Most devices dictate to a certain extent how they want to be used, e.g., TVs dictate a certain minimum distance for viewing by their size. Similarly, the size and form factor of mobile devices, in particular mobile phones, communicate handheld usage. Projections lack such affordances and thus require the product designer to communicate intended use else-wise, for instance through the position of the projection as explained before. The same lack of affordance on the other hand allows for very flexible and adaptable usage. Because projected displays can easily be moved and resized (by changing distance and angle) multiple projections can be merged and overlaid in an incomparably easy manner. Together with different positioning of users, different levels of privacy, collaboration, and even intimacy can be supported. Some of them have been leveraged in the SurfacePhone project, ranging from distant scenarios with users sitting on the opposite of each other, to collaborative scenarios with users sitting next to each other. In a broader sense, two perspectives on support for different postures have to be distinguished: on the one hand, the software should support and adapt to users taking different positions and angles during interaction to support different types of tasks, for instance ranging from selective picture presentation to collaborative puzzling as in SurfacePhone. And projected games can use these flexible transitions for attack and defense maneuvers as in SideBySide [269]. On the other hand, if the projection device was motorized, it could autonomously move the projection and its orientation and thereby force users to either take intimacy enforcing or intimacy disturbing positions to each other. We have used a similar concept for the autonomous device movement in the knight game of the HoverPad [W14].

11.2.5 Transfer techniques should use animation and only involve one hand if possible

- *Which human factors have to be considered when designing transfer techniques?*
- *What has to be considered regarding the visual style of transfer techniques?*

If displays are visually separated, i.e. not directly merged as the shared space of the SurfacePhone, animation should be used to visualize the transfer of content. The *proxy* technique of the SurfacePhonelet new content slide in from the side or animated content that already existed on the display on top of the existing item. Conversely, the *swipe* and *human link* techniques did not use animation but just removed the content from one display and added it to the other display. This sometimes confused users because they missed a transfer action the system had performed (on their own request or that of another user). Short animations help to increase awareness and presumably lead to less missed actions. In scenarios that include an *extended* reach like AMP-D this may involve the (three-dimensional) interpolation between different distances, which a physics engine is usually capable to provide.

As nomadic devices will usually be small and lightweight, they will likely not provide the necessary weight to stand robustly on a table. For devices that are used put down on a surface, one-handed interaction should thus be used as bi-manual interactions lead to accidental movements of the device, leading to inaccuracies.

11.2.6 Leverage (invisible) optical communication between multiple devices, possibly to set up a wireless communication channel.

- *Which special properties of projection technology must be considered during product and interaction design?*

When a device allows or requires device to device interaction between multiple projection devices or between a projection device and a public display for example, optical communication can be superior to ordinary communication channels like Bluetooth or Wi-Fi. Especially during the initial setup, bringing multiple devices into the same wireless network is often a tedious and error-prone task which repels users right away. Optical communication, particularly in the invisible infrared band, does not require any prior setup. It does assume cameras facing in the direction of the projection, but which projector systems typically require to recognize interaction anyway. In [269] Willis et al. have shown how both position and actions can be communicated solely via QR codes projected through infrared light alongside the visible content. Virolainen et al. [258] used the optical channel to transmit pictures. With the SurfacePhone it was shown how multiple devices can recognize each other's projections in the visible light space using natural feature image detection algorithms to support merging of multiple projections to a single combined one. As the optical bandwidth is more limited than the radio ones—mainly because of the much more limited update and frame rates of projector and camera—more complex communications must still rely on radio channels. But even then the optical communication can be used to communicate and negotiate connection parameters between multiple devices to facilitate an automated networking setup that is transparent to the user and therefore particularly supports the nomadic computing principles (Chapter 1).

11.3 GUIDELINES FOR NOMADIC PROJECTION WITHIN EXTENDED REACH AND ON THE GO

Analogously to the previous section, this section provided concrete guidelines for interacting at *extended reach*.

11.3.1 Consider privacy and possibly provide privacy-preserving mechanisms

- *Which privacy implications will a public projection for personal information management have?*

With the extended distance comes a bigger publicity of the projection. Different to some of the Nomadic Projection Within Reach scenarios like Penbook and IPC, any notion of privacy of the projection, at least at the extended distance, is given up. All the more this is the case in on-the-go scenarios. Thus private information on the projection within *extended* reach should probably be avoided and only disclosed when brought *within reach*. Furthermore, providing a means of instantly enabling/disabling the projection (such as AMP-D and SpiderLight have provided) enhances the user's sense and control of privacy.

11.3.2 Use mediated pointing techniques for object selection

- *How to interact with content out-of-reach?*

As touch interaction is not available at extended reach, alternatives have to be used. A suitable interaction technique to select distant objects are pointing gestures either using cursors (AMP-D) or shadows ([78, 176], SpiderLight). In on-the-go scenarios, these gestures must be robust against accidental movement, though. As found out by the case study on AMP-D, posture-based gestures work more reliably during movement to de- and activate selection out of reach as well as to switch between different modes. If possible, uni-manual gestures are to be preferred to not diminish one of the usual advantages of projected interfaces, namely hands-free operation.

11.3.3 Possibly leverage the extended reach for peripheral display

- *Which unique properties of projections can be leveraged?*

The extended distance allows to position the projected display in the visual periphery (like the floor display of AMP-D) and support peripheral interaction [W6]. Different to traditional peripheral displays, be they screen-based or projection-based and located within a room, the nomadic projection scenario requires some adaptations to consider:

1. Because the environmental light level can change any time, so does the contrast of the projected display. Pertaining to nomadic projection, however, the change may go unnoticed as the projection is not constantly perceived by the user. On the other hand, different amounts of animation on the peripheral display allow to control the amount of attention drawn from the user. Hence, if possible, the system should measure changing ambient light levels and adapt the amount of animation accordingly.
2. In nomadic and especially on-the-go scenarios, the user will be much more distracted than in traditional environments of peripheral displays. Thus, if possible, the system should track if and when the user has looked at the projection—or at least moved the head towards it—to adapt time and frequency of updating the display with (new) content.
3. Ways of implicit interaction such as the body movement utilized by AMP-D allow peripheral interaction that does not require the user to disconnect from the current task, which can be critical such as walking to destination in time.

As all technology, mobile projection has its limitations—limitations that despite good design guidelines may not be easily eliminated. A topic that has oftentimes turned up in paper reviews and discussions with colleagues was if and how other emerging display technologies like smart glasses and watches already outperform mobile projection or might do so in the future. Hence, before the thesis is concluded, the next chapter on future work will discuss this topic in detail. On the other hand, it will, of course, also discuss possible future improvements to Nomadic Projection Within (Extended) Reach and mobile projection.

12

THE FUTURE OF NOMADIC COMPUTING

This chapter will name some possible future improvements for Nomadic Projection Within (Extended) Reach and mobile projection in the next section. However, with many new wearable devices for nomadic computing appearing on the market right now, the future of mobile projection may equally lie in a meaningful combination with and between these than in an isolated advancement of projection technology. This perspective will be discussed by the subsequent section.

12.1 IMPROVING NOMADIC PROJECTION WITHIN (EXTENDED) REACH

Obvious are necessary requirements to advance Nomadic Projection Within (Extended) Reach in the future:

- The energy efficiency and performance of battery-powered projectors has to double or better triple to allow very small projectors to be integrated into standard mobile devices, to achieve a form factor that is considerably smaller than that of the prototypes presented in this thesis. To be usable indoors in direct sunlight or outdoors outside of direct sunlight, rather improvements in energy efficiency in the range of 5 to 10 times have to be achieved (only speaking of within (extended) reach distances).
- Flexibility in the projection setup. This thesis has presented several prototypes, each of which had been tailored to the specific use case regarding placement and orientation of the projector. But users, most likely, will not be willing to carry several devices for nomadic computing which would defeat its purpose. However, mechanical solutions that allow between different device setups are not very difficult to support. A simple means would allow to turn projector and camera 180°(see Figure 12.1a) and thereby support personal projection scenarios like the Penbook (components turned to the user) and collaborative ones like the SurfacePhone (components turned in order to face away from the user). Another approach that simplifies rotating camera and projector simultaneously, recently has been presented by Lenovo with the “Smart Cast” concept and a real prototype (Figure 12.1b)

(a) Rotatable cameras and projectors on smartphones would allow for different interaction setups using the same device. It would also render the additional front-facing camera of contemporary smartphones unnecessary.

(b) The Lenovo Smart Cast concept. A rotatable mirror at the top allows to flip the direction between forward and backward of camera and projector at once (does not allow arbitrary rotations, though). Image courtesy of Lenovo.

Figure 12.1: Solutions for serving different projection setups from the same device.

- Mobile devices with integrated depth cameras with a very good depth accuracy, especially in the near range between 10 and 50 cm. These will drastically improve the accuracy of touch detection of Nomadic Projection Within Reach and enable robust gesture detection for Nomadic Projection Within Extended Reach without sacrificing the mobile form factor.

12.2 THE FUTURE ROLE OF MOBILE PROJECTION IN NOMADIC COMPUTING

Taking a broader perspective, what has been tried to achieve in this thesis is to take the strengths of one mobile technology (mobile projection) to address the weaknesses of another mobile technology (e.g., smartphones and tablets) and vice versa (the private nature of the handheld screen has surmounted the very public nature of the projection). With mobile projection, we have successfully addressed four deficiencies of smartphones for nomadic computing, but also witnessed some of the limitations of mobile projection, foremost its unfavorable competition with ambient light. As Section 2.2 has explained, other nomadic technologies are less susceptible to ambient light and currently we witness their proliferation in form of all sorts of wearable devices such as fitbands and smart-rings, -watches, -shoes, -glasses, and -clothes. In a broader sense, all of these smart nomadic computing devices (see Figure 12.2) try to address each others' deficiencies (most notably those of smart phones) within the nomadic design space using their individual strengths and in spite of their own weaknesses.

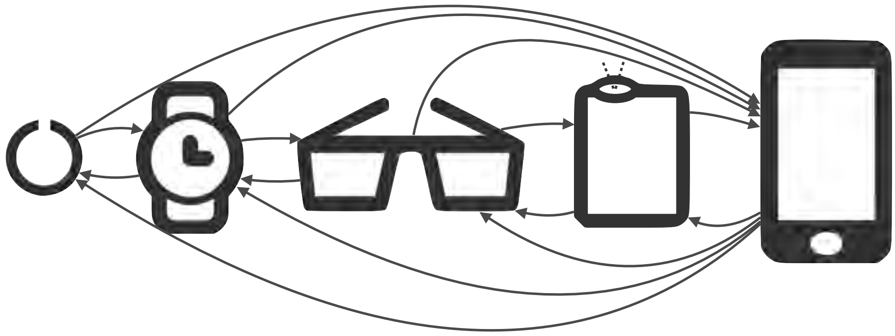


Figure 12.2: The device space of contemporary nomadic computing: smart ring, watch, glasses, projector, and phone. Each are able to address specific weaknesses of the smartphone but all the same between each other (some possible combinations are left out for clarity).

The prospect of this thesis should therefore take a closer look at a possible future role of mobile projection within this design space. As one example, smart glasses are pushing the output part of Nomadic Projection Within Reach to its extreme by putting the display extremely close to the user's eye where only very few ambient light comes in the way. As such, they are more robust towards brightly lit environments and also support an unmatched privacy for displayed content. At the same time, compared to mobile projection, interacting with information displayed on the glasses is way more difficult since no touch or pointing interaction is supported by the system. This is just one example of what seems to be a large new design space between wearable nomadic computing devices. The rest of this chapter will present a first proposal—to be extended by future research—of this design space by

1. presenting an overview of the unique deficiencies and strengths of contemporary smart wearables Subsection 12.2.1;
2. identifying two alternative approaches how multiple devices can be combined to surmount each other's deficiencies to aid the user in nomadic computing; both approaches will be exemplified by case studies, which already started research into that direction.
   - In the first approach, one device surmounts the deficiencies of another primary device, which aids the user in nomadic computing. The Glass Unlock case study in Subsection 12.2.2 depicts this type of approach.
   - In the second a framework provides seamless cooperation within the ecosystem to allow all available smart devices to achieve a common goal of the user. The Display Copy system in Subsection 12.2.3 briefly exemplifies this type of approach.

12.2.1 Deficiencies and Strengths of nomadic computing devices

The following two pages present a list of deficiencies and strengths (Table 12.1) of the currently most widespread nomadic computing devices and based on that, a device graph (Figure 12.3) that depicts opportunities for addressing deficiencies in nomadic computing by combining multiple devices.

These classifications are based on the personal experience of the author with these technologies that was gained during and after the conducted research on Nomadic Projection Within (Extended) Reach. It must be noted that the currently almost non-existent social acceptance of smart-glasses (as has been exemplified by Google Glass) and the dependability of mobile projection on ambient light did not receive a very large weight as to further decrease their negative score. The social acceptance of smart glasses may change quickly (such as the acceptance of headsets did) and Nomadic Projection Within Extended Reach has presented a concept how mobile projection is already usable today. Apart from that, these classifications are neither meant to be exhaustive nor objective and naturally require a formal evaluation to be conducted in the future. Nonetheless, they may already inform designers of future nomadic computing applications that consider combining several devices. The major purpose of these classifications, however, is to reflect upon the possible future role of mobile projection within this future ecosystem. Based on Table 12.1 and Figure 12.3 we can derive the following conclusions:

1. Mobile projection has a high nomadic suitability, similar to that of smart glasses and higher than other smart wearables. Except for smart watches and phones, that are already in the market, rings, glasses, and mobile projection require further technological advancement to decrease their size and thus *increase* their social acceptability to allow them to spread out in the end-consumer market.

2. Whereas other device categories share more similarities with each other, mobile projection distinguishes itself through several unique advantages but also disadvantages. Advantages include an unmatched support for creating large displays, for supporting collaboration, and constituting the only display technology that can easily be integrated to other nomadic devices. This not only allows the integration to phones and tablets as presented in this thesis, but to glasses, watches, rings and other nomadic computing devices in the future. The biggest disadvantages compared to other systems include the dependency on an available projection surface and the rivalry with ambient light as has already been discussed throughout this thesis.

As has been explained before, the remainder of this chapter will present two case studies (Glass Unlock and Display Copy) on two alternative approaches how nomadic devices can be combined.

			Range						Sum (CES)
Output	*Visibility*	**Robustness ambient light**	not robust (-2) to robust (2)	2	2	1	-2	2	5
		IO Coupling	possible (2) or not (0)	0	0	0	2	0	2
	Visual information	**Lead time**	long (-2) to short (2)	2	1	2	0	-2	3
		Display size	not there (0) to large (4)	0	1	2	4	2	9
		Fidelity/Resolution	low (0) to high (4)	0	1	3	3	2	9
		Collaboration	no (0) to high support (4)	0	0	1	4	2	7
		Privacy	no (0) to high support (4)	0	0	4	0	1	5
		Hands-free	possible (2) or not (0)	0	2	2	2	0	6
	Audio	**Distance/Publicity**	long/public (0) to short/private (4)	0	0	4	0	2	6
	Awareness (e.g. robustness notifications)	**Visual (periphery)**	no (0) to robust/peripheral (4)	0	1	3	4	0	8
		Audio	no (0) to robust/peripheral (3)	1	2	3	1	2	9
		Vibration	no (0) to robust/peripheral (2)	2	2	0	0	1	5
	AR	**Range**	close (0) to far (2)	0	0	2	0	2	4
		Independant from surface	possible (2) or not (0)	0	0	2	0	2	4
		Head-independent view	possible (2) or not (0)	0	0	0	2	2	4
		Floating support	possible (2) or not (0)	0	0	2	0	2	4
		Behind touch	possible (2) or not (0)	0	0	0	2	0	2
		3D depth support	possible (2) or not (0)	0	0	2	2	0	4
	Social acceptance	**Acceptance wearing**	low (-2) to high (2)	2	1	-1	-2	2	2
		Acceptance creating output	low (-2) to high (2)	2	1	0	-2	1	2
Input	*Modalities*	**Pointing/Touch**	no (0) to hight support (4)	0	2	0	3	4	9
		Gestures	no (0) to hight support (4)	2	3	3	3	3	14
		Speech	no (0) to hight support (2)	0	1	2	1	2	6
	Fidelity	**Space**	small (0) to high (2)	0	1	1	2	2	6
		Text input	small (0) to high (2)	0	1	0	1	2	4
	Body involvement	**Hands-free**	possible (2) or not (0)	0	0	2	1	0	3
		Eyes-free	possible (2) or not (0)	2	1	0	1	0	4
	Social acceptance	**Subtlety**	low (-2) to high (2)	2	0	-1	-1	0	0
Capabilities		**Run-time**	short (-2) to long (2)	0	1	-1	-2	2	0
		Connectivity	no (0) to high (4)	1	2	2	0	4	9
		Flexibility in positioning	no (0) to support (4)	1	0	0	4	2	7
		Integration into other devices	possible (2) or not (0)	0	0	0	2	0	2
			Sum (Nomadic Suitability Score)	19	26	40	35	44	

Table 12.1: This table lists strengths and weaknesses of current nomadic computing devices based on their state of the art. It does this by allocating points to each device across 32 aspects. The sum scores of each device depict "Nomadic Suitability" scores and are designed in a way that multiple devices can be accumulated to increase the overall suitability score. This should be done on an individual aspect level, though, clipping the accumulated score to the range of the aspect. This support required the score range of each aspect to be chosen individually based on two criteria: (1) if the addition of a device may have a negative impact on another device within the same aspect and therefore should start in the negative range (example: if one device has a significantly shorter lead time to interaction than the other one, their combined score should be worse). Conversely, two displays provide always more space than one and should not negatively impact each other. (2) A weight is introduced by making some ranges have a smaller stretch and/or maximum.

Scores are based on the personal experience of the author with all of these technologies, also considering related works and their state of the art. Hence, so far, they only serve as a proposal to spur further discussion and naturally require a formal evaluation in the future.

CES: Coverage of the aspect by the whole ecosystem of smart nomadic computing devices.

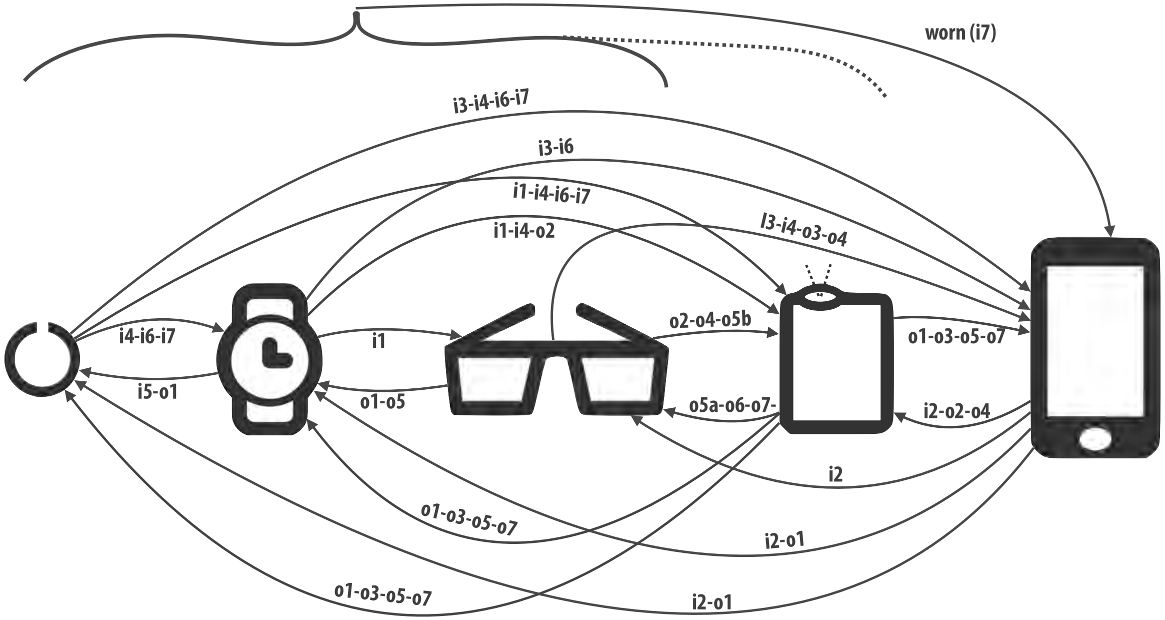


Figure 12.3: Nomadic device collaboration: A non-exhaustive graph of how nomadic devices can address each others' weaknesses and strengths. Combinations with a projector are highlighted in red.

Input
i1 Additional
i2 Superior
i3 Quick (micro)
i4 Subtle
i5 Hands-free
i6 Eyes-free
i7 Token

Output
o1 Larger
o2 Superior (contrast)
o3 Awareness
o4 Privacy
o5 AR
o5a behind
o5b before
o6 Publicity
o7 Collaboration

12.2.2 Combining smart phone and the very private display of smart glasses for a very secure smartphone unlock mechanism

This section is based on the previously published refereed conference paper:

[W2] **Winkler, C.**, Gugenheimer, J., De Luca, A., Haas, G., Speidel, P., Dobbelstein, D., Rukzio, E., "Glass Unlock: Enhancing Security of Smartphone Unlocking Through Leveraging a Private Near-eye Display." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, 2015, pp. 1407–1410

In addition, the following partially related thesis was supervised by the author:

- "Smart, Smarter, Smartest? An Exploration of the Design Space and Development of Interactions Between Multiple Smart Gadgets". Philipp Speidel. Bachelor's thesis. 2014

12.2.2.1 *Introduction*

Related video

Recent findings suggest that about 43% of smartphone users rely on some form of lock-screen to protect their phone from unwanted usage [102]. However, currently deployed smartphone authentication mechanisms like PIN and the Android unlock pattern are susceptible to different real world attacks such as smudge attacks [33], shoulder-surfing [82], or camera attacks. Especially the latter is becoming more and more of a threat with the increasing prevalence of video surveillance.

One way of protecting authentication from these attacks is to use biometric properties like fingerprints or input behavior [58]. While these are highly usable alternatives, they suffer from trust issues and the fact that they make the devices hard or impossible to share [77]. Indirect input or other kinds of software distractions [138, 140] suffer from highly reduced authentication speed and thus, negatively influence usability. As opposed to this, hardware based approaches rely on additional, external devices to provide invisible channels to the user which affect the input [47] or relocate the input to a less observable position [82]. While increasing usability, they require additional devices to be carried around.

With the advent of smart wearable devices such as smart watches and smart glasses on the consumer market, such devices are not an additional burden anymore as they are carried around anyway as part of the users' daily lives. We already see that they can be used to enhance the usability of lock-screens. For instance, Google's Android now offers to automatically disable the lock screen whenever the user's smart watch is in the near vicinity, meaningfully combining the devices and increasing the convenience of the user. However, it is only appropriate for less concerned users, as it enables new types of attacks like stealing

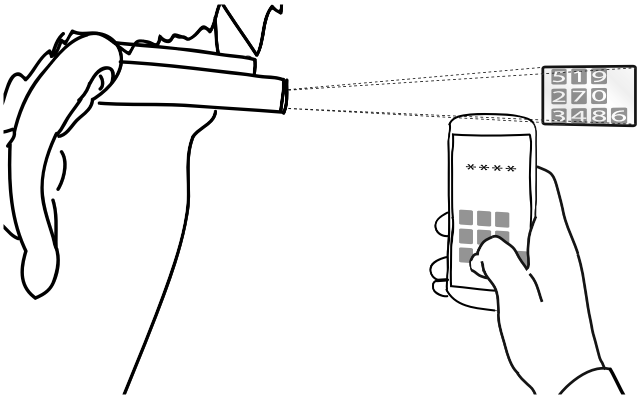


Figure 12.4: Glass Unlock concept: the scrambled PIN pad is only shown on the display of the user's smart glasses; input is performed on empty buttons on the smartphone, which does not give anything away to an attacker.

both devices together or leveraging moments when the phone is left unattended while still in range of the watch.

Glass Unlock introduces a similar approach for phone unlocking, combining smart glasses and their advantage of a very private near-eye display with the phone's lock screen. The basic idea is to hide the lock information (e.g. PIN digits) on the phone and instead show it on the glasses' display. For instance, in a standard 10-digits PIN screen the phone would show *empty* buttons while the same layout including the digits would be visible on the glasses as shown in Figure 12.4. The random order of digits is required to achieve the desired security as explained later. By precluding any attackers of making sense of the users' input on the phone, Glass Unlock is secure against smudge attacks, shoulder surfing, and camera attacks.

Interesting to investigate are the additional costs of this approach compared to the state-of-the-art of unlocking. According to Harbach et al. [102], this is PIN unlocking, which about a third of all smartphone users (78% of all lock screen users) rely upon. Besides the analogue 4 out of 10 digits implementation, two further alternative variations of Glass Unlock have been evaluated: one that proved to decrease the visual search time by reducing the number of digits from 10 to 6 (called 6Key); another that proved to support eyes-free input on the phone by requiring swipes instead of touches, thus removing any need to switch focus between the phone and the display of the glasses (called swipe).

12.2.2.2 *Glass Unlock Concept*

As people owning smart glasses will likely wear them most of the time, it makes sense to combine them with the people's phones to increase their security. While the whole phone unlock could be performed on

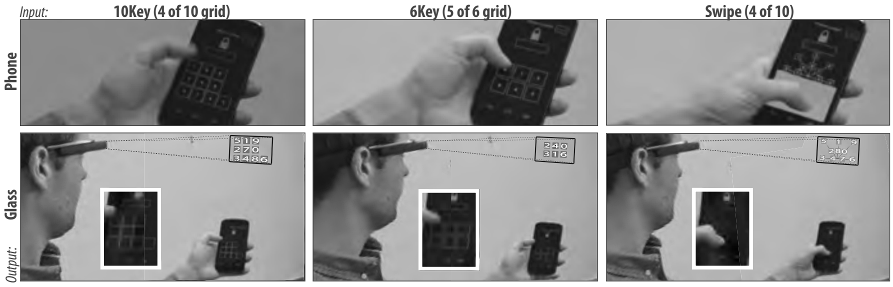


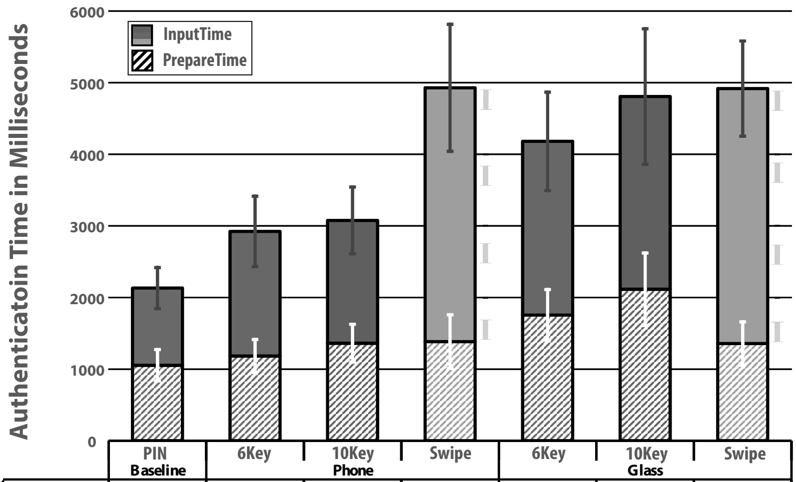
Figure 12.5: The 6 study systems: 3 input and 2 output methods (with and without Glass), additionally compared against standard PIN baseline (not shown).

the glasses alone, users should not be forced to switch input to another device while they interact with their phone. Hence, the Glass Unlock concept moves the authentication challenge to the near-eye display of the glasses (leveraging its ability to deliver the output in a very private manner) while retaining authentication input on the phone (leveraging the superiority of the phone regarding input). When the Glass is not available Glass Unlock gracefully degrades to scrambled PIN-entry with visible numbers on the phone.

By moving the authentication challenge from the (public) phone display to the (private) near-eye display, neither shoulder-surfing nor multiple synchronized camera observations give away the password simply because it is not shown on the phone. Small digits on the near-eye display are not visible to onlookers and cameras. In addition, Glass Unlock scrambles the order of digits after every successful unlock attempt, thus preventing attackers of merely repeating observed input on the phone, which also makes it resistant against smudge attacks.

As an attacker of Glass Unlock still has to acquire knowledge of the password, even stealing both devices together will not facilitate phone unlocking any more easily than without the glasses—this is a huge difference to previously mentioned automatic unlocking between multiple smart devices. Glass Unlock further assumes a secure Bluetooth connection between the devices, but even if the connection was compromised, an attacker of Glass Unlock would have to simultaneously record the digital transmission and observe the input on the phone. This is because no sensitive information is transmitted, only the randomized PIN layout.

It is important to note that the general Glass Unlock concept does not only relate to smartphone unlocking. People are required to enter secrets all the time, at the ATM, when paying with debit cards, etc. We can envision a general framework that would automatically transfer the challenge to the user's smart glasses.



	PIN (Baseline)	6Key (Phone)	10Key (Phone)	Swipe (Phone)	6Key (Glass)	10Key (Glass)	Swipe (Glass)
InputTime	1078	1740	1715	3544	2426	2691	3560
PrepareTime	1053	1182	1362	1384	1755	2115	1357

Figure 12.6: Authentication times divided into preparation time (time until first touch) and input time (remaining time). Swipe input time also shows— next to it—input times per digit and preparation time in-between.

12.2.2.3 *Study Results*

The description of the complete study setup, procedure, statistical analyses as well as their results can be found in [W2]. Figure 12.6 depicts the mean preparation (time until the first touch), input (rest of the time), and total times (preparation + input) of the 18 study participants using the seven compared systems. These comprised the standard unlock baseline and the new systems consisting of the three input techniques: *10Key*, *6Key*, and *Swipe*; and two output methods: *Phone* and *Glass* as independent variables (see Figure 12.5).

AUTHENTICATION SPEED The total time of the baseline was lower than total times of other phone methods. This is as expected since all other systems used a scrambled PIN pad that introduced a visual search task. The 6Key input method successfully reduced visual search time as it significantly decreased the preparation time compared to 10Key. Furthermore, we see that the preparation times of 6Key and 10Key significantly increase when used with Glass, but Swipe remains almost the same. This can be attributed to Swipe's support for eyes-free input when used with the Glass. In contrast, 6Key and 10Key require users to perform a mapping to the phone once they switched their focus. This leads to a higher preparation time. Also, to minimize attention shifts, users may have tried to find and remember multiple positions from the very beginning. Swipe on the other hand allowed the input to start as soon as the first digit was discovered.

Preparation time (search time) does not only happen before the first touch, but also between touches/swipes during the input time. This explains the significant rise in input times between phone and Glass methods. Again, like with preparation time, the times of 6Key and 10Key increase significantly when used with Glass as display switches occur during the input as well. Very interesting are the high input times of Swipe. They remain almost exactly the same between output methods, which gives strong evidence that Swipe supported eyes-free input, thus was not confounded by the separation of displays. However, swiping takes longer to perform and the unusual layout may have introduced a small disadvantage as well.

AUTHENTICATION ERRORS Errors were very low across all key input systems (overall 7 errors) and thus only the *Swipe* errors with and without Glass are worthwhile to discuss. Most errors occurred in the length (29 errors) of the swipe—too short or too long—or the angle (14 errors)—left or right slip. Using the Glass, participants produced more errors (19) in the length of the input than without (10). This can be attributed to the eyes-free input as the works of De Luca et al. [83] already revealed that users struggle with swipe input more when performed eyes-free. Surprisingly enough, introducing the Glass did not lead to any more errors with the key input methods, despite the required switching and the possible out-of-focus touching.

QUALITATIVE SYSTEM FEEDBACK Semi-structured interviews revealed users' experiences to be mainly in line with the quantitative results ([W2] provides some more details).

More interesting have been users comments to openly asked questions. Regarding 10Key, 13 participants criticized the annoying display switches while regarding Swipe, 12 users explicitly mentioned to cherish canceling out of display switches. On the negative side, 4 reported problems with distinguishing between short and long swipes, 3 found short swipes harder to perform than long swipes, and 3 found Swipe too slow in general. Interestingly, in the final ranking of the three input methods by output method (Figure 12.7) participants shifted their sympathy nonetheless even more towards the Swipe technique when used with Glass, followed by 6Key gaining only half the sympathy on rank 1. Thus, display switches seem to be a very annoying factor in this new type of multi-display system and users would rather choose a slower input technique but which is less demanding on the eye. Finally, ≈65% of participants stated they would entirely replace their current lock screen with their favorite Glass Unlock variant if they owned compatible glasses and additional ≈18% would do so only for security critical apps.

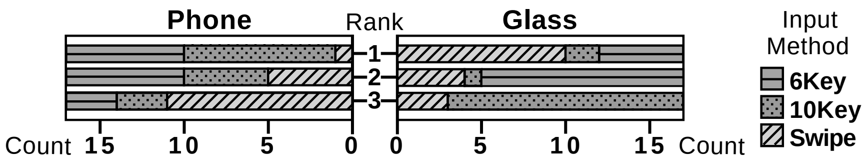


Figure 12.7: Final ranking of participants' preferred input method.

12.2.2.4 *Conclusion & Future Work*

Users spend much of their time on unlocking their phones. With mobile devices becoming more and more a medium for highly sensitive data, secure unlock methods are researched that yield acceptable input times without requiring additional hardware to carry. Glass Unlock depicts an authentication system, that is inherently secure against the most common visual attacks against mobile phone locks while increasing the unlock time only moderately. Glass Unlock achieves this by outsourcing the security critical output to the private near-eye display, which is believed to become a regular companion of many smartphone users, thus adding no additional hardware requirements. Thereby Glass Unlock leverages the advantages of the glass (private display) and the phone (accurate touch recognition) to unlock procedure, which is especially annoying in nomadic scenarios because it is more insecure and more difficult to perform while on-the-go.

However, we have seen that smart devices cannot be "just" combined and expected to work optimally together. Issues like visual separation of devices and displays leading to context-switches or reduced awareness, or the complexity of multi-modal input, show that the combination of devices may lead to new issues that require to be addressed by the interaction concept. In the case of Glass Unlock the biggest problem users faced seemed to be the required context switches between the displays, as well as the visual search task. These have been successfully addressed by the Swipe and 6Key techniques and ultimately should probably result in a combination of both as the right interaction concept for Glass Unlock.

Two other examples for surmounting the deficiencies between two devices are recent works like Chen et al. [73] (watch and phone) and the upcoming Benko et al. [41] (glasses and projector).

12.2.3 A Framework for Combining Multiple Smart Displays

Glass Unlock has presented how one nomadic device, in this case smart glasses, can surmount the deficiency of another one (the public display of a smartphone) to support its goal of secure unlocking. In the second approach, a framework combines all nomadic devices to solve a deficiency shared by all of them, which is that one device/display is not sufficient for multi-tasking.

12.2.3.1 *Framework Concept*

Because of the limited screen estate, all mobile operating systems apply the "single activity" concept where one application receives the whole screen space at a time. To switch between multiple applications, several interaction steps are usually required such as (1) holding a button for a longer time or tapping it twice, respectively, to then (2b) choose another application from a list (which may require several swipe gestures before (2a)). Switching back to the former application requires at least 1) and 2b) to be performed again. It must be assumed that users will perceive this interaction process as cumbersome in situations where several application switches are required within a short period of time. The framework approach described in this section tries to solve this issue by making multiple subsequent application switches unnecessary. Such as desktop systems allow to position multiple application windows side by side, nomadic users should be able to place different application windows within their own mobile display space which is constituted by the number of worn or available display devices (e.g., smart watch, smart glasses, mobile projection). This is not very different from the ongoing research on nomadic interaction with pervasive displays, however, as all involved devices belong to the user, trust and multi-user issues can be neglected. Instead, a very easy means of interaction that is quicker and more convenient than the aforementioned application switching on a mobile phone has to be found. The bachelor thesis cited at the top of this section provides a good example to start from:

> "A discussion among friends in a chat application: Two friends could arrange to meet at the cinema. A conversation about a visit to the cinema implies a discussion about the offered movies and the corresponding showtimes. To access the relevant information, several application switches between the chat application and the web browser are required which may become annoying."

Smart watches and smart glasses, however, neither provide good means for text entry in a chat application nor to browse the web—and that for good reason because of their small size and/or lacking input capabilities. However, if information from the smartphone, in this example the list of currently screening movies, could be easily put on one of the other displays, no further application switches would be necessary and chatting on the phone about the movie to choose may become more pleasant.

Figure 12.8: Display Copy system example: After the smartphone is touched long with the ring it displays a clipping frame to select the start view for a connected smart watch and finally stores a copy of the display and the selected region (left). Touching the smart watch with the ring for a short time is recognized (middle) and shortly afterwards the picture is retrieved from the server (phone) and can then possibly be panned and zoomed on the watch (right).

The basic idea of the proposed framework is to use a ring (which may be a smart ring but does not necessarily have to) as mediator between the displays. As all of the other display-equipped nomadic devices (watch, glasses, projector, phone) usually have built-in magnetometers, they are able to recognize when they are "touched" with a metallic (magnetic) ring. This can be leveraged to transfer content between the devices and their displays if these share a common network, e.g., over bluetooth (note that the ring is not required to be part of this network). This idea was implemented to investigate a suitable interaction design the and rest of this section describes its implementation.

12.2.3.2 *Framework Implementation*

Related video

First of all, the smart phone is chosen to act as server since it is believed to be always available and that most of the times, content is moved from the phone to other displays, albeit other transfers should be possible as well. Other smart devices connect to the server over Bluetooth and maintain a connection in the background. Further on, all devices constantly monitor differences in the magnetic field and interpret sharp (read quick and high) changes as touch. If two of these changes occur within a certain time interval, these are recognized as touch down and touch up events and taking the time lying in-between into account, long touches can be recognized as well. This comes in handy as at least two actions are required by the interaction design: selecting a display to be moved on one side and its destination on the other side. Although this could be achieved with a single touch (touch the first device then touch the second device), falsely recognized touches may quickly bring the system out of sync with the user's intentions. It seems more robust to use the copy&paste metaphor, i.e. to store display content in a clipboard (hosted on the server on the phone) and paste it to the destination device upon request. Consequently, a long touch is used to store the display content of one device and a normal touch to paste the content to another device's display (see Figure 12.8).

This almost suffices the desired functionality except for the differences in size between the displays. If, for instance, the cinema website is to be copied to the smart watch, how should it appear on the 20 times smaller screen of the watch? A simple solution that was applied is to copy the display content as an image and to send this to the watch. The user, then, can use the commonly available panning and zooming techniques to position the content to their likings. Watches usually support this by touch gestures and on the Google Glass the content is placed on a large space around the user's head which can be explored using head rotation. The simple means of creating an image of the content, of course, comes at the expense of losing the possibility to interact with the content (e.g., looking up details of the movies) and future work might want to research more suitable intermediate formats between nomadic devices. Apart from that, to reduce possible zooming and panning steps that are more cumbersome to perform, at least on a watch, than on the phone, multiple clipping frames are shown above the display whenever the screen shot of a display is taken (Figure 12.8 left). The part within the clipping frame is used as start segment when copied to one of the other devices. These rectangles can be moved altogether to select the desired start segment on the phone, which provides for more overview and thus a more convenient selection. If, however, the user is content with the start segment, no further action is required after the long touch and the clipping frames disappear by themselves after a short period of time.

Pasting the stored display content to another device, now, only requires briefly touching the target display with the ring, which triggers it to request the current screen content together with the initial segment definition from the server (the phone) and display it accordingly. Panning and zooming allow the whole display content to be accessed and thus, allow for multi-tasking and multi-application usage in nomadic scenarios. Regarding this second approach, no related works are known so far.

To conclude this chapter, one advantage worthwhile mentioning is that the perspective on nomadic computing outlined by this chapter acknowledges that people will more likely carry "smartified" versions of their traditional and socially accepted wearables (rings, watches, clothes, glasses) than carrying additional sensors and devices only for the purpose of enabling a new interaction metaphor. This makes looking into their meaningful combinations so essential. At the same time, it would be interesting to research the limitations of the "combined design space" to assess which new required sensors or other hardware are really necessary for nomadic computing devices in the future. The "CES" score in Table 12.1 can be regarded as a preliminary step into this direction.

13 CONCLUSION

Mobile screen displays can only increase so much to support the increasing demand of nomadic productivity and entertainment that yearns after more screen real estate. We have seen that mobile projection in principal, is able to provide these large displays in nomadic scenarios and from small physical form factors. Moreover, we have seen that the additional screen can be leveraged beyond its size to address further deficiencies of current mobile devices, namely lacking support for multi-tasking, (privacy-respectful) collaboration, and leveraging the environment by creating new types of AR experiences and increasing the awareness of the user. Nevertheless, these advantages do not just come by themselves. As the analysis and classification of the literature (Chapter 3) have shown, related works have mostly recreated traditional projection scenarios using projections at a distance and out-of-reach interaction techniques that are suitable for media broadcasting but not for nomadic projection (cf. arguments in Chapter 5) and information management (cf. Chapter 4).

Based on these observations, in Chapter 5 a framework called Nomadic Projection Within Reach was proposed which in contrast to most previous works, values the flexibility of the projected display much more than its size and hypothetically allows for many deficiencies of current mobile interaction to be successfully addressed and solved. In the subsequent chapters 6 to 8, this hypothesis has been systematically studied motivated by the previously identified deficiencies (Subsection 1.1.1), guided by the classification of related work (Chapter 3), trying to answer the research questions formulated by Section 1.2. Regarding these, the thesis found out that

R1 a better overview, shorter task completion times and lower error rates can be achieved when small targets are involved (Chapter 4);

R2 new input modalities include pen-input on real paper Chapter 6, implicit body movement as extension to the Spotlight metaphor for peripheral interaction (Chapter 9) and more generally device movement to maintain privacy (chapters 7 and 9), touch-input on tables with correction for over- and undershooting, a preference for uni-manual transfer techniques between displays (Chapter 7), and gestures robustly functioning during walking (Chapter 9);

R3 merging of projections to larger shared displays and support for different postures in colocated sharing allow for different setups of intimacy between collaborators (Chapter 7). In remote collaboration, creating awareness for each other's actions diminishes unnecessary meta-conversations regarding each other's actions (Chapter 8);

R4 real (Chapter 7) or artificial (Chapter 8) MMDEs provide private and shared spaces that address privacy concerns sufficiently for collaborating users. When not colocated, awareness about each other's actions is essential to a privacy-respectful experience (Chapter 8).

In a further step, chapters 9 and 10 in Part III expanded the within-reach interaction range to Nomadic Projection Within Extended Reach, covering the cross-distance interaction space and bridging the gap between Nomadic Projection Within Reach and many existing works on distant interaction with mobile projections. As we have seen, these case studies could prove that Nomadic Projection Within (Extended) Reach can successfully address further deficiencies of current mobile interaction. In particular,

R5 including digital information in the user's periphery using AR provides for an alternative input channel that can be leveraged beside and in spite of other primary tasks such as walking (Chapter 9);

R6 robust selection gestures (Chapter 9) or quick activation gestures (Chapter 10) allow for micro-interactions in nomadic and on-the-go scenarios that have the potential to outperform task completion times using smartphones (Chapter 10).

By example of the case studies, and based on precise calculations of required Lumens for differently lit environments (Subsection 2.4.1), we have further seen that

R7 the close projection distance has allowed, for the first time, to make mobile projection usable in unaltered indoor environments, proving the prevalent opinion wrong that mobile projection due to its limited brightness is not mature enough for nomadic use cases, yet.

The results of the case studies and the personal experience of the author have then been transferred to 12 practical guidelines on using mobile projection which have been summarized in Chapter 11. These inform the design of future mobile devices whether projection is suitable at all and if so, when and how to apply Nomadic Projection Within (Extended) Reach to the application scenario.

Finally, Chapter 12 elaborated a possible future role of nomadic projection. By classifying this work within a broader scope of nomadic device support (and presenting further case studies of this design space), the unique advantages of nomadic projection have been elicited. These include the unique support for enabling *collaboration* and integrating the *environment*, e.g. for increasing awareness. Moreover, different to other nomadic device categories, projectors can be integrated to existing nomadic devices which are already socially accepted. They have only little impact on the device size, but add many opportunities to enrich interaction.

Reflecting on the case studies presented by this thesis, the Penbook, SurfacePhone, and IPC concepts could presumably, be combined to a single concept that equips any handheld mobile device with better support for single-user multi-tasking and multi-user collaboration. On the other hand, concepts presented by AMP-D and SpiderLight could presumably, be integrated with smart watches or glasses, to increase the user's awareness while on-the-go and to address the collaboration deficiency of smart watches and glasses.

The contributions of this thesis can be summarized as follows (a more detailed list was already given by Section 1.2):

- A new classification of existing works considering *nomadicity* and *interaction distance*, revealing an opportunity for mobile projection in nomadic scenarios that has been mostly unexploited so far across many application domains, but especially regarding information management;
- the framework of Nomadic Projection Within (Extended) Reach and its underlying calculations;
- five case studies, each providing new concept(s), implementation(s), and evaluation(s), confirming the hypothesized advantages of the framework as well as revealing some of its limitations;
- 12 concrete guidelines for the application of the framework;

These contributions made nomadic projection come *within reach*— physically *and* figuratively—to provide large displays from small devices in nomadic usage scenarios. Drawing a line to the beginning of this thesis, this allows future nomadic interaction to be more *transparent* (only enabled when required), more *integrated* (part of existing accessories), more *adaptive* (enables small and large displays as required), and more *convenient* (small device size), providing users with richer single- and multi-user interaction when they are on the go.

PUBLICATIONS BY THE AUTHOR

[W1] Gugenheimer, J., **Winkler, C.**, Wolf, D., Rukzio, E., "Interaction with Adaptive and Ubiquitous User Interfaces." In: *Companion Technology - A Paradigm Shift in Human-Technology Interaction*. Red. by J. Carbonell, M. Pinkal, H. Uszkoreit, M. M. Veloso, W. Wahlster, and M. J. Wooldridge. Cognitive Technologies. Springer, 2016, to appear (cit. on p. 185).

[W2] **Winkler, C.**, Gugenheimer, J., De Luca, A., Haas, G., Speidel, P., Dobbelstein, D., Rukzio, E., "Glass Unlock: Enhancing Security of Smartphone Unlocking Through Leveraging a Private Near-eye Display." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. ACM, 2015, pp. 1407–1410 (cit. on pp. 228, 231 sq.).

[W3] Gugenheimer, J., Knierim, P., **Winkler, C.**, Seifert, J., Rukzio, E., "UbiBeam: Exploring the Interaction Space for Home Deployed Projector-Camera Systems." In: *Human-Computer Interaction – INTERACT 2015*. Vol. 9298. Lecture Notes in Computer Science. Springer International Publishing, 2015, pp. 350–366 (cit. on p. 35).

[W4] **Winkler, C.**, Seifert, J., Dobbelstein, D., Rukzio, E., "Pervasive Information Through Constant Personal Projection: The Ambient Mobile Pervasive Display (AMP-D)." In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14, **Honorable Mention Award**. ACM, 2014, pp. 4117–4126 (cit. on p. 160).

[W5] **Winkler, C.**, Löchtefeld, M., Dobbelstein, D., Krüger, A., Rukzio, E., "SurfacePhone: A Mobile Projection Device for Single- and Multiuser Everywhere Tabletop Interaction." In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 3513–3522 (cit. on p. 101).

[W6] **Winkler, C.**, "Peripheral Interaction On-The-Go." In: *Workshop on Peripheral Interaction: Shaping the Research and Design Space at CHI 2014*. CHI '14. 2014 (cit. on pp. 183, 212, 219).

[W7] **Winkler, C.**, Rukzio, E., "Projizierte tischbasierte Benutzungsschnittstellen." In: *Informatik-Spektrum* 37.5, Springer (2014), pp. 413–417.

[W8] **Winkler, C.**, Seifert, J., Reinartz, C., Krahmer, P., Rukzio, E., "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, **Best Note Award**. ACM, 2013, pp. 283–286 (cit. on p. 87).

[W9] **Winkler, C.**, Pfeuffer, K., Rukzio, E., "Investigating mid-air pointing interaction for projector phones." In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*. ITS '12. ACM, 2012, pp. 85–94 (cit. on pp. 57, 84).

[W10] **Winkler, C.**, Hutflesz, P., Holzmann, C., Rukzio, E., "Wall Play: A Novel Wall/Floor Interaction Concept for Mobile Projected Gaming." In: *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services Companion*. MobileHCI '12. ACM, 2012, pp. 119–124 (cit. on pp. 36, 39 sq., 52, 54, 109).

[W11] **Winkler, C.**, Reinartz, C., Nowacka, D., Rukzio, E., "Interactive phone call: synchronous remote collaboration and projected interactive surfaces." In: *Proceedings of the 2011 ACM international conference on Interactive tabletops and surfaces*. ITS '11. ACM, 2011, pp. 61–70 (cit. on p. 132).

[W12] **Winkler, C.**, Broscheit, M., Rukzio, E., "NaviBeam: Indoor Assistance and Navigation for Shopping Malls through Projector Phones." In: *MP2: Workshop on Mobile and Personal Projection*. CHI 2011. 2011 (cit. on pp. 170, 182).

[W13] Schaub, F., Könings, B., Lang, P., Wiedersheim, B.**, Winkler, C.**, Weber, M., "PriCal: Context-adaptive Privacy in Ambient Calendar Displays." In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '14. ACM, 2014, pp. 499–510.

[W14] Seifert, J., Boring, S.**, Winkler, C.**, Schaub, F., Schwab, F., Herrdum, S., Maier, F., Mayer, D., Rukzio, E., "Hover Pad: Interacting with Autonomous and Self-actuated Displays in Space." In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. UIST '14. ACM, 2014, pp. 139–147 (cit. on pp. 20, 216).

[W15] Derthick, K., Scott, J., Villar, N.**, Winkler, C.,** "Exploring smartphone-based web user interfaces for appliances." In: *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*. MobileHCI '13. ACM, 2013, pp. 227–236.

[W16] Valderrama Bahamondez, E. d. C.**, Winkler, C.**, Schmidt, A., "Utilizing multimedia capabilities of mobile phones to support teaching in schools in rural panama." In: *Proceedings of* *the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. ACM, 2011, pp. 935–944.

[W17] Shirazi, A., **Winkler, C.**, Schmidt, A., "SENSE-SATION: An extensible platform for integration of phones into the Web." In: *Internet of Things (IOT), 2010*. Internet of Things (IOT), 2010. 2010, pp. 1–8.

[W18] Shirazi, A. S., **Winkler, C.**, Schmidt, A., "Flashlight Interaction: A Study on Mobile Phone Interaction Techniques with Large Displays." In: *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '09. ACM, 2009, 93:1–93:2.

[19] Ahern, S., Eckles, D., Good, N. S., King, S., Naaman, M., Nair, R., "Over-exposed?: privacy patterns and considerations in online and mobile photo sharing." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. CHI '07. ACM, 2007, pp. 357–366 (cit. on p. 135).

[20] *AiRScouter Head Mounted Display | Brother UK - Brother at your side*. URL: http://www.brother.co.uk/business-solutions/communication-and-collaboration/airscouter (visited on 07/11/2015) (cit. on p. 19).

[21] Akiyama, Y., Miyashita, H., "Electronic Kit with No Current Flow That Uses Projection Mapping." In: *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '14. ACM, 2014, pp. 155–156 (cit. on pp. 52, 54).

[22] Alex, D., Gorur, P., Amrutur, B., Ramakrishnan, K., "Lamp-Top: Touch Detection for a Projector-camera System Based on Shape Classification." In: *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*. ITS '13. ACM, 2013, pp. 429–432 (cit. on p. 38).

[23] Aliaga, D. G., Yeung, Y. H., Law, A., Sajadi, B., Majumder, A., "Fast High-resolution Appearance Editing Using Superimposed Projections." In: *ACM Trans. Graph.* 31.2 (2012), 13:1–13:13 (cit. on p. 33).

[24] *Amadeus - Interactive Projection Mapped Piano - Projection Mapping Central*. URL: http://projection-mapping.org/amadeus-interactive-projection-mapped-piano/ (visited on 06/13/2015) (cit. on p. 52).

[25] *Android Beam*. Wikipedia, the free encyclopedia. 31, 2015. URL: https://en.wikipedia.org/w/index.php?title=Android_Beam&oldid=644955595 (visited on 07/06/2015) (cit. on p. 7).

[26] *ANSI/HFES 100-2007. Human Factors Engineering of Computer Workstations*. 2007 (cit. on p. 28).

[27] *Apple Retina Display – Jonesblog*. URL: http://prometheus.med.utah.edu/~bwjones/2010/06/apple-retina-display/ (visited on 09/11/2015) (cit. on p. 213).

[28] Arth, C., Reitmayr, G., Schmalstieg, D., "Full 6DOF pose estimation from geo-located images." In: *Proc. of*. Computer Vision–ACCV 2012. Springer, 2013, pp. 705–717 (cit. on p. 180).

[29] *Audi Production Lab: Reise in die Zukunft*. Audi MediaCenter. URL: https://www.audi-mediacenter.com/de/pressemitteilungen/audi-production-lab-reise-in-die-zukunft-4419 (visited on 09/22/2015) (cit. on p. 183).

[30] Augsten, T., Kaefer, K., Meusel, R., Fetzer, C., Kanitz, D., Stoff, T., Becker, T., Holz, C., Baudisch, P., "Multitoe: High-precision Interaction with Back-projected Floors Based on High-resolution Multi-touch Input." In: *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*. UIST '10. ACM, 2010, pp. 209–218 (cit. on p. 37).

[31] *Avegant Glyph*. URL: http://avegant.com/ (visited on 07/11/2015) (cit. on p. 19).

[32] *Average Hand Size - Your resource for Hand Size, Palm Size and Finger Length Averages*. URL: http://www.theaveragebody.com/average_hand_size.php (visited on 09/07/2015) (cit. on p. 187).

[33] Aviv, A. J., Gibson, K., Mossop, E., Blaze, M., Smith, J. M., "Smudge Attacks on Smartphone Touch Screens." In: *Proceedings of the 4th USENIX Conference on Offensive Technologies*. WOOT'10. USENIX Association, 2010, pp. 1–7 (cit. on p. 228).

[34] Bachl, S., Tomitsch, M., Kappel, K., Grechenig, T., "The effects of personal displays and transfer techniques on collaboration strategies in multi-touch based multi-display environments." In: *Human-Computer Interaction–INTERACT 2011*. Springer, 2011, pp. 373–390 (cit. on p. 110).

[35] *Balsamiq. Rapid, effective and fun wireframing software. - Balsamiq*. URL: https://balsamiq.com/ (visited on 07/10/2015) (cit. on p. 12).

[36] Banerjee, A., Burstyn, J., Girouard, A., Vertegaal, R., "Pointable: an in-air pointing technique to manipulate out-of-reach targets on tabletops." In: *Proceedings of the 2011 ACM international conference on Interactive tabletops and surfaces*. ITS '11. ACM, 2011, pp. 11–20 (cit. on p. 60).

[37] Bau, O., Poupyrev, I., "REVEL: Tactile Feedback Technology for Augmented Reality." In: *ACM Trans. Graph.* 31.4 (2012), 89:1–89:11 (cit. on p. 42).

[38] Bau, O., Poupyrev, I., Israr, A., Harrison, C., "TeslaTouch: Electrovibration for Touch Surfaces." In: *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*. UIST '10. ACM, 2010, pp. 283–292 (cit. on p. 42).

[39] Baudisch, P., Chu, G., "Back-of-device interaction allows creating very small touch devices." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. ACM, 2009, pp. 1923–1932 (cit. on p. 214).

[40] Beardsley, P., Baar, J., Raskar, R., Forlines, C., "Interaction using a handheld projector." In: *Computer Graphics and Applications, IEEE* 25.1 (2005), pp. 39–43 (cit. on pp. 32, 44, 135).

[41] Benko, H., Ofek, E., Zheng, F., Wilson, A. D., "FoveAR: Combining an Optically See-Through Near-Eye Display with Spatial Augmented Reality Projections." In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*. UIST '15. ACM, 2015, to appear (cit. on p. 233).

[42] Benko, H., Jota, R., Wilson, A., "MirageTable: freehand interaction on a projected augmented reality tabletop." In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. CHI '12. ACM, 2012, pp. 199–208 (cit. on p. 32).

[43] Benko, H., Wilson, A. D., "Multi-point interactions with immersive omnidirectional visualizations in a dome." In: *ACM International Conference on Interactive Tabletops and Surfaces*. ITS '10. ACM, 2010, pp. 19–28 (cit. on pp. 52, 54).

[44] Benko, H., Wilson, A. D., "Pinch-the-sky dome: freehand multi-point interactions with immersive omni-directional data." In: *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*. CHI EA '10. ACM, 2010, pp. 3045–3050 (cit. on pp. 52, 54).

[45] Benko, H., Wilson, A. D., Zannier, F., "Dyadic Projected Spatial Augmented Reality." In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. UIST '14. ACM, 2014, pp. 645–655 (cit. on p. 32).

[46] *Best Projection Mapping Softwares*. URL: http://projection-mapping.org/which-projection-mapping-software/ (visited on 06/13/2015) (cit. on p. 53).

[47] Bianchi, A., Oakley, I., Kostakos, V., Kwon, D. S., "The Phone Lock: Audio and Haptic Shoulder-surfing Resistant PIN Entry Methods for Mobile Devices." In: *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction*. TEI '11. ACM, 2011, pp. 197–200 (cit. on p. 228).

[48] Billinghurst, M., Bowskill, J., Dyer, N., Morphett, J., "Spatial information displays on a wearable computer." In: *IEEE Computer Graphics and Applications* 18.6 (1998), pp. 24–31 (cit. on pp. 165, 170).

[49] Bimber, O., Emmerling, A., Klemmer, T., "Embedded entertainment with smart projectors." In: *Computer* 38.1 (2005), pp. 48–55 (cit. on pp. 33, 40).

[50] Bimber, O., Wetzstein, G., Emmerling, A., Nitschke, C., "Enabling view-dependent stereoscopic projection in real environments." In: *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality, 2005. Proceedings*. Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality, 2005. Proceedings. 2005, pp. 14–23 (cit. on p. 35).

[51] Bimber, O., Coriand, F., Kleppe, A., Bruns, E., Zollmann, S., Langlotz, T., "Superimposing pictorial artwork with projected imagery." In: *ACM SIGGRAPH 2005 Courses*. SIGGRAPH '05. ACM, 2005 (cit. on pp. 33 sq., 52).

[52] Birnholtz, J., Reynolds, L., Luxenberg, E., Gutwin, C., Mustafa, M., "Awareness beyond the desktop: exploring attention and distraction with a projected peripheral-vision display." In: *Proceedings of Graphics Interface 2010*. GI '10. Canadian Information Processing Society, 2010, pp. 55–62 (cit. on p. 52).

[53] Blasko, G., Coriand, F., Feiner, S., "Exploring interaction with a simulated wrist-worn projection display." In: *Proceedings of the Ninth IEEE International Symposium on Wearable Computers*. 2005, pp. 2–9 (cit. on p. 187).

[54] Bränzel, A., Holz, C., Hoffmann, D., Schmidt, D., Knaust, M., Lühne, P., Meusel, R., Richter, S., Baudisch, P., "GravitySpace: Tracking Users and Their Poses in a Smart Room Using a Pressure-sensing Floor." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. ACM, 2013, pp. 725–734 (cit. on pp. 37, 40, 53).

[55] Brooke, J., "SUS-A quick and dirty usability scale." In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7 (cit. on p. 13).

[56] Brown, M. S., Song, P., Cham, T.-J., "Image Pre-Conditioning for Out-of-Focus Projector Blur." In: *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2*. CVPR '06. IEEE Computer Society, 2006, pp. 1956–1963 (cit. on p. 35).

[57] *Bullet Physics Simulation*. URL: http://bulletphysics.org/wordpress/ (visited on 05/08/2012) (cit. on p. 52).

[58] Burgbacher, U., Hinrichs, K., "An Implicit Author Verification System for Text Messages Based on Gesture Typing Biometrics." In: *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 2951–2954 (cit. on p. 228).

[59] *Burton Inc.* URL: http://burton-jp.com/en/index.htm (visited on 06/14/2015) (cit. on pp. 44, 51, 55).

[60] Cai, X., Xie, X., Li, G., Song, W., Zheng, Y., Wang, Z., "A new method of detecting fingertip touch for the projector-camera HCI system." In: *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2014 IEEE International Symposium on Circuits and Systems (ISCAS). 2014, pp. 526–529 (cit. on pp. 37, 40).

[61] Cao, X., "Handheld Projector Interaction." PhD thesis. University of Toronto, 2009 (cit. on pp. 43, 48).

[62] Cao, X., Forlines, C., Balakrishnan, R., "Multi-user interaction using handheld projectors." In: *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*. UIST '07. 2007, pp. 43–52 (cit. on pp. 44, 49 sq., 55, 85).

[63] Cao, X., Balakrishnan, R., "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. ACM, 2006, pp. 225–234 (cit. on pp. 38, 44, 49 sq., 54 sq.).

[64] Carter, T., Seah, S. A., Long, B., Drinkwater, B., Subramanian, S., "UltraHaptics: Multi-point Mid-air Haptic Feedback for Touch Surfaces." In: *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*. UIST '13. ACM, 2013, pp. 505–514 (cit. on p. 42).

[65] Casiez, G., Vogel, D., Balakrishnan, R., Cockburn, A., "The Impact of Control-Display Gain on User Performance in Pointing Tasks." In: *HUMAN–COMPUTER INTERACTION* 23.3 (2008), pp. 215–250 (cit. on p. 63).

[66] Cassinelli, A., Angesleva, J., Watanabe, Y., Frasca, G., Ishikawa, M., "Skin Games." In: *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*. ITS '12. ACM, 2012, pp. 323–326 (cit. on pp. 44, 51 sq., 54).

[67] Cauchard, J. R., Fraser, M., Alexander, J., Subramanian, S., "Offsetting Displays on Mobile Projector Phones." In: *Ubiprojection Workshop*. Pervasive. 2010 (cit. on p. 30).

[68] Cauchard, J. R., Löchtefeld, M., Irani, P., Schoening, J., Krüger, A., Fraser, M., Subramanian, S., "Visual separation in mobile multi-display environments." In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*. UIST '11. ACM, 2011, pp. 451–460 (cit. on pp. 60, 105).

[69] Chan, L., Hsieh, C.-H., Chen, Y.-L., Yang, S., Huang, D.-Y., Liang, R.-H., Chen, B.-Y., "Cyclops: Wearable and Single-Piece Full-Body Gesture Input Devices." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. ACM, 2015, pp. 3001–3009 (cit. on p. 41).

[70] Chapanis, A., Ochsman, R., Parrish, R., Weeks, G., "Studies in Interactive Communication: I. The Effects of Four Communication Modes on the Behavior of Teams During Cooperative Problem-Solving." In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 14.6 (1972), pp. 497–505 (cit. on p. 134).

[71] Chehimi, F., "Storybook Augmentation through Portable Projectors, 3D Augmented Reality and Anaglyph." In: *Papercomp Workshop*. Ubicomp. 2010 (cit. on p. 34).

[72] Chen, N., Guimbretiere, F., Dixon, M., Lewis, C., Agrawala, M., "Navigation techniques for dual-display e-book readers." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. ACM, 2008, pp. 1779–1788 (cit. on p. 103).

[73] Chen, X. '., Grossman, T., Wigdor, D. J., Fitzmaurice, G., "Duet: Exploring Joint Interactions on a Smart Phone and a Smart Watch." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 159–168 (cit. on p. 233).

[74] *cocos3d « The Brenwill Workshop*. URL: http://brenwill.com/cocos3d/ (visited on 05/08/2012) (cit. on p. 52).

[75] Costanza, E., Inverso, S. A., Pavlov, E., Allen, R., Maes, P., "eye-q: eyeglass peripheral display for subtle intimate notifications." In: *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. MobileHCI '06. ACM, 2006, pp. 211–218 (cit. on pp. 163 sq.).

[76] Cotting, D., Naef, M., Gross, M., Fuchs, H., "Embedding imperceptible patterns into projected images for simultaneous acquisition and display." In: *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*. ISMAR 2004. 2004, pp. 100–109 (cit. on p. 41).

[77] Coventry, L., De Angeli, A., Johnson, G., "Usability and Biometric Verification at the ATM Interface." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '03. ACM, 2003, pp. 153–160 (cit. on p. 228).

[78] Cowan, L. G., Li, K. A., "ShadowPuppets: supporting collocated interaction with mobile projector phones using hand shadows." In: *Proceedings of the 2011 annual conference on Human factors in computing systems*. CHI '11. ACM, 2011, pp. 2707–2716 (cit. on pp. 36, 40, 212, 218).

[79] Cowan, L. G., Weibel, N., Griswold, W. G., Pina, L. R., Hollan, J. D., "Projector phone use: practices and social implications." In: *Personal Ubiquitous Comput.* 16.1 (2012), pp. 53–63 (cit. on pp. 44, 48, 50).

[80] Dancu, A., Franjcic, Z., Fjeld, M., "Smart Flashlight: Map Navigation Using a Bike-mounted Projector." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 3627–3630 (cit. on pp. 44, 47, 49 sq.).

[81] Dao, V. N., Hosoi, K., Sugimoto, M., "A semi-automatic real-time calibration technique for a handheld projector." In: *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*. ACM, 2007, pp. 43–46 (cit. on p. 32).

[82] De Luca, A., Harbach, M., Zezschwitz, E., Maurer, M.-E., Slawik, B. E., Hussmann, H., Smith, M., "Now You See Me, Now You Don'T: Protecting Smartphone Authentication from Shoulder Surfers." In: *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 2937–2946 (cit. on p. 228).

[83] De Luca, A., Zezschwitz, E., Nguyen, N. D. H., Maurer, M.-E., Rubegni, E., Scipioni, M. P., Langheinrich, M., "Back-of-device Authentication on Smartphones." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. ACM, 2013, pp. 2389–2398 (cit. on p. 232).

[84] *Did you know that smartphone screens nearly doubled in size since 2007?* Phone Arena. URL: `http://www.phonearena.com/news/Did-you-know-that-smartphone-screens-nearly-doubled-in-size-since-2007_id52067` (visited on 07/05/2015) (cit. on p. 5).

[85] Fanello, S. R., Keskin, C., Izadi, S., Kohli, P., Kim, D., Sweeney, D., Criminisi, A., Shotton, J., Kang, S. B., Paek, T., "Learning to Be a Depth Camera for Close-range Human Capture and Interaction." In: *ACM Trans. Graph.* 33.4 (2014), 86:1–86:11 (cit. on p. 41).

[86] Feiner, S., MacIntyre, B., Haupt, M., Solomon, E., "Windows on the world: 2 D windows for 3 D augmented reality." In: *Proceedings of the 6 th annual ACM symposium on User interface software and technology*. UIST '93. 1993, pp. 145–155 (cit. on pp. 163, 165).

[87] Feiner, S., MacIntyre, B., Höllerer, T., Webster, A., "A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment." In: *Personal Technologies* 1.4 (1997), pp. 208–217 (cit. on p. 163).

[88] Ferreira, D., Goncalves, J., Kostakos, V., Barkhuus, L., Dey, A. K., "Contextual Experience Sampling of Mobile Application Micro-usage." In: *Proceedings of the 16th international conference on Human computer interaction with mobile devices and services*. MobileHCI '14. ACM, 2014, pp. 91–100 (cit. on p. 187).

[89] Follmer, S., Leithinger, D., Olwal, A., Hogge, A., Ishii, H., "inFORM: Dynamic Physical Affordances and Constraints Through Shape and Object Actuation." In: *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*. UIST '13. ACM, 2013, pp. 417–426 (cit. on p. 38).

[90] *Fujitsu to start selling laptops with optional pico-projector accessories*. URL: http://www.picoprojector-info.com/fujitsu-start-selling-laptops-optional-pico-projector-accessories (visited on 11/15/2011) (cit. on p. 44).

[91] Garriss, S., Cáceres, R., Berger, S., Sailer, R., Doorn, L., Zhang, X., "Trustworthy and personalized computing on public kiosks." In: *Proceeding of the 6th international conference on Mobile systems, applications, and services*. MobiSys '08. ACM, 2008, pp. 199–210 (cit. on p. 135).

[92] Glaser, B. G., Strauss, A. L., *The discovery of grounded theory: Strategies for qualitative research*. Transaction Publishers, 2009 (cit. on p. 13).

[93] *Google Glass*. Wikipedia. 5, 2015. URL: https://de.wikipedia.org/w/index.php?title=Google_Glass&oldid=143770647 (visited on 07/06/2015) (cit. on pp. 6, 163).

[94] Greaves, A., Hang, A., Rukzio, E., "Picture browsing and map interaction using a projector phone." In: *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. MobileHCI '14. 2008, pp. 527–530 (cit. on p. 49).

[95] Greaves, A., Rukzio, E., "Evaluation of picture browsing using a projector phone." In: *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. MobileHCI '08. ACM, 2008, pp. 351–354 (cit. on pp. 44, 48 sq., 60).

[96] Grundhofer, A., Bimber, O., "Real-Time Adaptive Radiometric Compensation." In: *IEEE Transactions on Visualization and Computer Graphics* 14.1 (2008), pp. 97–108 (cit. on p. 33).

[97] *Guide to DIN EN 12464-1*. URL: http://en.licht.de/en/info-and-service/publications-and-downloads/detailview/product/guide_to_din_en_12464_1/bp/398/noc/1/ (visited on 07/02/2015).

[98] Gunaratne, J. A., Bernheim Brush, A. J., "Newport: enabling sharing during mobile calls." In: *Proceedings of the 28th international conference on Human factors in computing systems*. CHI '10. 2010, pp. 343–352 (cit. on pp. 132 sq., 135 sq., 154).

[99] Gupta, S., Morris, D., Patel, S. N., Tan, D., "AirWave: Non-contact Haptic Feedback Using Air Vortex Rings." In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '13. ACM, 2013, pp. 419–428 (cit. on p. 42).

[100] Gurevich, P., Lanir, J., Cohen, B., Stone, R., "TeleAdvisor: a versatile augmented reality tool for remote assistance." In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. CHI '12. ACM, 2012, pp. 619–622 (cit. on pp. 44, 50, 85).

[101] Hainich, R., Bimber, O., *Displays: Fundamentals and Applications*. CRC Press, 2014 (cit. on p. 19).

[102] Harbach, M., Zezschwitz, E. v., Fichtner, A., Luca, A. D., Smith, M., "It's a Hard Lock Life: A Field Study of Smartphone (Un)Locking Behavior and Risk Perception." In: *Symposium On Usable Privacy and Security*. SOUPS 2014. USENIX Association, 2014, pp. 213–230 (cit. on pp. 228 sq.).

[103] Hardy, J., Alexander, J., "Toolkit support for interactive projected displays." In: *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*. MUM '12. ACM, 2012, 42:1–42:10 (cit. on pp. 44, 53).

[104] Harrison, C., Benko, H., Wilson, A. D., "OmniTouch: wearable multitouch interaction everywhere." In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*. UIST '11. ACM, 2011, pp. 441–450 (cit. on pp. 41, 44, 48, 54 sq., 64, 85, 120, 167, 187).

[105] Harrison, C., Ramamurthy, S., Hudson, S. E., "On-body interaction: armed and dangerous." In: *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*. TEI '12. ACM, 2012, pp. 69–76 (cit. on p. 37).

[106] Harrison, C., Tan, D., Morris, D., "Skinput: Appropriating the Body As an Input Surface." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. ACM, 2010, pp. 453–462 (cit. on pp. 37, 41, 167).

[107] Hart, S., Staveland, L., "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." In: *Advances in Psychology* 52 (1988), pp. 139–183 (cit. on p. 69).

[108] Hashida, T., NIshimura, K., Naemura, T., "Hand-rewriting: automatic rewriting similar to natural handwriting." In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*. ITS '12. ACM, 2012, pp. 153–162 (cit. on p. 92).

[109] Hashimoto, N., Tashiro, A., Saito, H., Ogawa, S., "Multifocal Projection for Dynamic Multiple Objects." In: *ACM SIGGRAPH 2014 Posters*. SIGGRAPH '14. ACM, 2014, 32:1–32:1 (cit. on p. 35).

[110] Henze, N., Rukzio, E., Boll, S., "100,000,000 taps: analysis and improvement of touch performance in the large." In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '11. 2011, pp. 133–142 (cit. on pp. 14, 125, 128).

[111] Hinckley, K., Bi, X., Pahud, M., Buxton, B., "Informal information gathering techniques for active reading." In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. CHI '12. ACM, 2012, pp. 1893–1896 (cit. on p. 91).

[112] Hinckley, K., Dixon, M., Sarin, R., Guimbretiere, F., Balakrishnan, R., "Codex: a dual screen tablet computer." In: *Proceedings of the 27th international conference on Human factors in computing systems*. CHI '09. ACM, 2009, pp. 1933–1942 (cit. on pp. 91, 102–105, 130).

[113] Hinckley, K., Pausch, R., Goble, J. C., Kassell, N. F., "A survey of design issues in spatial input." In: *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology*. UIST '94. ACM, 1994, pp. 213–222 (cit. on p. 61).

[114] Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P., Smith, M., "Stitching: pen gestures that span multiple displays." In: *Proceedings of the working conference on Advanced visual interfaces*. AVI '04. ACM, 2004, pp. 23–31 (cit. on pp. 104, 129).

[115] Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., Benko, H., Buxton, B., "Pen + touch = new tools." In: *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. UIST '10. ACM, 2010, pp. 27–36 (cit. on p. 91).

[116] *Holographic Protest Against Gag Law At Spanish Parliament*. URL: http://www.gettyimages.com/detail/news-photo/holographic-images-representing-a-protest-are-projected-in-news-photo/469208638?Language=en-US (visited on 06/26/2015) (cit. on p. 53).

[117] Holz, C., Baudisch, P., "Understanding Touch." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. ACM, 2011, pp. 2501–2510 (cit. on p. 128).

[118] Hosoi, K., Dao, V. N., Mori, A., Sugimoto, M., "VisiCon: a robot control interface for visualizing manipulation using a handheld projector." In: *Proceedings of the international conference on Advances in computer entertainment technology*. ACE '07. ACM, 2007, pp. 99–106 (cit. on pp. 44, 50, 52, 54).

[119] Hu, J., Li, G., Xie, X., Lv, Z., Wang, Z., "Bare-fingers Touch Detection by the Button's Distortion in Projector-camera System." In: *IEEE Transactions on Circuits and Systems for Video Technology* Early Access Online (2013) (cit. on p. 40).

[120] Huber, J., Steimle, J., Liao, C., Liu, Q., Mühlhäuser, M., "LightBeam: Interacting with Augmented Real-world Objects in Pico Projections." In: *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*. MUM '12. ACM, 2012, 16:1–16:10 (cit. on pp. 44, 51, 53, 55).

[121] *InfoComm International, the Audiovisual (AV) Association - ANSI / INFOCOMM 3M-2011, Projected Image System Contrast Ratio*. InfoComm International. 2011. URL: http://www.infocomm.org/cps/rde/xchg/infocomm/hs.xsl/32932.htm (visited on 08/11/2015) (cit. on p. 28).

[122] *Intel shows off their OASIS pico-projector object recognition system*. URL: http://www.picoprojector-info.com/intel-shows-their-oasis-pico-projector-object-recognition-system (visited on 06/24/2014) (cit. on pp. 52, 54).

[123] *ISO 9241-210:2010 - Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems*. URL: http://www.iso.org/iso/catalogue_detail.htm?csnumber=52075 (visited on 07/07/2015) (cit. on p. 11).

[124] *ISO 9241-302:2008 - Ergonomics of human-system interaction – Part 302: Terminology for electronic visual displays*. 2008. URL: http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=40097 (visited on 08/11/2015) (cit. on p. 28).

[125] Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., Fitzgibbon, A., "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. ACM, 2011, pp. 559–568 (cit. on pp. 40 sq., 202).

[126] James R. Lewis, *IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use*. Technical 54.786. Human Factors Group (cit. on p. 114).

[127] *JigLibX - Home*. URL: http://jiglibx.codeplex.com/ (visited on 03/27/2013) (cit. on p. 176).

[128] Jones, B. R., Benko, H., Ofek, E., Wilson, A. D., "IllumiRoom: Peripheral Projected Illusions for Interactive Experiences." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. ACM, 2013, pp. 869–878 (cit. on pp. 32, 52, 54).

[129] Jones, B., Sodhi, R., Murdock, M., Mehra, R., Benko, H., Wilson, A., Ofek, E., MacIntyre, B., Raghuvanshi, N., Shapira, L., "RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-camera Units." In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. UIST '14. ACM, 2014, pp. 637–644 (cit. on pp. 32, 44, 46, 52, 54 sq., 85).

[130] Jota, R., Nacenta, M. A., Jorge, J. A., Carpendale, S., Greenberg, S., "A comparison of ray pointing techniques for very large displays." In: *Proceedings of Graphics Interface 2010*. GI '10. Canadian Information Processing Society, 2010, pp. 269–276 (cit. on pp. 60, 68).

[131] Kajastila, R., Hämäläinen, P., "Augmented Climbing: Interacting with Projected Graphics on a Climbing Wall." In: *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*. CHI EA '14. ACM, 2014, pp. 1279–1284 (cit. on pp. 52, 54).

[132] Kajiwara, Y., Tajimi, K., Uemura, K., Sakata, N., Nishida, S., "ClippingLight: a method for easy snapshots with projection viewfinder and tilt-based zoom control." In: *Proceedings of the 2nd Augmented Human International Conference*. AH '11. ACM, 2011, 14:1–14:8 (cit. on pp. 44, 47).

[133] Kane, S. K., Avrahami, D., Wobbrock, J. O., Harrison, B., Rea, A. D., Philipose, M., LaMarca, A., "Bonfire: a nomadic system for hybrid laptop-tabletop interaction." In: *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. 2009, pp. 129–138 (cit. on pp. 37, 40 sq., 44, 48, 55, 103 sqq., 120, 215).

[134] Karitsuka, T., Sato, K., "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. IEEE Computer Society, 2003, p. 321 (cit. on p. 48).

[135] Kaufmann, B., Hitz, M., "Eye-Shield: protecting bystanders from being blinded by mobile projectors." In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ITS '11. ACM, 2011, pp. 31–34 (cit. on pp. 23, 30).

[136] Kauko, J., Häkkilä, J., "Shared-screen social gaming with portable devices." In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. MobileHCI '10. ACM, 2010, pp. 317–326 (cit. on pp. 104, 129).

[137] Khalilbeigi, M., Lissermann, R., Mühlhäuser, M., Steimle, J., "Xpaaand: Interaction Techniques for Rollable Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Com-*

puting Systems. CHI '11. ACM, 2011, pp. 2729–2732 (cit. on p. 53).

[138] Khot, R. A., Kumaraguru, P., Srinathan, K., "WYSWYE: Shoulder Surfing Defense for Recognition Based Graphical Passwords." In: *Proceedings of the 24th Australian Computer-Human Interaction Conference*. OzCHI '12. ACM, 2012, pp. 285–294 (cit. on p. 228).

[139] Kim, J., Seo, J., Han, T.-D., "AR Lamp: Interactions on Projection-based Augmented Reality for Interactive Learning." In: *Proceedings of the 19th International Conference on Intelligent User Interfaces*. IUI '14. ACM, 2014, pp. 353–358 (cit. on p. 52).

[140] Kim, S.-H., Kim, J.-W., Kim, S.-Y., Cho, H.-G., "A New Shoulder-surfing Resistant Password for Mobile Environments." In: *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*. ICUIMC '11. ACM, 2011, 27:1–27:8 (cit. on p. 228).

[141] Kjeldskov, J., Skov, M. B., "Was It Worth the Hassle?: Ten Years of Mobile HCI Research Discussions on Lab and Field Evaluations." In: *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*. MobileHCI '14. ACM, 2014, pp. 43–52 (cit. on p. 14).

[142] Kleinrock, L., "Nomadic Computing—an Opportunity." In: *SIGCOMM Comput. Commun. Rev.* 25.1 (1995), pp. 36–40 (cit. on p. 3).

[143] Kleinrock, L., "Breaking Loose." In: *Commun. ACM* 44.9 (2001), pp. 41–46 (cit. on pp. 3, 43, 46).

[144] Knibbe, J., Benko, H., Wilson, A. D., *Juggling the Effects of Latency: Motion Prediction Approaches to Reducing Latency in Dynamic Projector-Camera Systems*. 35. 2015 (cit. on p. 30).

[145] Köhler, M., Patel, S. N., Summet, J. W., Stuntebeck, E. P., Abowd, G. D., "TrackSense: infrastructure free precise indoor positioning using projected patterns." In: *Proceedings of the 5th international conference on Pervasive computing*. PERVASIVE'07. Springer-Verlag, 2007, pp. 334–350 (cit. on p. 41).

[146] Koike, H., Matoba, Y., Takahashi, Y., "AquaTop Display: Interactive Water Surface for Viewing and Manipulating Information in a Bathroom." In: *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*. ITS '13. ACM, 2013, pp. 155–164 (cit. on p. 38).

[147] Korn, O., Schmidt, A., Hörz, T., "The Potentials of In-situ-projection for Augmented Workplaces in Production: A Study with Impaired Persons." In: *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '13. ACM, 2013, pp. 979–984 (cit. on pp. 44, 52, 54 sq.).

[148] Krum, D. M., Suma, E. A., Bolas, M., "Augmented reality using personal projection and retroreflection." In: *Personal Ubiquitous Comput.* 16.1 (2012), pp. 17–26 (cit. on pp. 33, 44, 54).

[149] Ladetto, Q., Merminod, B., "Digital magnetic compass and gyroscope integration for pedestrian navigation." In: *9th International Conference on Integrated Navigation Systems, St. Petersburg*. 2002, pp. 111–120 (cit. on p. 178).

[150] Laput, G., Xiao, R., Chen, X. '., Hudson, S. E., Harrison, C., "Skin Buttons: Cheap, Small, Low-powered and Clickable Fixed-icon Laser Projectors." In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. UIST '14. ACM, 2014, pp. 389–394 (cit. on pp. 44, 85).

[151] Laramee, R. S., Ware, C., "Rivalry and interference with a head-mounted display." In: *ACM Trans. Comput.-Hum. Interact.* 9.3 (2002), pp. 238–251 (cit. on p. 19).

[152] *Laser safety*. Wikipedia, the free encyclopedia. 15, 2015. URL: https://en.wikipedia.org/w/index.php?title=Laser_safety&oldid=667092394 (visited on 06/17/2015) (cit. on p. 23).

[153] *Le Petit Chef: Projection Mapping on your Dinner Plate! - Projection Mapping Central*. URL: http://projection-mapping.org/le-petit-chef-projection-mapping-on-your-dinner-plate/ (visited on 06/13/2015) (cit. on p. 53).

[154] Leithinger, D., Follmer, S., Olwal, A., Luescher, S., Hogge, A., Lee, J., Ishii, H., "Sublimate: State-changing Virtual and Physical Rendering to Augment Interaction with Shape Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. ACM, 2013, pp. 1441–1450 (cit. on p. 38).

[155] *Lenovo Magic View: Prototyp der Dual Screen-Smartwatch im Hands on-Video*. URL: http://www.mobilegeeks.de/artikel/lenovo-magic-view-prototyp-der-dual-screen-smartwatch-im-hands-on-video/ (visited on 09/07/2015) (cit. on p. 202).

[156] *Lenovo unveils laser projector smartphone concept*. Android Central. URL: http://www.androidcentral.com/lenovo-unveils-smart-cast-smartphone-concept (visited on 09/19/2015) (cit. on p. 47).

[157] Leung, M., Tomitsch, M., Vande Moere, A., "Designing a personal visualization projection of online social identity." In: *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*. CHI EA '11. ACM, 2011, pp. 1843–1848 (cit. on pp. 44, 48, 162, 215).

[158] Lewis, J. R., "Psychometric evaluation of the post-study system usability questionnaire: The PSSUQ." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 36. SAGE Publications, 1992, pp. 1259–1260 (cit. on p. 13).

[159] Ljiljana Udovicic, "Damit nichts ins Auge geht... - Schutz vor Laserstrahlung." In: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) 2. Auflage. 2010 (cit. on p. 23).

[160] *Lumo Interactive Projector | Indiegogo*. URL: https://www.indiegogo.com/projects/lumo-interactive-projector#/story (visited on 06/06/2015) (cit. on pp. 52, 54).

[161] Lyons, K., Pering, T., Rosario, B., Sud, S., Want, R., "Multi-display Composition: Supporting Display Sharing for Collocated Mobile Devices." In: *Human-Computer Interaction – INTERACT 2009*. Red. by D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, and G. Weikum. Vol. 5726. Springer Berlin Heidelberg, 2009, pp. 758–771 (cit. on pp. 104, 129).

[162] *Magic Leap*. URL: http://www.magicleap.com (visited on 07/06/2015) (cit. on p. 6).

[163] Martinez Plasencia, D., Knibbe, J., Haslam, A. D., Latimer, E. J., Dennis, B., Lewis, G. J., Whiteley, M., Coyle, D., "ReflectoSlates: Personal Overlays for Tabletops Combining Camera-projector Systems and Retroreflective Materials." In: *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '14. ACM, 2014, pp. 2071–2076 (cit. on pp. 44, 50).

[164] Mathur, A., Ramachandran, D., Cutrell, E., Balakrishnan, R., "An Exploratory Study on the Use of Camera Phones and Pico Projectors in Rural India." In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '11. ACM, 2011, pp. 347–356 (cit. on pp. 44, 47, 50).

[165] Matoba, Y., Tokui, T., Sato, R., Sato, T., Koike, H., "SplashDisplay: Volumetric Projection Using Projectile Beads." In: *ACM SIGGRAPH 2012 Emerging Technologies*. SIGGRAPH '12. ACM, 2012, 19:1–19:1 (cit. on pp. 44, 51, 55).

[166] Matsuda, D., Sakata, N., Nishida, S., "Wearable input/output interface for floor projection using hands and a toe." In: *2013 23rd International Conference on Artificial Reality and Telexistence (ICAT)*. 2013 23rd International Conference on Artificial Reality and Telexistence (ICAT). 2013, pp. 122–128 (cit. on pp. 37, 44, 48, 55).

[167] Matsumoto, T., Hashimoto, S., Okude, N., "The embodied Web: embodied Web-services interaction with an umbrella for augmented city experiences." In: *Computer Animation and Virtual Worlds* 19.1 (1, 2008), pp. 49–66 (cit. on pp. 44, 50).

[168] Matsumoto, T., Horiguchi, D., Nakashima, S., Okude, N., "Z-agon: mobile multi-display browser cube." In: *CHI '06 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '06. ACM, 2006, pp. 351–356 (cit. on p. 103).

[169] McCallum, D. C., Irani, P., "ARC-Pad: absolute+relative cursor positioning for large displays with a mobile touchscreen." In: *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. UIST '09. ACM, 2009, pp. 153–156 (cit. on p. 60).

[170] McFarlane, D. C., Wilder, S. M., "Interactive dirt: increasing mobile work performance with a wearable projector-camera system." In: *Proceedings of the 11th international conference on Ubiquitous computing*. Ubicomp '09. ACM, 2009, pp. 205–214 (cit. on pp. 44, 47, 49, 52, 54).

[171] McGookin, D. K., Brewster, S. A., Christov, G., "Studying Digital Graffiti As a Location-based Social Network." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 3269–3278 (cit. on pp. 44, 50, 53).

[172] Miller, T., Stasko, J., "Artistically conveying peripheral information with the InfoCanvas." In: *Proceedings of the Working Conference on Advanced Visual Interfaces*. AVI '02. ACM, 2002, pp. 43–50 (cit. on p. 52).

[173] Mistry, P., Maes, P., Chang, L., "WUW-wear Ur world: a wearable gestural interface." In: *CHI '09 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '09. 2009, pp. 4111–4116 (cit. on pp. 36, 44, 48–51, 61, 64, 188).

[174] *Moject | Motion Projection by D&A*. URL: http://moject.com/ (visited on 06/17/2011) (cit. on pp. 52, 54).

[175] Molyneaux, D., Gellersen, H., Finney, J., "Cooperative Augmentation of Mobile Smart Objects with Projected Displays." In: *ACM Trans. Interact. Intell. Syst.* 3.2 (2013), 7:1–7:35 (cit. on p. 41).

[176] Molyneaux, D., Izadi, S., Kim, D., Hilliges, O., Hodges, S., Cao, X., Butler, A., Gellersen, H., "Interactive Environment-Aware Handheld Projectors for Pervasive Computing Spaces." In: *Pervasive 2012*. Springer, 2012 (cit. on pp. 32, 36, 40 sq., 44, 46, 51, 55, 212, 218).

[177] Morris, M. R., Wobbrock, J. O., Wilson, A. D., "Understanding Users' Preferences for Surface Gestures." In: *Proceedings of Graphics Interface 2010*. GI '10. Canadian Information Processing Society, 2010, pp. 261–268 (cit. on p. 14).

[178] Myers, B. A., "Using handhelds for wireless remote control of PCs and appliances." In: *Interacting with Computers* 17.3 (2005), pp. 251–264 (cit. on p. 135).

[179] Myers, B. A., Stiel, H., Gargiulo, R., "Collaboration using multiple PDAs connected to a PC." In: *Proc. CSCW 1998*. ACM Press, 1998, pp. 285–294 (cit. on p. 58).

[180] *Myo Gesture Control Armband - Take Control of Presentations and more*. URL: https://www.thalmic.com/myo/ (visited on 06/18/2015) (cit. on pp. 37, 41).

[181] *Next-gen Siri versus Google Now on Tap: The battle to read your mind best*. CNET. URL: http://www.cnet.com/news/proactive-siri-versus-google-now-on-tap-compared/ (visited on 07/01/2015) (cit. on p. 48).

[182] Ni, T., Karlson, A. K., Wigdor, D., "AnatOnMe: facilitating doctor-patient communication using a projection-based handheld device." In: *Proceedings of the 2011 annual conference on Human factors in computing systems*. CHI '11. ACM, 2011, pp. 3333–3342 (cit. on pp. 44, 50 sq.).

[183] Nielsen, M., Störring, M., Moeslund, T., Granum, E., "A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI." In: *Proc. GW 2003*. Springer, 2004, pp. 409–420 (cit. on p. 14).

[184] *Nikon | Imaging Products | Product Archive - COOLPIX S1200pj*. URL: http://imaging.nikon.com/lineup/coolpix/s/s1200pj/index.htm (visited on 07/06/2015) (cit. on p. 6).

[185] *Nomad*. In: *Wikipedia, the free encyclopedia*. 24, 2015 (cit. on p. 3).

[186] Norman, G., "Likert scales, levels of measurement and the 'laws' of statistics." In: *Advances in Health Sciences Education* 15.5 (1, 2010), pp. 625–632 (cit. on p. 13).

[187] Nozaki, H., "Flying Display: A Movable Display Pairing Projector and Screen in the Air." In: *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '14. ACM, 2014, pp. 909–914 (cit. on pp. 38, 51).

[188] Ochiai, Y., Hoshi, T., Rekimoto, J., "Pixie Dust: Graphics Generated by Levitated and Animated Objects in Computational Acoustic-potential Field." In: *ACM Trans. Graph.* 33.4 (2014), 85:1–85:13 (cit. on pp. 38, 44, 51, 55).

[189] Ochiai, Y., Kumagai, K., Hoshi, T., Rekimoto, J., Hasegawa, S., Hayasaki, Y., "Fairy Lights in Femtoseconds: Aerial and Volumetric Graphics Rendered by Focused Femtosecond Laser Combined with Computational Holographic Fields." In: *ACM SIGGRAPH 2015 Emerging Technologies*. SIGGRAPH '15. ACM, 2015 (cit. on p. 51).

[190] Okude, E., Kakehi, Y., "Rainterior: An Interactive Water Display with Illuminating Raindrops." In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ITS '11. ACM, 2011, pp. 270–271 (cit. on pp. 44, 55).

[191] Okumura, K., Oku, H., Ishikawa, M., "Lumipen: Projection-Based Mixed Reality for Dynamic Objects." In: *2012 IEEE International Conference on Multimedia and Expo (ICME)*. 2012 IEEE International Conference on Multimedia and Expo (ICME). 2012, pp. 699–704 (cit. on p. 30).

[192] Olivier, P., Cao, H., Gilroy, S. W., Jackson, D. G., "Crossmodal Ambient Displays." In: *People and Computers XX — Engage*. Springer London, 2007, pp. 3–16 (cit. on p. 52).

[193] Olsen, D. R., Nielsen, T., "Laser pointer interaction." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. CHI '01. ACM, 2001, pp. 17–22 (cit. on p. 60).

[194] Oswald, P., Tost, J., Wettach, R., "I.Ge: Exploring New Game Interaction Metaphors with Interactive Projection." In: *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*. TEI '15. ACM, 2015, pp. 733–738 (cit. on pp. 44, 50).

[195] Oyamada, Y., Saito, H., "Defocus Blur Correcting Projector-Camera System." In: *Advanced Concepts for Intelligent Vision Systems*. Lecture Notes in Computer Science 5259. Springer Berlin Heidelberg, 1, 2008, pp. 453–464 (cit. on p. 35).

[196] *Peripheral vision*. Wikipedia, the free encyclopedia. 25, 2015. URL: https://en.wikipedia.org/w/index.php?title=Peripheral_vision&oldid=668624753 (visited on 08/06/2015) (cit. on p. 21).

[197] Pinhanez, C., "The everywhere displays projector: A device to create ubiquitous graphical interfaces." In: *Proceedings of the 3rd international conference on Ubiquitous Computing*. Ubicomp '01. 2001, pp. 315–331 (cit. on p. 49).

[198] Plaue, C., Stasko, J., "Animation in a peripheral display: distraction, appeal, and information conveyance in varying display configurations." In: *Proceedings of Graphics Interface 2007*. GI '07. ACM, 2007, pp. 135–142 (cit. on p. 52).

[199] *Pointer Ballistics for Windows XP*. URL: http://msdn.microsoft.com/en-us/windows/hardware/gg463319.aspx (visited on 06/10/2012) (cit. on p. 65).

[200] Poon, A. D., Fagan, L. M., "PEN-Ivory: the design and evaluation of a pen-based computer system for structured data entry." In: *Proceedings of the Annual Symposium on Computer Application in Medical Care* (1994), pp. 447–451 (cit. on p. 91).

[201] Pouli, T., Subramanian, S., "Mobile projectors versus mobile displays: an assessment of task performance." In: *Proceedings of the ACM Symposium on Applied Perception*. SAP '12. ACM, 2012, pp. 103–110 (cit. on p. 49).

[202] Pousman, Z., Stasko, J., "A taxonomy of ambient information systems: four patterns of design." In: *Proceedings of the working conference on Advanced visual interfaces*. AVI '06. ACM, 2006, pp. 67–74 (cit. on p. 164).

[203] *Project Tango*. Google. URL: https://www.google.com/atap/project-tango/ (visited on 06/19/2015) (cit. on p. 40).

[204] Qin, Q., Rohs, M., Kratz, S., "Dynamic Ambient Lighting for Mobile Devices." In: *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology*. UIST '11 Adjunct. ACM, 2011, pp. 51–52 (cit. on pp. 44, 52, 55, 215).

[205] *QuickStart Game Engine*. URL: http://quickstartengine.codeplex.com/ (visited on 09/13/2012) (cit. on p. 176).

[206] Ramakers, R., Schöning, J., Luyten, K., "Paddle: Highly Deformable Mobile Devices with Physical Controls." In: *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 2569–2578 (cit. on p. 53).

[207] Rapp, S., "Spotlight Navigation: a pioneering user interface for mobile projection." In: *Proc. Ubiprojection 2010, Workshop on Personal Projection at Pervasive*. 2010, p. 4 (cit. on p. 38).

[208] Raskar, R., Baar, J., Beardsley, P., Willwacher, T., Srinivas, R., Forlines, C., "iLamps: geometrically aware and self-configuring projectors." In: *ACM Transactions on Graphics* (2003) (cit. on p. 31).

[209] Raskar, R., Beardsley, P., Van Baar, J., Wang, Y., Dietz, P., Lee, J., Leigh, D., Willwacher, T., "RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors." In: *ACM Transactions on Graphics (TOG)* 23.3 (2004), pp. 406–415 (cit. on pp. 44, 51).

[210] Raskar, R., Beardsley, P., "A self-correcting projector." In: *Computer Vision and Pattern Recognition*. Vol. 2. CVPR '01. IEEE Computer Society, 2001, pp. 504–508 (cit. on pp. 31 sq.).

[211] Rekimoto, J., Saitoh, M., "Augmented surfaces: a spatially continuous work space for hybrid computing environments." In: *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*. CHI '99. ACM, 1999, pp. 378–385 (cit. on pp. 44, 50).

[212] Robinson, S., Jones, M., "HaptiProjection: multimodal mobile information discovery." In: *Ubiprojection Workshop at Pervasive*. Vol. 10. 2010 (cit. on pp. 44, 52).

[213] Roeber, H., Bacus, J., Tomasi, C., "Typing in thin air: the canesta projection keyboard - a new method of interaction with electronic devices." In: *CHI '03 extended abstracts on Human factors in computing systems*. CHI EA '03. ACM, 2003, pp. 712–713 (cit. on pp. 40, 44, 47, 51).

[214] *RoomAlive Toolkit - Microsoft Research*. 2015. URL: http://research.microsoft.com/en-us/projects/roomalivetoolkit/ (visited on 09/19/2015) (cit. on p. 53).

[215] Rosenthal, S., Kane, S. K., Wobbrock, J. O., Avrahami, D., "Augmenting on-screen instructions with micro-projected guides: when it works, and when it fails." In: *Proceedings of the 12th ACM international conference on Ubiquitous computing*. Ubicomp '10. ACM, 2010, pp. 203–212 (cit. on pp. 44, 52, 55).

[216] Rukzio, E., Holleis, P., Gellersen, H., "Personal projectors for pervasive computing." In: *Pervasive Computing, IEEE* 11.2 (2012), pp. 30–37 (cit. on pp. 35, 43, 48, 58).

[217] Sahami Shirazi, A., Abdelrahman, Y., Henze, N., Schneegass, S., Khalilbeigi, M., Schmidt, A., "Exploiting Thermal Reflection for Interactive Systems." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 3483–3492 (cit. on p. 41).

[218] Sajadi, B., Gopi, M., Majumder, A., "Edge-guided Resolution Enhancement in Projectors via Optical Pixel Sharing." In: *ACM Trans. Graph.* 31.4 (2012), 79:1–79:122 (cit. on p. 33).

[219] Sakata, N., Konishi, T., Nishida, S., "Mobile Interfaces Using Body Worn Projector and Camera." In: *Virtual and Mixed Reality*. Lecture Notes in Computer Science 5622. Springer Berlin Heidelberg, 2009, pp. 106–113 (cit. on pp. 44, 51, 55).

[220] Samosky, J. T., Nelson, D. A., Wang, B., Bregman, R., Hosmer, A., Mikulis, B., Weaver, R., "BodyExplorerAR: Enhancing a Mannequin Medical Simulator with Sensing and Projective Augmented Reality for Exploring Dynamic Anatomy and Physiology." In: *Proceedings of the Sixth International Conference*

on Tangible, Embedded and Embodied Interaction. TEI '12. ACM, 2012, pp. 263–270 (cit. on pp. 44, 50, 52, 54).

[221] *Samsung Galaxy Beam2 - Full phone specifications*. URL: http://www.gsmarena.com/samsung_galaxy_beam2-6328.php (visited on 07/06/2015) (cit. on p. 6).

[222] Sarvas, R., Viikari, M., Pesonen, J., Nevanlinna, H., "MobShare: Controlled and Immediate Sharing of Mobile Images." In: *Proceedings of the 12th Annual ACM International Conference on Multimedia*. MULTIMEDIA '04. ACM, 2004, pp. 724–731 (cit. on p. 135).

[223] Sato, A., Watanabe, K., Rekimoto, J., "MimiCook: A Cooking Assistant System with Situated Guidance." In: *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*. TEI '14. ACM, 2013, pp. 121–124 (cit. on pp. 52, 54).

[224] Sato, M., Poupyrev, I., Harrison, C., "Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. ACM, 2012, pp. 483–492 (cit. on pp. 37, 41).

[225] Sato, Y., Kobayashi, Y., Koike, H., "Fast tracking of hands and fingertips in infrared images for augmented desk interface." In: *International conference on automatic face and gesture recognition*. IEEE Computer Society, 2000, pp. 462–467 (cit. on p. 40).

[226] Scheible, J., "Interactive Snow Sculpture Painting." In: *CHI '11 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '11. ACM, 2011, pp. 503–504 (cit. on pp. 44, 52).

[227] Schmidt, A., "Implicit human computer interaction through context." In: *Personal Technologies* 4.2-3 (1, 2000), pp. 191–199 (cit. on p. 38).

[228] Schmidt, D., Molyneaux, D., Cao, X., "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. ACM, 2012, pp. 379–388 (cit. on pp. 37, 41, 44, 51).

[229] Schmidt, D., Seifert, J., Rukzio, E., Gellersen, H., "A cross-device interaction style for mobiles and surfaces." In: *Proceedings of the Designing Interactive Systems Conference*. DIS '12. ACM, 2012, pp. 318–327 (cit. on pp. 103 sq.).

[230] Schöning, J., Löchtefeld, M., Rohs, M., Krüger, A., "Projector Phones: A New Class of Interfaces for Augmented Reality." In: *International Journal of Mobile HCI* (2010) (cit. on p. 48).

[231] Seah, S. A., Martinez Plasencia, D., Bennett, P. D., Karnik, A., Otrocol, V. S., Knibbe, J., Cockburn, A., Subramanian, S., "SensaBubble: A Chrono-sensory Mid-air Display of Sight and Smell." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. ACM, 2014, pp. 2863–2872 (cit. on pp. 44, 51, 55).

[232] Seifert, J., Dobbelstein, D., Schmidt, D., Holleis, P., Rukzio, E., "From the Private Into the Public: Privacy-Respecting Mobile Interaction Techniques for Sharing Data on Surfaces." In: *Personal and Ubiquitous Computing* (2013), pp. 1–14 (cit. on p. 50).

[233] Seifert, J., Simeone, A., Schmidt, D., Holleis, P., Reinartz, C., Wagner, M., Gellersen, H., Rukzio, E., "MobiSurf: improving co-located collaboration through integrating mobile devices and interactive surfaces." In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*. ITS '12. ACM, 2012, pp. 51–60 (cit. on pp. 84, 104).

[234] Sharp, T., Keskin, C., Robertson, D., Taylor, J., Shotton, J., Kim, D., Rhemann, C., Leichter, I., Vinnikov, A., Wei, Y., Freedman, D., Kohli, P., Krupka, E., Fitzgibbon, A., Izadi, S., "Accurate, Robust, and Flexible Real-time Hand Tracking." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. ACM, 2015, pp. 3633–3642 (cit. on p. 41).

[235] Shilkrot, R., Hunter, S., Maes, P., "PoCoMo: Projected Collaboration Using Mobile Devices." In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '11. ACM, 2011, pp. 333–336 (cit. on pp. 44, 48, 50, 52, 54).

[236] Shirmohammadi, B., Taylor, C. J., "Self-localizing smart camera networks." In: *ACM Trans. Sen. Netw.* 8.2 (2012), 11:1–11:24 (cit. on p. 109).

[237] Siriborvornratanakul, T., Sugimoto, M., "Clutter-aware Adaptive Projection Inside a Dynamic Environment." In: *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*. VRST '08. ACM, 2008, pp. 241–242 (cit. on p. 30).

[238] *Smartphone-Funktionen - Nutzung 2015 | Umfrage*. Statista. URL: http://de.statista.com/statistik/daten/studie/166150/umfrage/nutzung-von-smartphone-funktionen-in-deutschland/ (visited on 07/05/2015) (cit. on p. 4).

[239] Smith, A., *U.S. Smartphone Use in 2015*. Pew Research Center: Internet, Science & Tech. URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ (visited on 07/05/2015) (cit. on pp. 3 sq.).

[240] Sodhi, R., Benko, H., Wilson, A., "LightGuide: projected visualizations for hand movement guidance." In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. CHI '12. ACM, 2012, pp. 179–188 (cit. on pp. 44, 51, 55, 167, 187).

[241] Sodhi, R., Poupyrev, I., Glisson, M., Israr, A., "AIREAL: Interactive Tactile Experiences in Free Air." In: *ACM Trans. Graph.* 32.4 (2013), 134:1–134:10 (cit. on p. 42).

[242] Song, H., Grossman, T., Fitzmaurice, G., Guimbretiere, F., Khan, A., Attar, R., Kurtenbach, G., "PenLight: combining a mobile projector and a digital pen for dynamic visual overlay." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. ACM, 2009, pp. 143–152 (cit. on pp. 38, 44, 47, 51, 55).

[243] Song, H., Guimbretiere, F., Grossman, T., Fitzmaurice, G., "MouseLight: bimanual interactions on digital paper using a pen and a spatially-aware mobile projector." In: *Proceedings of the 28th international conference on Human factors in computing systems*. CHI '10. ACM, 2010, pp. 2451–2460 (cit. on pp. 38, 44, 47 sq., 51, 55).

[244] Soukoreff, R. W., MacKenzie, I. S., "Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI." In: *International Journal of Human-Computer Studies* 61.6 (2004), pp. 751–789 (cit. on pp. 67 sq.).

[245] Steimle, J., Jordt, A., Maes, P., "Flexpad: Highly Flexible Bending Interactions for Projected Handheld Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. ACM, 2013, pp. 237–246 (cit. on p. 41).

[246] Sugimoto, M., Dao, V. N., Seki, E., "An Adaptive Projection Technique for Monochromatic Non-flat Surfaces using a Handheld Projector–Camera System." In: *MP 2 Workshop CHI'11*. 2011 (cit. on p. 32).

[247] Sugimoto, M., Miyahara, K., Inoue, H., Tsunesada, Y., "Hotaru: intuitive manipulation techniques for projected displays of mobile devices." In: *Human-Computer Interaction-INTERACT 2005* (2005), pp. 57–68 (cit. on pp. 44, 50).

[248] Summet, J., Flagg, M., Cham, T. .-.-J., Rehg, J. M., Sukthankar, R., "Shadow Elimination and Blinding Light Suppression for Interactive Projected Displays." In: *IEEE Transactions on Visualization and Computer Graphics* 13.3 (2007), pp. 508–517 (cit. on p. 30).

[249] *Tablet Displays Under Bright Ambient Lighting Shoot-Out*. URL: http://www.displaymate.com/Tablet_Brightness_ShootOut_1.htm (visited on 07/03/2015) (cit. on p. 28).

[250] Torning, K., Oinas-Kukkonen, H., "Persuasive system design: state of the art and future directions." In: *Proceedings of the 4th International Conference on Persuasive Technology*. Persuasive '09. ACM, 2009, 30:1–30:8 (cit. on p. 171).

[251] *Touch Jet*. URL: http://www.touchjet.com (visited on 01/23/2014) (cit. on pp. 38, 52, 54).

[252] Treisman, A., "Preattentive processing in vision." In: *Computer Vision, Graphics, and Image Processing* 31.2 (1985), pp. 156–177 (cit. on p. 164).

[253] *Ultrahaptics*. Ultrahaptics. URL: http://ultrahaptics.com/ (visited on 06/20/2015) (cit. on p. 42).

[254] *Understanding Success Criterion 1.4.3 | Understanding WCAG 2.0*. URL: http://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-contrast.html (visited on 08/11/2015) (cit. on p. 28).

[255] Vartan Piroumian, *Nomadic Computing with Mobile Devices*. URL: http://www.cognizant.ch/InsightsWhitepapers/Nomadic-Computing-with-Mobile-Devices.pdf (cit. on p. 46).

[256] Vatavu, R.-D., "There's a World Outside Your TV: Exploring Interactions Beyond the Physical TV Screen." In: *Proceedings of the 11th European Conference on Interactive TV and Video*. EuroITV '13. ACM, 2013, pp. 143–152 (cit. on pp. 44, 50, 52, 54).

[257] Vatavu, R.-D., Wobbrock, J. O., "Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. ACM, 2015, pp. 1325–1334 (cit. on p. 14).

[258] Virolainen, A., Akerman, P., Häkkilä, J., "Burn-to-share: content sharing with mobile projectors." In: *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*. MUM '10. ACM, 2010, 18:1–18:4 (cit. on pp. 44, 47, 51, 217).

[259] Vogel, D., Balakrishnan, R., "Distant freehand pointing and clicking on very large, high resolution displays." In: *Proceedings of the 18th annual ACM symposium on User interface software and technology*. UIST '05. ACM, 2005, pp. 33–42 (cit. on p. 60).

[260] Waldner, M., Grasset, R., Steinberger, M., Schmalstieg, D., "Display-adaptive Window Management for Irregular Surfaces." In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ITS '11. ACM, 2011, pp. 222–231 (cit. on pp. 44, 50).

[261] *WalkAbout Projection - Wearable Projection Mapping - Projection Mapping Central*. URL: http://projection-mapping.org/walkabout-projection-wearable-projection-mapping/ (visited on 06/13/2015) (cit. on p. 53).

[262] Wecker, A., Kuflik, T., Lanir, J., Stock, O., "Evaluating Mobile Projectors As a Shared Display Option for Small Groups." In: *CHI '12 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '12. ACM, 2012, pp. 2543–2548 (cit. on p. 50).

[263] Weigel, M., Boring, S., Steimle, J., Marquardt, N., Greenberg, S., Tang, A., "ProjectorKit: Easing Rapid Prototyping of Interactive Applications for Mobile Projectors." In: *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services*. MobileHCI '13. ACM, 2013, pp. 247–250 (cit. on pp. 44, 53).

[264] Weiss, M., Wacharamanotham, C., Voelker, S., Borchers, J., "FingerFlux: Near-surface Haptic Feedback on Tabletops." In: *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*. UIST '11. ACM, 2011, pp. 615–620 (cit. on p. 42).

[265] Willis, K., "Ubiquitous Projection: New Interfaces using Mobile Projectors." In: *Dissertations* (19, 2013) (cit. on pp. 17, 43).

[266] Willis, K. D., "A Pre-history of Handheld Projector-based Interaction." In: *Personal Ubiquitous Comput.* 16.1 (2012), pp. 5–15 (cit. on pp. 17 sq.).

[267] Willis, K. D. D., Poupyrev, I., Shiratori, T., "Motionbeam: a metaphor for character interaction with handheld projectors." In: *Proceedings of the 2011 annual conference on Human factors in computing systems*. CHI '11. ACM, 2011, pp. 1031–1040 (cit. on pp. 18, 39 sqq., 44, 47, 49, 52, 54).

[268] Willis, K. D. D., Shiratori, T., Mahler, M., "HideOut: Mobile Projector Interaction with Tangible Objects and Surfaces." In: *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. TEI '13. ACM, 2013, pp. 331–338 (cit. on pp. 41, 44, 52, 55).

[269] Willis, K. D., Poupyrev, I., Hudson, S. E., Mahler, M., "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. ACM, 2011, pp. 431–440 (cit. on pp. 18, 41, 44, 48, 50 sqq., 54, 209, 216 sq.).

[270] Wilson, A. D., "PlayAnywhere: a compact interactive tabletop projection-vision system." In: *Proceedings of the 18th annual ACM symposium on User interface software and technology*. 2005, pp. 83–92 (cit. on pp. 37, 40, 44, 50, 54 sq., 64, 104, 120).

[271] Wilson, A. D., Benko, H., "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. ACM, 2010, pp. 273–282 (cit. on pp. 44, 46, 49, 84, 110).

[272] Wilson, A., Benko, H., Izadi, S., Hilliges, O., "Steerable augmented reality with the beamatron." In: *Proceedings of the 25th annual ACM symposium on User interface software and technology*. UIST '12. ACM, 2012, pp. 413–422 (cit. on pp. 30, 32, 44, 49).

[273] Wiltse, H., Nichols, J., "PlayByPlay: collaborative web browsing for desktop and mobile devices." In: *Proceedings of the 27th international conference on Human factors in computing systems*. CHI '09. ACM, 2009, pp. 1781–1790 (cit. on pp. 135, 137).

[274] Wishnitzer, R., Laiteerapong, T., Hecht, O., "Subcutaneous implantation of magnets in fingertips of professional gamblers—Case report." In: *The Journal of Hand Surgery* 6.5 (1981), pp. 473–474 (cit. on p. 42).

[275] Wobbrock, J. O., Aung, H. H., Rothrock, B., Myers, B. A., "Maximizing the Guessability of Symbolic Input." In: *CHI '05 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '05. ACM, 2005, pp. 1869–1872 (cit. on p. 14).

[276] Wobbrock, J. O., Morris, M. R., Wilson, A. D., "User-defined Gestures for Surface Computing." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. ACM, 2009, pp. 1083–1092 (cit. on p. 14).

[277] Wobbrock, J. O., Shinohara, K., Jansen, A., "The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. ACM, 2011, pp. 1639–1648 (cit. on p. 67).

[278] Wobbrock, J. O., Wilson, A. D., Li, Y., "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. ACM, 2007, pp. 159–168 (cit. on p. 122).

[279] Xiao, R., Harrison, C., Hudson, S. E., "WorldKit: rapid and easy creation of ad-hoc interactive applications on everyday surfaces." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. ACM, 2013, pp. 879–888 (cit. on pp. 44, 53).

[280] Xiao, R., Harrison, C., Willis, K. D., Poupyrev, I., Hudson, S. E., "Lumitrack: Low Cost, High Precision, High Speed Tracking with Projected M-sequences." In: *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*. UIST '13. ACM, 2013, pp. 3–12 (cit. on pp. 41, 51).

[281] Yamamoto, G., Sato, K., "PALMbit: A Body Interface Utilizing Light Projection onto Palms." In: *Inst. Image Information and Television Engineers* 61.6 (2007), pp. 797–804 (cit. on p. 187).

[282] Ye, Z., Khalid, H., "Cobra: flexible displays for mobilegaming scenarios." In: *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*. CHI EA '10. ACM, 2010, pp. 4363–4368 (cit. on pp. 52 sq.).

[283] Yee, K.-P., "Peephole Displays: Pen Interaction on Spatially Aware Handheld Computers." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '03. ACM, 2003, pp. 1–8 (cit. on p. 38).

[284] Yoshida, T., Hirobe, Y. Nii, H. Kawakami, N. Tachi, S. "Twinkle: Interacting with physical surfaces using handheld projector." In: Virtual Reality Conference (VR) (2010), pp. 87–90 (cit. on pp. 41, 44, 48, 51 sq., 54).

[285] Zhao, Y., Xue, C., Cao, X., Shi, Y., "PicoPet: "Real World" Digital Pet on a Handheld Projector." In: *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology*. UIST '11 Adjunct. ACM, 2011, pp. 1–2 (cit. on pp. 44, 48, 51 sq., 54).

[286] *ZTE Spro 2 (AT&T) - ZTE USA*. URL: http://www.zteusa.com/zte-spro2-265 (visited on 06/03/2015) (cit. on p. 22).

CURRICULUM VITAE

Christian was born on 20th of November, 1982 in Lemgo, NRW, Germany. From 1989 to 2002 he attended primary and high school. After that, he studied media informatics as a major and design as a minor at the University of Bielefeld and received his B.Sc. (with distinction) in 2007. Before and during the studies, he worked for Media Zone AG (2002–2004) as web developer and designer.

Christian continued his studies at the University of Duisburg-Essen where he graduated in Applied Informatics—Systems Engineering and received his M.Sc. degree in 2010. In his master thesis, he developed a social web platform of phones called sense-sation. The thesis was supervised by Prof. Dr. Albrecht Schmidt, who led the Pervasive Computing group in which Christian worked first as student assistant (2009) and later as research associate (2010).

In late 2010, Christian then started to work as a research associate in the Mobile HCI group led by Prof. Dr. Enrico Rukzio, fist located at the University of Duisburg-Essen and since 2012 at Ulm University, where he received his doctorate (with distinction) in January 2016. During his doctoral studies, he further worked as research intern and contractor for Microsoft Research Cambridge in the UK. In October 2015, he started working in the Advanced Visual Solutions group at Daimler Protics GmbH.

Christian regularly published his work at the major international HCI conferences such as ACM CHI, ACM UIST, and ACM ITS. His work was awarded Best Note Award at ACM ITS and Honorable Mention Award at ACM CHI and also granted a patent. From 2011 to 2015 he further regularly served as reviewer and PC member in (inter)national conferences on HCI and Ubiquitous Computing.

His primary focus is to naturally integrate mobile computing interfaces into everyday life by leveraging new display and interaction technology.