Advanced User Interfaces for Spatial Information

Johannes Schöning

Department of Computer Science

Saarland University

A thesis submitted for the degree of

Doktor-Ingenieur – Dr.-Ing.

February 2009

Dean: Prof. Dr. Joachim Weickert

1. Reviewer: Prof. Dr. Antonio Krüger

2. Reviewer: Dr. Michael Rohs

Head of Committee: Prof. Dr. Dr. h.c. mult. Wolfgang Wahlster

Committee member: Dr. Jörg Baus

Day of defense: $30^{th} March, 2010$

Abstract

For 6000 years, humans have used maps to navigate through space and complete other spatial tasks. These depictions of the Earth or other phenomenon (e.g the Moon and cognitive maps) as a set of symbols and at a scale, whose representative fraction was less than 1:1, were created by humans to support daily activities. For nearly all of these 6000 years, maps were drawn or printed on a piece of paper or on material like stone or papyrus. In the 1960s, the digital map arrived. At the beginning of the digital map revolution, the desktop terminal had an effective monopoly on the display and interaction with digital maps. Recently, devices spanning the entirety of Mark Weiser's ubiquitous computing spectrum – from tiny cell phones to enormous wall-sized displays – are beginning to be used for digital mapping activities. However, as this transition has occurred, it has become obvious that the map user interfaces implemented on desktop computers fail to function in these new ubiquitous computing contexts.

In this PhD thesis, we investigate new user interfaces that allow intuitive interaction with digital spatial information on a wide range of ubiquitous computing devices. We identify user interfaces (UI) problems, design and build innovative new UIs, and test whether these new UIs contributed solutions to the problems. All different sizes of devices are considered, from mobile augmented reality interfaces to mobile projection interfaces to large-scale interactive multi-touch surfaces. New findings about how people interact with analogue spatial information in combination with digital spatial information are derived. All interfaces follow the core design principle of *simple is fun*. Digital spatial information will play an increasingly larger and larger role in many of our day-to-day and longer-term decisions. Therefore, good user interfaces are needed to interact with this information.

Kurzfassung

Seit über 6000 Jahren benutzen Menschen Karten um sich zu orientieren. Diese Darstellungen der Erde oder anderer Phänome (z. B. die Darstellung des Mondes oder kognitive Karten) als eine Zusammenstellung von Symbolen und mit einem Maßstab von weniger als 1:1, wurden von Menschen geschaffen, um tägliche Aktivitäten zu unterstützen. Fast die gesamten 6000 Jahre bestanden Karten aus Papier, Papyrus oder Stein. Ab den 1960ziger hatten digitale Karten ihren Anfang. Am Begin der "digitalen Revolution" wurden diese allein auf Desktop-Computer dargestellt. Der Technologie Trend "Ubiquitous Computing", von Mark Weiser in den 1990zigern skizziert, und die damit verbundene Miniaturisierung von Computertechnologie erlaubt es heute, Nutzern eine Vielzahl von oft unterschiedlich großen Geräten für die Kartendarstellung bereitzustellen. An diesem Wendepunkt ist es offensichtlich, dass die Benutzerschnittstellen von Desktop Computern nicht auf diesen neuen Gerätenklassen übertragen werden können.

In dieser Doktorarbeit werden neue Benutzerschnittstellen für diese Geräteklassen betrachtet, die es erlauben intuitive mit räumlichen Daten zu interagieren. Es werden Probleme der heutigen Benutzerschnittstellen aufgezeigt, neue und innovative Benutzerschnittstellen gestaltet und entwickelt. Es wird getestet ob diese die Probleme beheben können. Dabei werden Benutzerschnittstellen für alle Gerätegrößen betrachtet, angefangen von mobilen Augmented Reality Benutzerschnittstellen über mobile Projektion Benutzerschnittstellen, zu großen und interaktiven multi-touch Oberflächen. Alle vorgestellten Benutzerschnittstellen folgen der Grundidee *"simple is fun"*.

"Bitte oszillieren Sie!
Pingpong ohne Hierarchie!
Bitte oszillieren Sie!
Ich bitte Sie: Genießen Sie!"

from *Bitte oszillieren Sie!* by Tocotronic

Contents

List of Figures

List of Tables

1

Introduction & Motivation

"This spatial stuff is going to get bigger, faster than anybody thinks." Kevin Kelly, former WIRED[1] editor.

"Maps become a user interface to many things [...] Geography is another way, a different way, to organize information." Financial Times, issue May 21^{th} 2008.

Both of these quotes are illustrative of the growing importance of digital spatial information for our daily lives. While neither are rooted in firm statistics, both capture the fact digital spatial information is undeniably a part of the zeitgeist of our times. In the past decade, digital maps, geotagged photos, virtual globes and other forms of digital spatial information have become enormously powerful cultural and economic forces. However, while the accessibility of this information has drastically increased, its basic interaction paradigms have not improved at the same rate. In the thesis that follows, we address this problem by designing, implementing, and experimentally validating entirely new interaction paradigms and user interfaces for digital spatial information.

We define *spatial information* as information that is described in a reference system. In spatial information the focus of this thesis the reference system normally covers the range of our world, i.e. latitude and longitude[2]. An extremely popular and ubiquitous

[1]Wired is a monthly American magazine and on-line periodical, published since March 1993, that reports on how technology affects culture, the economy, and politics see `http://www.wired.com/`.

[2]A structured and brief description of references systems is given by Werner et al. [240].

Figure 1.1: Map of Çatalhöyük in Turkey, complete with erupting "volcano" in the background, is know as the world's oldest map (6200 BC); (located in the Museum at Konya, Turkey) (left). "Imago Mundi" world map of the 6th century BC Babylonia (located in the British Museum, UK) (right). Harley and Woodward[83] provide further details.

interface to spatial information is the *map*. Following our definition, we can define a map as a visual representation of spatial information for a certain region of a coordinate system. Symbolic depiction is used to highlight relationships between information elements of that space such as objects, regions, and themes.

Humans have used maps as abstract representations of space for more than 6000 years, mainly to facilitate orientation and navigation in different scenarios [131] (see figure 1.1). Various researchers investigated the different functions of maps. Freitag [67] distinguish, based on the work of Papay [166] and Board [23], four functions of maps. Freitag defines [67]:

- The cognitive function encompasses all processes and operations and all models which generate and enhance spatial knowledge. All processes of map analysis, transformations, generalization, simulations, animations, etc. should be listed here, if possible in a sequence of operations leading from near-reality models to very abstract models of space.

- The communication function, which includes demonstration, encompasses all processes and operations of spatial knowledge transfer from a map maker to a user. It may be divided into several sub-functions according to the extent of transferred knowledge, the level of pre-knowledge, and the form and means of knowledge transfer. Educational communication, mass media communication, academic communication, administrative communication represents the dimensions of this function.

- The decision support function encompasses all processes and operations which – based on the evaluation of spatial phenomena – result in spatial decisions and spatial actions. Examples of these types of functions include navigation, planning, and persuasion.

- Social functions: The social function encompasses all processes which result not in spatial, but in social behavior and actions. One form of this involves the professional map maker in relation to other persons in the mapping process, including the users. Maps can also be seen as tools of social power, exercised through the access or the denial of access to spatial information, through copyrights or the monopoly on mapping equipment. Furthermore there is the ability to consider mapping as a cultural activity.

Creating these abstractions and mappings of real environments into new abstract spaces involve lots of human interaction and human intelligence [171]. Today map design is a complex task. Robinson, an American cartographer, claimed, when considering all aspects of cartography, that “map design is perhaps the most complex” [111; 190]. In addition, Robinson stated that a map must be designed explicitly for its purpose and its desired audience.

1.1 Problem

Recently, maps have been changing from static information products printed on paper to dynamic digital representations of space displayed on various kinds of interactive electronic devices such as mobile devices or large-scale interactive displays. Building user interfaces (UIs) for this new generation of maps is far from a trivial pursuit owing to the aforementioned complexity of map design.

A key paradigm that has informed digital map design is ubiquitous computing. Ubiquitous computing (Ubicomp) is a post-desktop paradigm of human-computer interaction [239] in which information processing has been thoroughly integrated into everyday objects and activities. In the course of ordinary activities, users of ubiquitous computing engage many computational devices and systems simultaneously, and may not necessarily even be aware that they are doing so. According to Weiser [239], the first wave of ubiquitous computing brings us three different size classes of computing devices. The first class includes tiny computers (pads), with the main goal of simulating analog Post-Its. The second describes so-called "notepads" (tabs), envisioned not as a personal computer but as analogous to scrap paper to be grabbed and used easily, with many in use by a person at once. The third class is wall-sized interactive surfaces (boards) (see figure 1.2). He also claims that the most technologies will *disappear*: "They weave themselves into the fabric of everyday life until they are indistinguishable from it" and defined other properties of Ubicomp in his article [239]. Transferring UIs from the desktop computing area to ubiquitous computing often fails, and standard interaction techniques cannot be applied for these new kinds of devices. As such important new UI questions emerge, especially in the difficult domain of spatial information.

The combination of digital spatial information with devices from the ubiquitous computing paradigm is potentially hugely beneficial for users, but numerous major challenges arise in the process. As such, paper maps are still quite superior in some categories to their digital counterparts. For instance, they provide high resolution at a large-scale with zero power consumption. In this thesis we focus on user aspects of theses interfaces, because we think, that the user aspects are very important to design seamless transitions between spatial information, devices (with user interfaces) and users. The device should preserve the functions of spatial information as highlighted earlier.

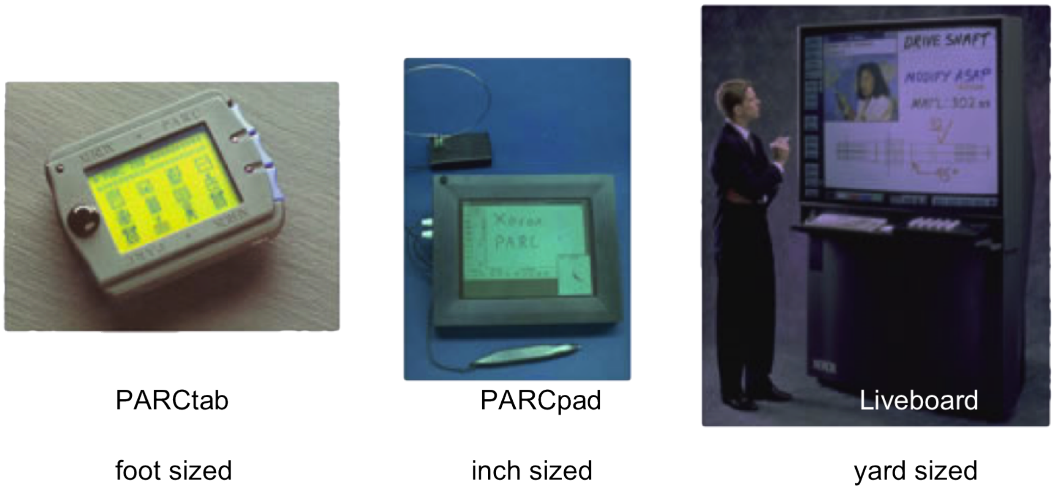

Figure 1.2: Weiser's device classes ranging from tab(foot)- sized [236] to pad(inch)-sized [107] to board(yard)-sized interfaces [58]. All different prototypes were developed at Xerox PARC form 1988 to 1996.

1.2 Goals

In this PhD thesis, we investigate new ubiquitous computing user interfaces that allow intuitive interaction with digital spatial information transferring and/or preserving the physical reality. Rather than borrowing everything from interaction with physical paper maps, we close the gap between analogue and digital maps from both sides - even going in the opposite direction (e.g. labelling a light switch with *1* and *0* as proposed by Gentner [70]). UIs for all the size classes of the 'first wave of Ubicomp" are considered, including handheld devices (pads), mobile augmented reality interfaces, mobile projection interfaces, mobile interface for tab-sized devices and even large scale interactive multi-touch surfaces (boards). Our core design idea for Ubicomp UIs is *simple is fun.*

1.3 Method

This thesis follows the method a of deployment-based research [50] and also described and used in the CASCO project[1] with a strong focus on the user (user-centered de-

[1] `http://www.comp.lancs.ac.uk/~fittond/index-old.html`

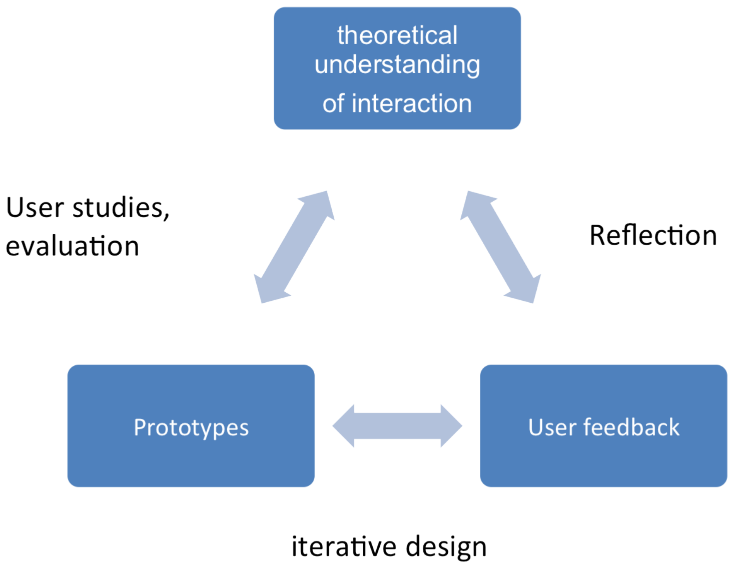

Figure 1.3: Deployment based research cycle adapted from CASCO project description[1].

sign). This approach will is based on a combination of theoretical research, collection of empirical data sets (e.g. arising from use of cultural probes) and prototyped application development as illustrated in figure 1.3. This methodology involves a tight cycle where theoretical issues and understanding, developed through reflection on empirical observations, are used to design deployed systems that test and explore the theory. These deployed systems then create a new context for observation of user behavior and thus lead to fresh insights, discoveries and refinement of theoretical understanding. Deployment-based research serves to gain user insight as well as technical insight [157]. The observation and involvement of users will serve the dual purpose of traditional user centered design and source for more theoretical analysis. Put more simply, operating within the spatial information domain, we identify UI problems, design and build innovative new UIs, and test whether these new UIs contributed solutions to the problems. This combination of technical and user insights makes this work unique in the spatial domain. For doing so we use classical methods that are used in the HCI community [175]. In summary, operating within the spatial information domain, we identify UI problems, design and build innovative new UIs, and test whether these new UIs contributed solutions to the problems.

1.4 Thesis Outline

The remainder of this thesis is structured as followed. The second chapter contains a complete survey of related work. In the following chapter, advanced UIs for pad-sized devices are discussed, namely Mobile Map Interaction (MMI), mobile projection interfaces and the PhotoMap application. This chapter is followed by a chapter describing interfaces for tab-sized interfaces. Next, board-sized interfaces are discussed. While research contributions (theoretical and applied) and statements of future work specific to individual projects are discussed in the context of the write-up of each project, the thesis is concluded with a statement of the three overarching theoretical and applied contributions as well as broader themes of future work.

Portions of this thesis have been published in journals, conference and workshop articles as well as research deliverables and patents, most notably [204; 208; 209; 211; 213; 215].

1.4.1 Notes on Writing Style

To improve the readability of the thesis:

- Neutral persons will be referred to using the female pronoun only (she instead of, e.g., she or he).
- All links are placed as footnotes directly into the text. All links were last visited on December 24^{th}, 2009).
- All links to companies, products, websites and others are listed only at first mention.
- Trademarks symbols are not used.
- The parts of the thesis are named as follows: a single number X denotes a chapter, X.X is referred to as a section and all other parts like X.X.X or X.X.X.X are called subsections.

2

Background and Related Work

In this chapter, background and general related work is presented, focusing on (mobile) augmented reality (AR) in the spatial domain. First, the term AR is defined. This is followed by a section centered on AR's role in the field of ubiquitous computing. In addition a classification and demarcation of this work is presented. In the third part of this chapter, related work in the area of AR and maps is presented. This area is highlighted because it is the most central to the work presented in this thesis. This thesis is also more peripherally related to research in other areas such as natural user interfaces, mobile multi-touch, and mobile human computer interaction (HCI) in general. These topics are covered in the final section of this chapter.

2.1 Augmented Reality Definition

AR is an extension of a user's perception with virtual information. It is defined by the following three essential characteristics:

- real and virtual elements must be combined,
- interactivity must occur in real-time and,
- operation must occur in three-dimensions [4].

AR presents the opportunity for many new approaches and interfaces, especially for geospatial information. Central to the power of AR in the geospatial context is that spatial information can be directly displayed "on the spot", and the interaction can take place in a simple and intuitive way [4]. In contrast to Virtual Reality (VR),

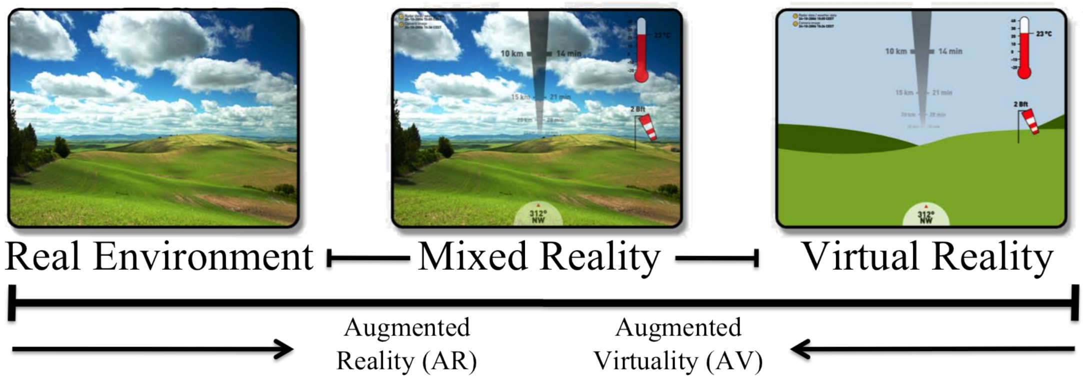

Figure 2.1: Milgram's Reality-Virtuality continuum [142], which is a scale ranging from the completely real (left) to the completely virtual (right). The area between the two extremes, where both the real and the virtual are mixed, is referred to as Mixed Reality. This can be in turn further subdivided into Augmented Reality, where the virtual augments the real, and Augmented Virtuality, where the real augments the virtual (left).

which completely immerses a user in a computer-generated environment, AR adds information to the user's view and thereby allows the user to experience both real and virtual information at the same time. AR has close connections to the fields of VR and Mixed Reality (MR), and Augmented Virtuality (AV). The Reality-Virtuality continuum of Milgram et al. [142] provides a good overview (see figure 2.1) of the relationships between these related approaches, separating the MR section into the two subsections of AR and AV. Since the term "Augmented Virtuality" is rarely used nowadays, AR and MR are now sometimes used as synonyms. The continuum has been somewhat incorrectly described as a concept in new media and computer science, when in fact it should belong closer to anthropology.

2.2 Classification & Demarcation

Weiser stated that Ubicomp is roughly the opposite of virtual reality [239]. However, when one considers that VR is merely at one extreme of the Reality-Virtuality Continuum postulated by Milgram, then one can see that Ubicomp and VR are not strictly opposite one another but rather orthogonal as described and illustrated by Newman et al. [155]. This new dimension was named by Newman the "Weiser's Continuum" and has Ubicomp at one extreme and the concept of terminal-based computing at the

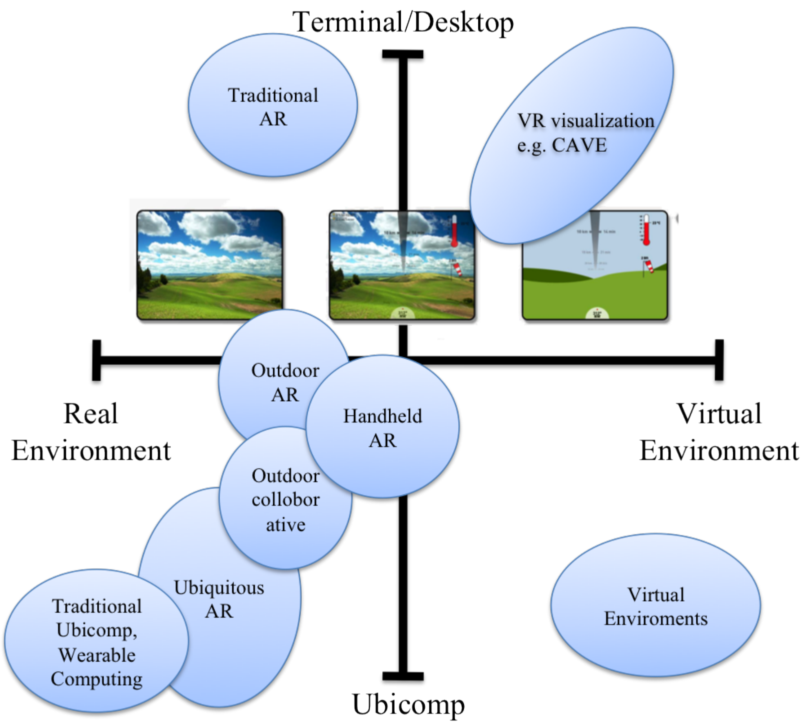

Figure 2.2: The Milgram-Weiser continuum of Newman et al. [155].

other. The terminal is the antithesis of the Disappearing Computer and a palpable impediment to intuitive interaction between user and computing environment. Placing both continua, the Reality-Virtuality (see figure 2.2) and the "Weiser's Continuum" at the two axes opens a 2D space shown in figure 2.2, in which different application domains represent areas in this space.

Maps – both paper maps and virtual maps – are a widespread medium deployed in many recent applications, especially location based systems (LBS) [69]. Analogous to Milgram's Reality-Virtuality continuum, maps can have a varying degree of realness in the spectrum from reality to virtuality. This area in between both endpoints, using maps to augment a visual representation of an area is an exciting category of AR interfaces and applications. For realizing this class of AR interfaces, different methods are used. The most commonly used are the video-see-through-display, the optical-see-through-display, and the projection-AR-display, which are explained later.

We will concentrate on the third quadrant because we are interested in applications that enhance the real environment with registered virtual information overlays, especially for geographic information systems, mobile mapping services and a variety of other geospatial applications. There is increasing interest in linking AR with cartography and the geospatial domain. For example, Schmalstieg and Reitmayr describe how to employ AR as a medium for cartography [202].

When not only looking at the development of technology and we should also look at the development of user interface paradigms. Looking back at the history of user interfaces, the first computer interfaces were command line user interfaces (CLI) in which users enter text to interact with the computers. These interfaces were around for a long time, and are still quite common today. The Graphical User Interface (GUI) ended the domination of CLI, which is currently the most-used user interface. However, it is predicted that soon more and more so-called natural user interfaces (NUI) will arise. As we will enter the next phases of Ubicomp, these NUIs may replace most GUIs in the next few decades. In this thesis, many of the interfaces introduced and investigated are NUIs that allow intuitive interaction with digital spatial information transferring and/or preserving the physical reality.

The characteristics of NUIs are reflected in the design and the development of the user interfaces throughout this thesis without reusing the term NUI again and again. We frequently adopted the highlighted NUI paradigm in our prototypes because it is often the optimal choice for geospatial interfaces in the third quadrant of the Milgram-Weiser diagram for all Ubicomp device classes.

2.3 Augmented Reality & Maps

With the advent of both pervasive environments and ubiquitous computing infrastructures, AR interfaces are evolving in various application domains. AR provides a set of methods to enhance the real environment with registered virtual information overlays, which has promising applications in many domains. Several approaches for AR interfaces have been successfully implemented ranging from projector-based to video-see-through metaphors.

AR has a very close relationship to geovisualization in that AR systems deal with large volumes of inherently spatial data. The development of scalable AR systems therefore draws on many infrastructures and algorithms from GIScience (e.g. efficient management and retrieval of spatial data, precise positioning, etc.). As such, it is not surprising that the use of AR as a user interface paradigm also has great potential for the development of novel geospatial applications: most importantly, AR provides intuitive mechanisms for interaction with spatial data and bridges the gap between the real environment and abstracted map representations in a way that can be flexibly adapted according to the current requirements of the users.

In this thesis section, we give an overview of the actual state of the art on geospatial AR technology and provide a categorization of different geospatial AR applications from a theoretical and technical point of view. We highlight in particular one of the most important geospatial AR application development areas: augmented maps and augmenting the real world with spatial information. In addition we intend to extrapolate current developments in AR to future trends over the next decade. It is structured as follows. First, we introduce the fields of AR and ubiquitous computing and highlight the connection to the geospatial domain. Section two describes various AR displays and tracking approaches that allow the augmentation of real environments with registered virtual information overlays especially for GIS, mobile mapping services and a variety of other applications. In section three the role of AR as an interface for Web & GIS services is discussed. Next, we summarize in section four key visualization and interaction techniques. The state of the art is discussed in section five. Finally, section six provides some concluding remarks and outlines future AR interfaces and technical and interaction challenges of the next generation AR interfaces in the geospatial domain.

2.3.1 Augmented Reality Displays

Generally, AR displays can be split into head-mounted displays (HMD), handheld displays and projection displays, the latter being stationary but potentially able to accommodate multiple users [202]. Also, for image generation and merging with the real world, two approaches can be distinguished [202]: optical see-through systems, which allow the user to see through the display onto the real world, and video see-through systems, which use video cameras to capture an image of the real world and provide the user with an augmented video image of her environment. As a result, five major classes of AR can be distinguished by their display type and their merging approach: optical see-through HMD AR, video see-through HMD AR, handheld display AR, projection-based AR with video augmentation, and projection-based AR with physical surface augmentation.

In the last decade, AR has grown out of infancy and is continuously showing its applicability and usefulness in today's society for various application domains including engineering, tourism or architecture. Additionally, AR is emerging with web based technologies thanks to the concurrent improvement in wireless technologies. The web provides an unprecedented amount of content with a location or geospatial component to serve for registered overlays in the user's view. In order to register, or align, virtual information with the physical objects, AR requires accurate position and orientation tracking. Therefore, tracking devices are necessary to deliver all Six Degrees of Freedom (6DOF) accurately to determine the location of a user and her orientation. A wide range of tracking technologies exists. A widely adopted technique for AR is optical tracking, which uses video cameras and advanced computer vision software to detect targets, so called markers, in the camera image and calculates their position and orientation. Furthermore, markerless tracking approaches rely on detecting natural features in the environment and do not need any physical infrastructure. The increasing availability of video cameras in today's computer devices has led to their use as means for tracking the position and orientation of a user as described by Klein in more detail [112]. Optical tracking can be applied in both indoor and outdoor environments, but is not the only form of tracking available. For outdoor applications, the global positioning system (GPS) is the predominant tracking system for delivering position estimates. Usually, GPS receivers are combined with inertial trackers and magnetic compasses, which deliver the orientation of the user. For indoor environments, infrared (IR) or ultra wide band (UWB) and other sensors such as electromagnetic-trackers (e.g. Polhemus) can be used for tracking the user. However, these approaches require the physical preparation of the environment with sensors. Finally, various combinations of tracking technologies have been integrated in hybrid tracking approaches for AR.

This section briefly overviews different AR displays by describing possible hardware configurations. The categorization of Bimber and Raskar [21] (see figure 2.2 (right)) illustrates the different possibilities of where the image can be formed, where the displays are located with respect to the observer and the real object, and what type of image is produced (i.e., planar or curved). Our goal is to place them into different

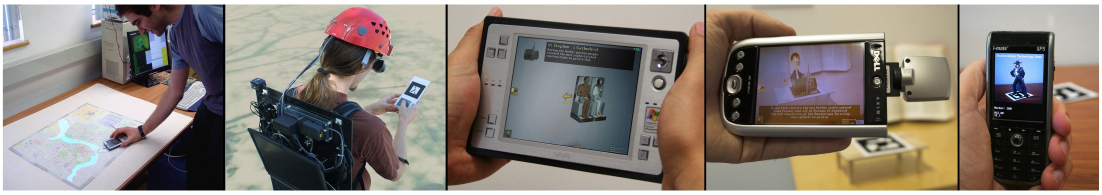

Figure 2.3: Fixed projection-based AR display and the evolution of mobile AR hardware from backpack systems to Ultra-Mobile PCs, personal digital assistants and mobile phones (from left to right).

categories so that it becomes easier to understand the state of the art and to help to identify new directions of research.

2.3.1.1 Projection-based AR Displays

Projection-based AR with video augmentation uses video projectors to display the image of an external video camera augmented with computer graphics on the screen whereas Projection-based AR with physical surface augmentation projects light onto arbitrarily shaped real world objects. It uses the real world objects as the projection surface for the virtual environments. Ordinary surfaces have varying reflectance, color, and geometry. Limitations of mobile devices, such as low resolution and small field of view, focus constraints, and ergonomic issues can be overcome in many cases by the utilization of projection technology. Thus, applications that do not require mobility can benefit from efficient spatial augmentations.

Projection-based AR with physical surface augmentation has applications in industrial assembly, product visualization, etc. Examples range from edutainment in museums (such as storytelling projections onto natural stone walls in historical buildings) to architectural visualizations (such as augmentations of complex illumination simulations or modified surface materials in real building structures). Both types of the projection-based AR are also well suited to multiple user situations.

The recent availability of cheap, small, and bright projectors has made it practical to use them for a wide range of applications such as creating large seamless displays and immersive environments. By introducing a camera into the system, and applying techniques from computer vision, the projection system can operate taking its environment into account. For example, it is possible to allow users to interact with the

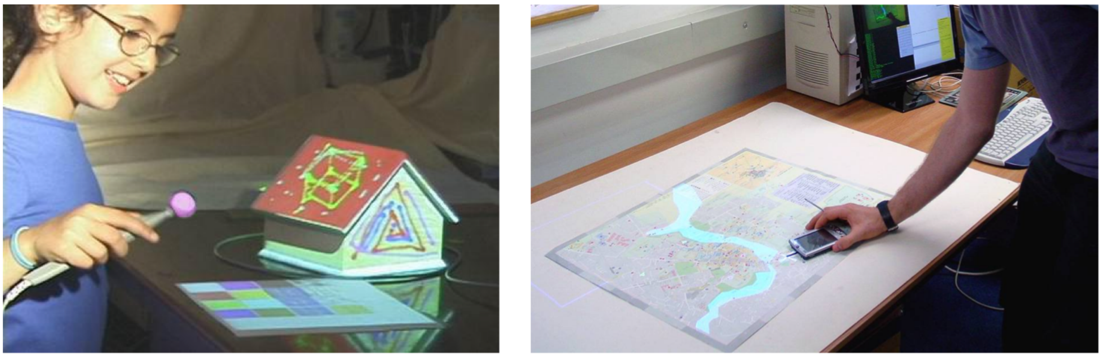

Figure 2.4: The *shader lamps* project is using projection technology to change the original surface appearance of real world object [179] (left). A AR flood control application for the city of Cambridge (UK) to demonstrate possible features of augmented maps, in which a map of interest is augmented with an overlaid area representing the flooded land at a certain water level [185] (right).

projected image, thus creating *projected interfaces*. The idea of *shader lamps* [179] is to use projection technology to change the original surface appearance of real world objects (see figure 2.4 left). A new approach to combine real TV studio content and computer-generated information was introduced in the Augmented Studio project. For this purpose projectors are used as studio point light sources. This allows the determination of camera pose or surface geometry and enables the real-time augmentation of the video stream with digital content [20].

An example of large spatially augmented environments is the *Being There* project, where a walk-through environment is constructed with styrofoam blocks and is augmented by projecting view-dependent images. Thus a realistic simulation of the interior of a building can be realized. Since the user is able to walk around in the augmented building a strong sense of immersion can be provided. To allow the user to freely move around in the setup, a wide area tracking system (*3rdTech's HiBall*) is used to track the head position [128].

Reitmayr et al. [185] have implemented a flood control application for the city of Cambridge (UK) to demonstrate possible features of augmented maps, in which a map of interest is augmented with an overlaid area representing the flooded land at a certain water level (see figure 2.4 right). The overall system centers around a table top environment where users work with maps. A camera mounted above the table tracks

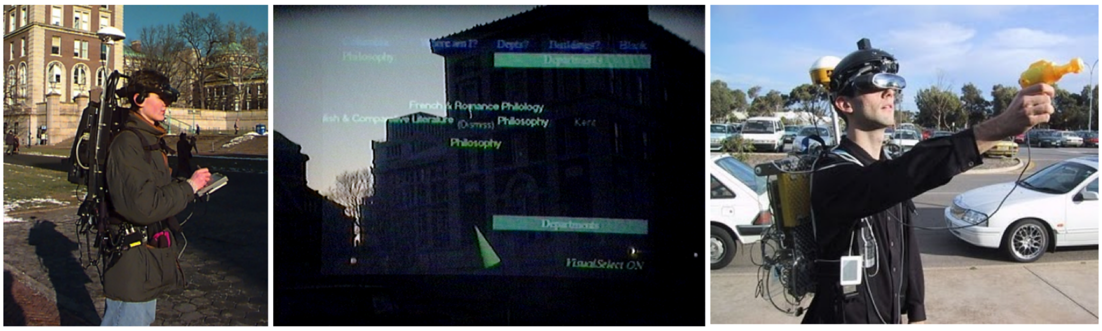

Figure 2.5: First "mobile" AR systems: The Touring Machine used backpacks with laptop computers and head-mounted displays [62] (left and middle). Similar setup was later used by the *Tinmith* system [172] (right).

the maps' locations on the surface and registers interaction devices placed on them. A projector augments the maps with projected information from overhead. Tracking and localization is done via visual matching of templates, which are stored for each map. Moreover, with the increasing compactness of modern projectors, new and more flexible possibilities of their usage arise. For example, miniaturized handheld projectors can be combined with mobile AR interfaces serving as output device.

2.3.1.2 HMD-based AR Displays

Head-mounted displays (HMDs) are usually worn by the user on her head and provide two image-generating devices, one for each eye. Optical see-through HMD AR uses a transparent HMD to blend together virtual and real content. Prime examples of an optical see-through HMD AR system are various augmented medical systems [3]. Video see-through HMD AR uses an opaque HMD to display merged video of the virtual environment with and view from cameras on the HMD. By overlaying the video images with the rendered content before displaying both to the user, virtual objects can appear fully opaque and occlude the real objects behind them. The drawback of video-based systems is that the viewpoint of the video camera does not completely match the user's viewpoint [202]. This approach is a bit more complex than optical see-through AR, requiring proper location of the cameras. For safety reasons, these systems cannot be used in applications where the user has to walk around or perform complex or dangerous tasks, since judgment of distances is distorted. Early work on mobile AR, such as the Touring Machine [62] used backpacks with laptop computers and head-mounted displays Höllerer et al. [93] built a series of *Mobile AR systems* (MARS) prototypes, starting with extensions to the *Touring Machine* from Feiner et al. (see figure 2.5). Similar AR prototypes have been built by Piekarsky et al. in form of the *Tinmith system* [172] (see figure 2.5). The *Tinmith-Metro* application is the main application, demonstrating the capture and creation of 3D geometry outdoors in real-time, leveraging the user's physical presence in the world. Furthermore, systems such as Signpost, which is a prototypical AR tourist guide for the city of Vienna, have been built by Reitmayr and Schmalstieg [186] allowing for indoor/outdoor tracking, navigation and collaboration based on hybrid user interfaces (2D and 3D). For tracking the mobile user usually GPS, inertial sensors and marker based tracking is applied. However, these systems are rather cumbersome for mobile applications deployed over longer working periods.

2.3.1.3 Handheld AR Displays

With the advent of handheld devices featuring cameras, the video-see-through metaphor has been widely adopted for mobile AR systems [202]. Devices that use the video-see-through metaphor can be built from tablet PCs, Ultra-Mobile PCs, and of course phones with cameras such as the iPhone. Therefore, recently handheld display AR have become very popular and can be potentially used in ubiquitous computing, such as location based services (LBS).

Ultra-Mobile PC displays This approach was originally conceived by Fitzmaurice and Buxton [63], and later refined into a see-through AR device by Rekimoto [187] (see figure 2.6). UMPCs are basically small mobile PCs running standard operation systems. A number of researchers have started employing them in AR simulations such as Wagner et al. [235], Newman et al. [154]. At this times the most popular UMPCs for AR was the *Sony Vaio* U70 and UX180, as well as *Samsung* Q1. Elmqvist et al. [57] have employed the wearable computer *Xybernaut* Mobile Assistant, which, although shares some common characteristics with UMPCs, does not belong in the UMPC category. This has started a strong trend towards handheld AR [234]. Handheld AR prototype devices of this category have been designed and built, for example, by [231]. The tracking approaches are similar to that of HMD-based AR setups. Moreover, Reitmayr et al.

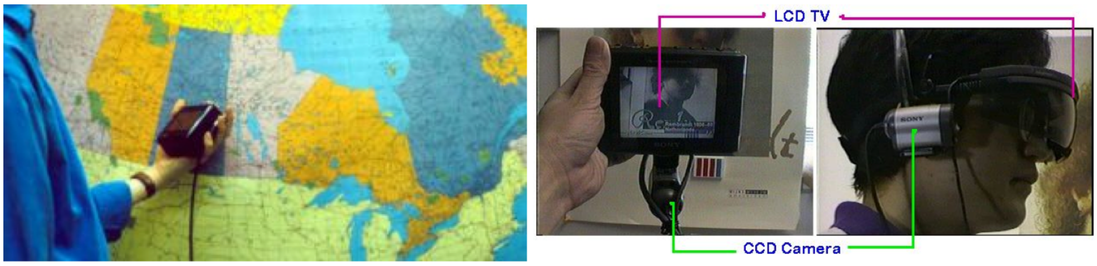

Figure 2.6: Pioneer AR interfaces: The so called dynamic peephole interfaces of Fitzmaurice [63] (left.) The NaviCam application, a magnifying glass approach to AR by Rekimoto et al. [187] (right).

have shown that even highly robust natural feature tracking from IMU/vision sensor fusion is possible on a UMPC, if a detailed model of the environment is available [184].

Mobile device displays Before the recent introduction of UMPCs or cell phones with CPUs of significant computing power, PDAs were the only true mobile alternative for AR researchers. PDAs now have enhanced color displays, wireless connectivity, web-browser and GPS system. A number of computational issues made the use of older PDAs for AR difficult, primarily the lack of dedicated 3D capability and the lack of a floating point computational unit. However, some researchers were able to use the PDA for AR with some success. Wagner et al. demonstrated the Invisible Train [232], which used PDAs as handheld displays for AR applications (see figure 2.7). Makri et al. [133] allowed for a custom-made connection with a special micro-optical display as an HMD.

Smart phones are fully featured high-end cell phones featuring PDA capabilities, so that applications for data processing and connectivity can be installed on them. As the processing capability of smart phones is improving, this enables a new class of AR applications that use the smartphone's camera for vision based tracking. Notable examples are from Wagner et al. [233], Henrysson et al. [91] and Olwal [160] utilizing them as final mobile AR displays. The Wikitude[1] project implements a mobile AR travel guide with AR functionality based on Wikipedia[2] or Panoramio[3] running on

[1]http://www.mobilizy.com/wikitude.php
[2]http://www.wikipedia.org/
[3]http://www.panoramio.com/

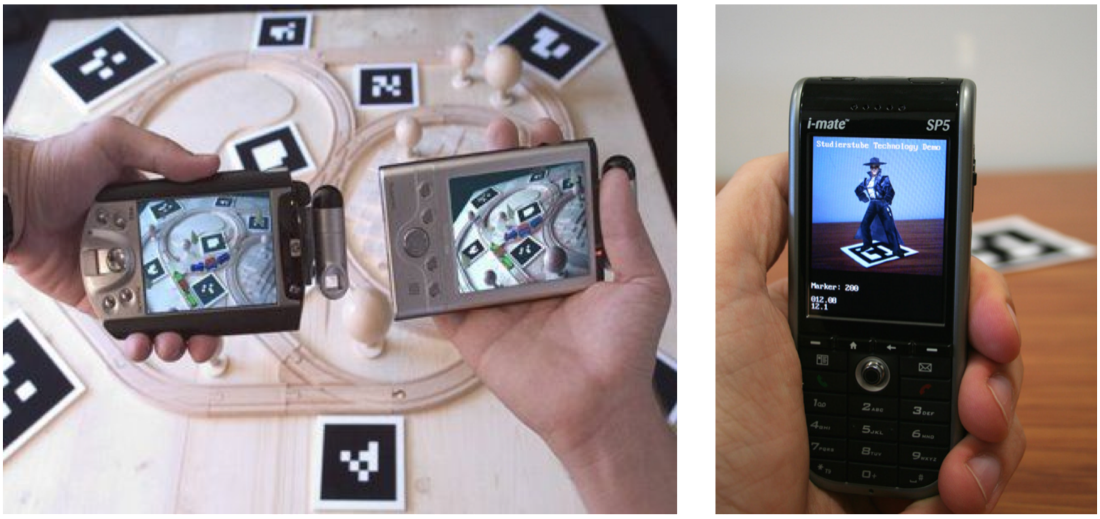

Figure 2.7: First handheld AR Displays on mobile devices: The Invisible Train demonstration by Wagner et al. [232], (left) and a first marker-based tracking system for a mobile devices by again by Wagner [233] (right).

Android smartphones. Looking at the mobile device screen, the user sees an annotated landscape in an AR camera view. For instance, looking at a mountain range, the user will see toponyms and descriptions of the mountains. The user can then download additional information about a chosen location from the Web, say, the names of businesses in the local shopping center. The tracking of the mobile device is done by the built-in GPS sensor and orientation sensor. Also the Nokia research team has demonstrated a prototype phone equipped with *MARA* [104] (Mobile Augmented Reality Applications) software and the appropriate hardware: a GPS, an accelerometer, and a compass. The phone is able to identify restaurants, hotels, and landmarks and provides Web links and basic information about these objects on the phone's screen [84]. The latest research on smart phones focuses on vision based tracking of natural features allowing for the tracking of the user in unprepared and unconstrained environments. Rohs et al. used smart phones for markerless tracking of magic lenses on paper maps in real-time [193]. Furthermore, Wagner et al. already made major advances in pose tracking from natural features on mobile phones [233].

2.3.2 Visualization & Interaction Techniques

Visualization and interaction techniques play a central role in mixed reality systems. Especially in systems that use head mounted displays, careful visualization design is essential as the virtual information integrated into the user's view may obscure important parts of the real-world environment or can distract the user significantly. Similarly, careful design of the interaction techniques in a MR system is required to ensure that the potential of MR systems to provide an intuitive and usable interface is realized. While new approaches to MR visualization and interaction are still emerging, there is also a growing consensus regarding the applicability and usability of established techniques. The following sections aim to provide a brief overview of the available design space and introduce some common MR visualization and interaction techniques.

2.3.2.1 Visualization Techniques

Visualization techniques for use in MR applications can be characterized by

- their spatial reference,
- their integration with the real-world environment and
- the amount of visual realism.

The spatial reference frame describes how the 3D graphics objects that are rendered into the augmented graphics display are spatially bound to the real-world environment. Typical reference frames are the world, objects, the body of the user and the screen of the display device. Using the world as a reference frame, virtual objects are bound to a geo-spatial location and their visualization behaves like a physical object located at this position. Using this method, many well known AR applications can be implemented, e.g. the visualization of planned or historic buildings integrated into the current real-world environment or the visualization of hidden infrastructures. The use of objects as a reference frame defines a local coordinate system, where the visualizations move with the object. This reference frame is commonly used in marker-based AR systems (where the absolute geo-spatial position is not known), to implement tangible user interfaces (discussed in the following section) or to display instructions in maintenance or assembly applications (where only relative locations with respect to the object under consideration are relevant). The use of the user's body as the spatial reference is used to make virtual tools easily accessible in virtual reality systems (in a fixed location with respect to the user, as in a virtual tool belt), but less common in mixed reality setups. Finally, the augmentation information can be bound to the screen of the display device, resulting in overlays that always appear in the same display location. A common use of this is the implementation of head-up-displays.

Regarding the integration with the real-world environment, MR visualization techniques can replace, enhance or mediate the real world environment. In the simplest case, the graphics rendered from an MR visualization simply replace the real environment in a part of the display. This approach prevents a tight integration of virtual content with the real-environment, but has the advantage that all existing visualization techniques can be embedded into an MR application in this way. More typical are visualization techniques where the added information is used to enhance the real-world view, which remains visible. By adjusting transparency, the display can be seamlessly blended between virtual and real objects. Visualization techniques that mediate the real world environment can also filter information or objects from the environment; carried to the extreme, complete real-world objects could be removed from the user's view in a setup known as "diminished reality".

With respect to the visual realism of visualization techniques, the design space spans a continuum from abstract to photo-realistic graphics. The central set of available design parameters are the depth-cues (occlusion, shading, shadows, parallax) [95] but artificial meta-objects (illustration techniques) and visual abstraction techniques (e.g. NPR) are also possible [68]. Drasic and Milgram [52] examined the impact of stereoscopic vision in AR displays and Surdick et al. [222] discuss the impact of various depth cues. The central challenge in the development of visualization techniques for mixed reality applications is to design techniques that are perceptually easy to interpret for the user as well as efficient to model and render. Additional constraints can arise from the display device used, e.g. in optical see-through devices where the real background always remains visible [195]. Depending on the application the level of realism has to be adjusted to either clearly convey the difference between virtual and real objects or to mix them as seamlessly as possible. The management of the amount of information to be displayed (filtering) [102] and the spatial layout of visualization objects are additional relevant issues [13].

2.3.2.2 Interaction Techniques

Established GUIs are based on the WIMP (windows, icons, menu, pointer) concept in which a user interacts with a graphical interface representation using the mouse. This approach has become the standard means of interaction in desktop applications and has the key advantage that a limited set of standardized interaction hardware (mouse or similar 2D pointing devices) in combination with standardized graphical interface objects (widgets or controls) enables users to control arbitrary applications. The direct transfer of such techniques to mixed reality applications is possible [223], but mixed reality applications are not limited to such techniques. A potential benefit of mixed reality interaction techniques is to enable more direct manipulation in which the user manipulates real world objects, exploiting his everyday physical manipulation skills. While conventional GUIs are limited to indirect manipulation of virtual objects using a 2D pointer, mixed reality applications offer a larger design space for the development of task specific interaction techniques in which user interactions with objects in the real world control the application. A simple example for a common mixed reality interaction technique in which a physical object is used both in direct interaction and to control the application are so called magic books, introduced by [19]. The user turns pages in a physical book (direct interaction). This manipulation is tracked by the system (e.g. using image recognition on special markers on the pages) and additional actions like the display of augmentation information are triggered by the application.

The use of a physical object to control an application is commonly referred to as a tangible user interfaces (TUI). Ullmer and Ishii [100] define TUIs as systems relating to the use of physical artifacts as representations and controls for digital information. A similar approach under the name of "Graspable User Interfaces" was introduced by Fitzmaurice et al. in the Bricks project [65]. TUIs remove indirections in the interaction and can exploit real world skills of users like bimanual manipulation. However, they require careful design and must be tailored to each application to exploit this potential. The need for application specific development and sensor hardware can be problematic in some application contexts. In the Bricks project small wooden pegs were assigned as physical controls to virtual objects. Using a number of Bricks, users could manipulate objects. The physical embodiments are not application specific and can be reassigned in another application context [65].

A tangible user interface for the manipulation and query of spatial data was introduced in the “Tangible Geospace” application as part of the metaDESK project by Ullmer et al. [227]. Several TUI prototypes have developed in the domains of urban and architectural planning. In the Illuminating Clay system [174], users were able to physically model a landscape on a table surface. The geometry was then acquired using a laser-scanner and used within the modeling application. Such an approach enables very intuitive modeling interaction, if the desired information can be specified by the clay model. However, this interaction approach is obviously highly specialized. A more general, but less direct approach was introduced in the “Magic Cup” system by Kato et al. in which physical marker objects that can be spatially tracked were assigned to virtual buildings, benches, lamp posts, etc. to enable simple spatial manipulation in a city planning application [109]. Another important category of interaction techniques for mixed reality applications is based in the recognition of user gestures. In addition to approaches that are similar to general gesture recognition in that finger, hand or body gestures of the user are tracked and recognized (e.g. [30]), the use of camera gestures in an inside-out-vision setup is of special interest due to the proliferation of mixed reality on camera equipped mobile devices like smartphones and PDAs. In this setup, the user moves the camera to signal the gesture that is interpreted from the video stream. Reimann and Paelke [183] have given an overview of how common interaction tasks like selection and quantify can be implemented in such an approach.

2.3.3 Applications & Examples

Below, we overview a broad range of important AR applications in the geospatial domain. We discuss demonstration and research prototypes as well as real world applications. This section does not list all noteworthy applications, but rather aims to provide a representative overview of current work in this context.

Project Name	Institutions	Project Description, Reference
Augmented Maps Cambridge	University of Cambridge, UK	At the University of Cambridge a system to augment printed maps with digital graphical information and user interface components was developed. These augmentations complement the properties of the printed information in that they are dynamic, permit layer selection and provide complex computer mediated interactions with geographically embedded information and user interface controls. Two methods are presented which exploit the benefits of using tangible artifacts for such interactions. The overall system centers around a table top environment where users work with maps. One or more maps are spread out on a table or any other planar surface. A camera mounted above the table tracks the maps' locations on the surface and registers interaction devices placed on them. A projector augments the maps with projected information from overhead [185].
AR PRISM	University of Washington, USA and Hiroshima City University, Japan	This system presents the user geographic information on top of real maps, viewed with a head-tracked HMD. It allows collaborative work of multiple users (via multiple HMDs) and gesture-based interaction [90].
ARMobile	VTT Technical Research Centre of Finland, Finnland	ARMobile technology is mobile software (Symbian, Java) that adds user-defined 3D objects to the camera view of the mobile phone. The application enables placing, for example, virtual furniture on mobile phone's camera image [246].
Enkin	University of Koblenz–Landau, Germany	Enkin displays location-based content in a unique way that bridges the gap between reality and classic map-like representations. It combines GPS, orientation sensors, 3D graphics, live video, several web services and a novel user interface into an intuitive and light navigation system for mobile devices[1].

[1] http://www.enkin.net/

GeoScope	Leibniz University Hannover, Germany	This application is a telescope like novel mixed reality I/O device tailored to the requirements of interaction with geo-spatial data in the immediate environment of the user. The I/O device is suitable for expert and casual users, integrates with existing applications using spatial data and can be used for a variety of applications that require geo-visualization including urban planning, public participation, large scale simulation, tourism, training and entertainment [163].
Handheld Augmented Reality	Graz University of Technology, Austria	This projects aims at providing AR anywhere and anytime. It mainly focuses on developing a cost-effective and lightweight hardware platform for AR. Based on this platform, they developed some applications [234].
IPCity - Interaction and Presence in Urban Environments	EU funded Sixth Framework program	The vision of the IPCity project is to provide citizens, visitors, as well as professionals involved in city development or the organization of events with a set of technologies that enable them to collaboratively envision, debate emerging developments, experience past and future views or happenings of their local urban environment, discovering new aspects of their city[1].
Localization and Interaction for Augmented Maps	University of Cambridge, UK	It augments printed maps with digital graphical information and user interface components. These augmentations complement the properties of the printed information in that they are dynamic, permit layer selection and provide complex computer mediated interactions with geographically embedded information and user interface controls [185].
MARA - Sensor Based Augmented Reality System for Mobile Imaging	Nokia Research Center, Finnland	MARA utilizes camera equipped mobile devices as platforms for sensor-based, video see-through mobile augmented reality. It overlays the continuous viewfinder image stream captured by the camera with graphics and text in real time, annotating the user's surroundings [104].

[1] http://www.ipcity.eu/

MARQ - Mobile Augmented Reality Quest	Graz University of Technology, Austria	This project aims at developing an electronic tour guide for museums based on a self-contained, inexpensive PDA, that delivers a fully interactive 3D AR to a group of visitors[1].
MOBVIS - Vision Technologies and Intelligent Maps for Mobile Attentive Interfaces in Urban Scenarios	EU funded Sixth Framework program	The MOBVIS project identifies the key issue for the realization of smart mobile vision services to be the application of context to solve otherwise intractable vision tasks. In order to achieve this challenging goal, MOBVIS claims that three components, (1) multimodal context awareness, (2) vision based object recognition, and (3) intelligent map technology, should be combined for the first time into a completely innovative system - the attentive interface [2].
Overlaying Paper Maps with Digital Information Services for Tourists	ETH Zurich, Swiss	It implements interactive paper maps based on emerging technologies for digitally augmented paper. A map of the Zurich city centre was printed using the Anoto pattern and a PDA used to visualize the supplementary digital information. It also developed an Interactive Map System for Edinburgh Festivals [158].
Signpost	Vienna University of Technology, Austria	It is a prototypical tourist guide application for city of Vienna covering both outdoor city areas as well as indoor areas of buildings. It provides a navigation model and an information browser mode. His low-cost indoor navigation system runs on off-the-shelf camera phones. More than 2,000 users at four different large-scale events have already used it. The system uses built-in cameras to determine user location in real time by detecting unobtrusive fiduciary markers. The required infrastructure is limited to paper markers and static digital maps, and common devices are used, facilitating quick deployment in new environments [151].

[1] `http://studierstube.icg.tu-graz.ac.at/handheld_ar/marq.php`
[2] `http://www.mobvis.org`

Situated Documentaries	Columbia University, USA	It is an experimental wearable AR system that enables users to experience hypermedia presentations that are integrated with the actual outdoor locations to which they are are relevant. The system uses a tracked see-through head-worn display to overlay 3D graphics, imagery, and sound on top of the real world, and presents additional, coordinated material on a hand-held pen computer [94].
The Touring Machine	Columbia University, USA	It presents information about Columbia university's campus, using a head-tracked, see-through, head-worn, 3D display, and an untracked, opaque, handheld, 2D display with stylus and trackpad [62].
Timmi	University of Münster, Germany	The main idea behind the Timmi application is that the camera image of the physical map is augmented with dynamic content, for example locations of ATM machines on the map. By moving a tracked camera device over the physical map users can explore requested digital content available for the whole space of the map by just using their mobile PDA or smartphone as a see-through device. For this purpose the mobile camera device has to be tracked over the physical map using AR Toolkit Markers [210].
Tinmith	University of South Australia, Australia	It supports indoor and outdoor tracking of the user via GPS and fiducial marker. Interaction with the system is brought by the use of custom tracked gloves. The display of overlap is delivered by a video see-through HMD. The main application area of Tinmith is outdoor geometric reconstruction [172].
Urban Sketcher	Graz University of Technology, Austria	It describes how mixed reality (MR) technology is applied in the urban reconstruction process and can be used to share the sense of place and presence. It introduces Urban Sketcher, an MR prototype application designed to support the urban renewal process near or on the urban reconstruction site [198].

Vidente	Graz University of Technology, Austria	Vidente is a handheld outdoor system designed to support field staff of utility and infrastructure companies in their everyday work, such as maintenance, inspection and planning. Hence, hidden underground assets including their semantic information, projected objects and abstract information such as legal boundaries can easily be visualized and modified on-site [199; 200].
WalkMap	Nokia Research Center, Finnland	WalkMap is targeted at a walking user in an urban environment, and offers the user both navigational aids as well as contextual information. WalkMap uses AR techniques to display a map on the surrounding area on the user's head-worn display [122].
Wikitude - AR Travel Guide	Mobilizy, Austria	Wikitude is a mobile travel guide for the Android platform based on location-based Wikipedia and Qype content. It is a handy application for planning a trip or to find out information about landmarks in surroundings; 350,000 world-wide points of interest may be searched by GPS or by address and displayed in a list view, map view or camera view[1].

Table 2.1: Related Augmented Reality Projects in Geospatial applications (in alphabetic order).

[1] `http://www.mobilizy.com/wikitude.php`

In the following, we use the above five major AR types to categorize the AR Geo-applications. We categorize using the different interface types (optical see-through HMD, video see-through HMD, handheld display AR, projection-based AR with video augmentation, projection-based AR with physical surface augmentation); see also section 2.3.1) and the tracking system (GPS, marker based, marker less (optical tracking), IR, RFID, Inertial, UWB) that is used.

AR display type	**Geo-applications: "Application name (Developer)"**
Optical see-through HMD	• AR WalkMap (Nokia Research Center) • The Touring Machine (Cambridge University) • AR PRISM (University of Washington and Hiroshima City University) • Situated Documentaries (Columbia University) • AR/GPS/INS for Subsurface Data Visualization (University of Nottingham) • Signpost (Vienna University of Technology)
Video see-through HMD	• HMD AR Tinmith (University of South Australia)

Hand-held display AR	• Wikitude (Mobilizy) • MARA - Sensor Based Augmented Reality System for Mobile Imaging (Nokia Research Center) • Enkin (University at Koblenz-Landau) • MOBVIS (EU funded Sixth Framework programme) • MARQ - Mobile Augmented Reality Quest (TU Graz) • Handheld Augmented Reality (TU Graz) • ARMobile (Technical Research Centre of Finland) • Vidente (TU Graz) • Signpost (TU Graz)
Projection-based AR with video augmentation	• Urban Sketcher (TU Graz)
Projection-based AR with physical surface augmentation	• Augmented Map System (Cambridge University) • Digitally augmented paper maps (ETH Zurich), IPCity - Interaction and Presence in Urban Environment (EU funded Sixth Framework programme)

Table 2.2: Geo AR examples categorized by interface type.

In conclusion, AR promises to combine the interactive nature of computer generated content with real world objects. This allows the creation of new forms of interactive maps. By using AR as an interface, maps gain a dimension of realness and interactivity in the spectrum from reality to virtuality. The past five years have seen an increasing use of maps in mobile and ubiquitous computing applications, in which they are visually presented as realistically or as virtually as suits the users needs. More recent forms of maps are already building on online access to geographic information and leverage geospatial Web services. By using AR such geographic information can be represented intuitively. The increasing demand from the general public for prompt and effective geospatial services is being satisfied by the revolution of web mapping from major IT vendors. Since maps are two-dimensional representations of the three-dimensional real world, maps will continue to evolve towards more integrated and higher-dimensional representations of the real world. Moreover, the latest technological developments for handheld devices allow AR to become mobile and ubiquitous. Both developments combine well and will also lead to social implications of where and how information is consumed by users.

2.4 Peripheral Research

In the following section, work related to the specific content of later chapters is summarized. It is divided into related work for our (1) pad-sized interfaces, (2) tab-sized interfaces, and (3) board-sized interfaces. Research relevant to the application areas of specific interface prototypes is also highlighted.

2.4.1 Related Work in the Area of Pad Sized Interfaces

In this section, work pertinent to our research into pad-sized interfaces and to the our application areas for these interfaces is summarized. In particular, we discuss wayfinding in general, location-aware mobile guides, mobile map interaction, UI for games on pad-sized devices, work on mobile projection systems and work in the crisis response domain.

2.4.1.1 Wayfinding

Lynch defines the term *wayfinding* in his book The "Image of the city" [130] as a consistent use and organization of definite sensory cues from the external environment?. In his book, he also considers how users perceive and organise spatial information while performing wayfinding tasks in the city by forming mental maps typically comprising five elements including landmarks. In [226], the concept of spatial mental models is introduced. Tversky highlights the ability of individuals to carry a mental model that

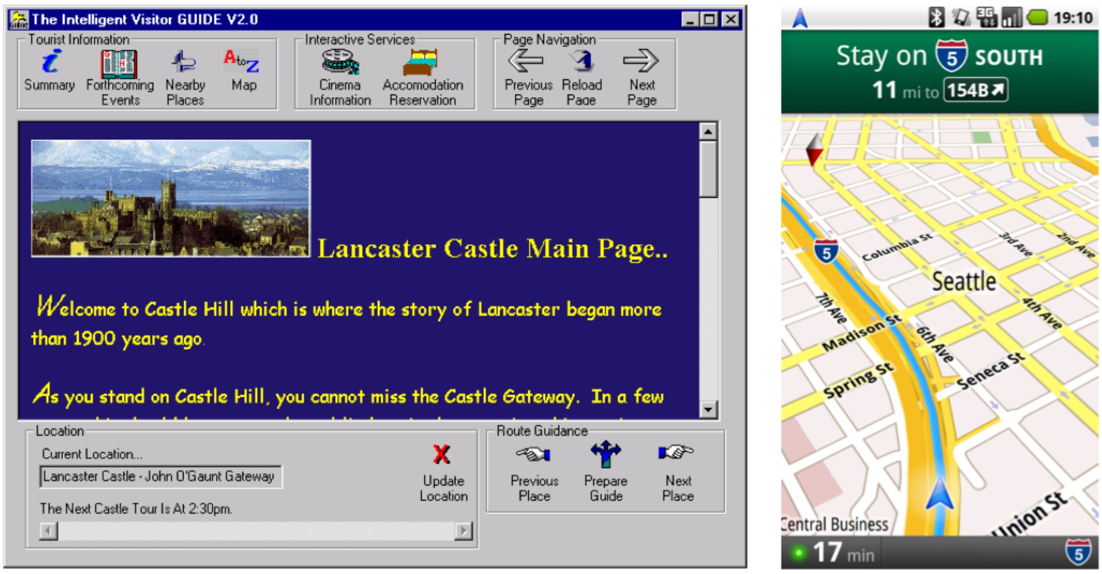

Figure 2.8: The GUIDE LBS developed at Lancaster University [1] (left). Google Maps Mobile running on an Android mobile device (right).

supports spatial relations amongst elements, e.g. landmarks, but does not allow accurate metric judgments to be made. Traditional 'You are here' maps i.e. maps that explicitly indicate the position of the map reader [113; 123; 161] enable users to orient themselves spatially to their current environment and acquire survey knowledge. Being a form of map, they support wayfinding and provide their readers with an appreciation of nearby landmarks.

2.4.1.2 Location-Aware Mobile Guides

Various kinds of maps and other geospatial content can be found on the Internet. With the growth of the mobile Internet, geospatial data can be easily retrieved from mobile devices and, utilizing LBS, filtered based on position. Therefore, it is now possible to use geospatial data on the fly without a desktop computer. Gartner et al. [176], Küpper [120] and Urquhart et al. [229] provide an overview of the major LBS prototypes that allow users to display maps and interact with them on a mobile device through ordinary interaction techniques, i.e., joystick or key input. *Google Maps Mobile*[1] is a prominent commercial example of a mobile geospatial service (see figure 2.8).

[1] `http://www.google.de/gmm`

A thorough overview of location-aware mobile guides that rely on maps or map–like representations in providing their services is presented in [11]. Research into the general area of mobile device support for pedestrian navigation/wayfinding has undergone an evolution over the last 15 years from the early research prototypes to the now common "mobile maps" standard on many new GPS-enabled mobile phones. Fundamental research into the technical and human factors issues arising from the use of mobile devices and into the concept of location-awareness to support pedestrian navigation was conducted under the CyberGuide [1] and GUIDE [38; 46] projects. Both utilized some form of graphical map representation as part of their functionality (see figure 2.8 (left)). The Cyberguide project [1] was responsible for two mobile guides, one to support navigation inside and the other to support outdoor navigation. Both enabled the user to view a schematic map of the area, automatically updated according to the position of the user. The position was determined by means of infrared sensors (indoors) or GPS (outdoors). The GUIDE project [38; 46] led to the development of different prototypes of mobile tourist guides for the city of Lancaster. GUIDE allowed the user to request a sketch-like map of her/his surroundings that had a simple animation highlighting the path to their next chosen attraction. One of the later studies relating to GUIDE involved a Wizard of Oz format in which the user could select an attraction, e.g. Lancaster Castle, by taking a photo of the attraction with a camera equipped PDA device (see [45] for more details). Other system are described and developed by Wasinger et al. [237], Krüger et al. [117] and Baus et al. [12]. With the M3I again Wasinger et al. [238] describe a near-complete Pocket PC implementation of a Mobile Multi-Modal Interaction platform for pedestrian navigation. The platform is designed to easily support indoor and outdoor navigation tasks, and uses the combination of several modalities for presentation output and user input. Whereas 2D/3D-graphics and synthesized speech are used to present useful information on routes and places, fused input from embedded speech and gesture recognition engines allow for situated user interaction. Krüger et al. [118] describe the concepts behind the BPN (BMW Personal Navigator), an entirely implemented system that combines a desktop event and route planner, a car navigation system, and a multi-modal, in- and outdoor pedestrian navigation system for a PDA.

A detailed investigation into the implications of supporting Wayfinding through the use of mobile maps is presented by Willis et al. [125]. In this paper, Willis and colleagues

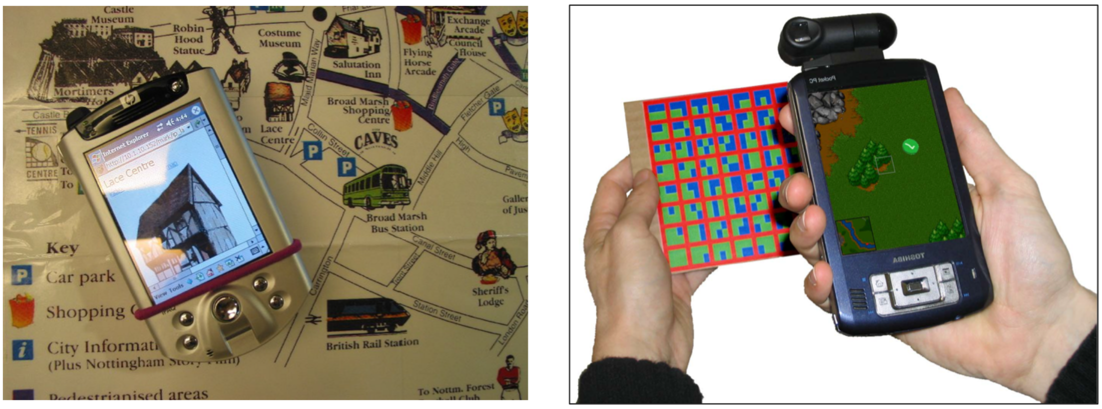

Figure 2.9: The Marked Up Maps Project by Reilly et al. [182]; they use maps equipped with an array of RFID (Radio Frequency Identification) tags to realize the physical hyperlink (left). Hachet et al. [77] realize a two-handed magic lens interface by tracking a piece of cardboard that the user moves behind a camera-equipped device (right).

describe the results of an experiment that examined the types of knowledge acquired by an individual depending on whether they used a physical paper map or a mobile map (a Nokia 6630 mobile phone running the Route 66 mobile mapping software application). Her study revealed that mobile map users tended to perform worse than paper map users, particularly in wayfinding and route distance estimation. In other words, the study showed that the format and presentation of spatial information has an effect on task performance. Kray et al. [115] present results from a series of experiments, where relevant factors for the use of path prepositions were examined. We were especially interested in the concepts behind the German prepositions “entlang” and “vorbei” (similar to “along” and “past”).

2.4.1.3 Mobile Map Interaction

The merging of the richness and dynamism of digital geospatial content with physical maps is known as mobile map interaction (MMI). In MMI, mobile devices serve as sources for additional information when interacting with a given map. Some approaches add hyperlink information to locations of a physical map. The mobile device can then display specific points of interest while the physical map provides a static overview. For example, Reilly et al. [182] use maps equipped with an array of RFID (Radio Frequency Identification) tags to realize the physical hyperlink (see figure 2.9 (left)).

Disadvantages of this approach are the low spatial resolution (because of the size of the RFID tags) and the high map production costs. The method was enhanced by computer vision techniques [181] to achieve higher spatial resolution.

Applying the principle of magic lenses to paper maps is natural. When positioned over the map at a certain distance, map features recorded by the camera can be perceived on the display in real time. Following the classical video see-through augmented reality approach, digital geospatial data can then be overlaid over the images, delivering information to the user originally not available on the paper map. Geospatial data can be retrieved via a Web Map or Web Feature Service (WFS). Both are standardized protocols that allow geospatial data to be requested across the web. This allows users to personalize each paper map with content of their interest without having the map cluttered with too much information.

The magic lens metaphor is related to various forms of navigating virtual information on a plane. Mehra et al. [141] compare the *dynamic* and *static peephole* metaphors. They extend the work of Fitzmaurice [64] and Yee [250], and provide empirical evidence that a dynamic peephole interface is superior for tasks in which spatial relationships are important and the display size is limited. Mehra et al. use a mouse-based PC interface and focus on line-length discrimination whereas we focus on map exploration, target localization, and the remembering of a specific target attribute, hence requiring spatial memorization. The physical context of the interaction is different, since motor control issues of thumb movement vs. arm movement in 3D space play a role. Baudisch et al. [9] investigate the use of a high resolution focus display in combination with a lower resolution context display (see figure 2.10). Hachet et al. [77] realize a two-handed magic lens interface by tracking a piece of cardboard that the user moves behind a camera-equipped device (see figure 2.9 (right)). Sanneblad and Holmquist [197] use ultrasonic tracking to align a small display with a large overview in the context of a map application.

In our previous work [211], we have used a marker-based approach with an UMPC and a Symbian smartphone. Marker-based tracking has the severe drawback of requiring markers of considerable size scattered all over the map and thus disturbing its appearance, aesthetics, and overall usability, but the advantage that geospatial data can be precisely augmented on the map. Wikeye [88] is a follow-up to Timmi that aimed at improving the understanding of places through the combination of digital

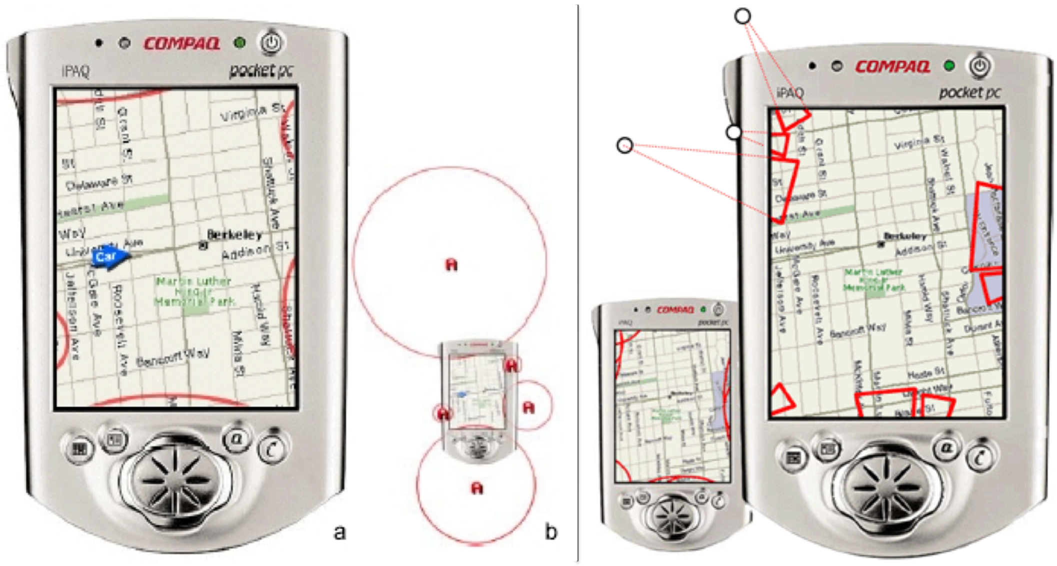

Figure 2.10: Off-screen visualization techniques of Baudisch et al.: Halo technique [10] (left) and Wedges technique [76] (right).

Wikipedia content with a paper-based map. The goal of WikEye is to help users to understand more about their surroundings via an easy to use mobile interface. When the user views a small portion of a map through her mobile device, Wikipedia-derived content relating to spatial objects in view is offered to the user. For example, when a WikEye user examines a small portion of a Berlin paper map through the camera phone like the area containing the Reichstag building and the Brandenburg Gate Wikipedia content is overlaid on the camera image of the map, highlighting these spatial objects and their relationships. Following a clock metaphor, rotating the device about the camera axis switches to a different time in history and delivers an overview of content related to that time period. Depending on the spatial extent of the map visible on the camera display unit, the system responds by offering Wikipedia data about spatial objects with larger area footprints. In other words, as the cartographic scale decreases, the threshold area of a spatial object to be featured with Wikipedia data increases. While WikEye [88] concentrates on the interaction within a single map, WikEar is a bridge between two or more maps, a concept that will be elaborated upon in the next section. WikEar also differs in its approach to content. While WikEar utilizes whole narratives from the Minotour [86] project, WikEye is limited to much simpler forms of

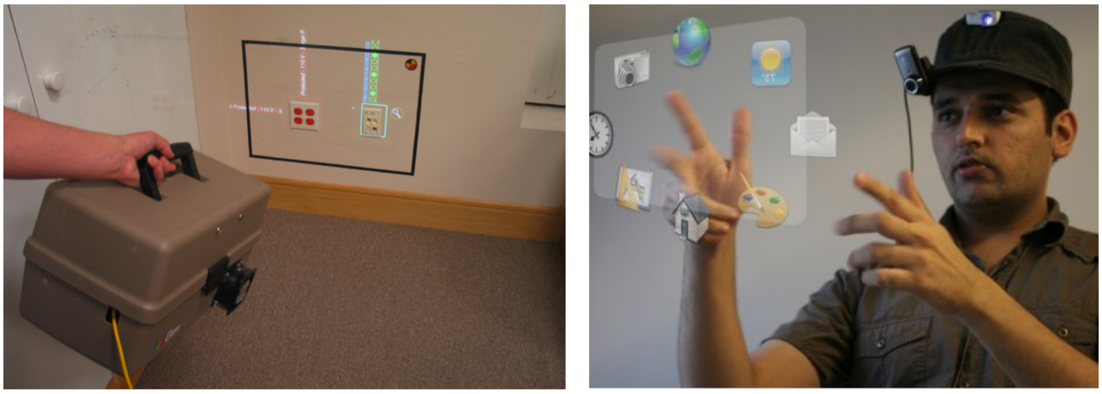

Figure 2.11: RFIGLamps were able to create object adaptive projections [177] (left). The concepts of the Wear Ur World (WUW) application show that mobile projection could be utilized in every day life [145] (right).

educating the user with Wikipedia data. Later in this thesis WikEar is explained in more detail.

2.4.1.4 Mobile AR Games

A seminal augmented reality game is the *Invisible Train* by Wagner et al. [232; 234], in which players control virtual trains on a real model railroad track with a touch screen of a PDA. The games *Penalty Kick*, *Impera Visco* and *SmartMemory* use two-dimensional barcodes to realize the interaction with the real world objects. *SmartMemory* is a version of the classical "memory" game and Impera Visco is a multi-user extension. *Penalty Kick* is a simple penalty shootout game that consists of a printed soccer field, a goal, the penalty spot, and spectators in the background. The display of the camera phone shows the virtual goalkeeper and the user can control the virtual ball with the mobile device to score a goal. All three games are described in detail in [191]. A good overview of augmented reality game and techniques can be found in [164].

2.4.1.5 Mobile Projection

Initial research on mobile projection interfaces was conducted by Raskar et al. with the iLamps [178]. While the iLamps mainly focused on creating distortion free projection on various surfaces and using multiple projectors to create a larger projection, the follow-up of the iLamps, the RFIGLamps [177], were used to create object adaptive projections. Set in a warehouse scenario, the RFIGLamps could be used for example to mark products for which the expiration date is close to the actual date (see Figure 2.11). Blaskó et al. [22] explored the interaction with a wrist-worn projection display by simulating the mobile projector with a steerable projector in a lab. To examine the possibilities of multi-user interaction with a mobile projector, Cao et al. used an instrumented room to create several information spaces, which could be explored with handheld projectors [35]. Hang et al. [73; 81] have outlined the advantages of projected displays over mobile phone displays for exploring large-scale information. With the Wear Ur World (WUW) prototype of the SixthSense project of the MIT, Mistry et al. [145] showed that mobile projection could be utilized in every day life (see 2.10). According to Mistry et al. WUW was built from parts that are available for around $ 350 (in addition, laptop in a backpack is needed to run the services provided in WUW), showing that mobile projectors could reach mass markets in the near future. As a result, a rich design space for projection-based mobile augmented reality applications could emerge. With a built-in projector, not only the graphical scale of the applications can be increased but also the range of application possibilities can widen. In combination with the built-in camera of a mobile device, mobile camera projector units become a powerful interface for mobile AR applications. For instance, they can replace head-mounted displays for AR games, which suffer from poor mobility and comfort. Similarly, projectors could be more scalable to multiple users than current magic lens applications.

2.4.1.6 Crisis Response

The design of effective collaborative visualization and decision-making tools for crisis response are well known problems in digital cartography, geospatial visualization and GIS. In addition, other disciplines, such as the HCI and ubiquitous computing, have tried to tackle various portions of these problems. Most of the existing work focuses on large format map applications that support decision-making, for example, in an emergency operation center (EOC) and are discussed later in section 2.4.3. All this work concentrates on supporting decision-making and group collaboration at a single command center, but does not concentrate on the problem of the communication between mobile units and the command center. In the HCI community, Palen and

Liu [165] describe new perspectives on citizen-based activities that arise out of peer-to-peer communications in disaster activities. These activities serve important tactical, community-building, and emotional functions. Yuan et al. [251] propose a concept for an intelligent mobile crisis response system (CRS) that would have facilitated a more effective response in a given scenario. They list six critical tasks of a CRS (Monitoring and Reporting, Identification, Notification, Organization, Operation, Assessment and Investigation) and outline several ideas to improve performance on these tasks.

2.4.2 Related Work in the Area of Tab Sized Interfaces

This section mainly focuses on providing background information on *geocaching*. This information is important background for chapter four. Other peripheral work in the area of intelligent narrative technologies important for chapter four and five is described at the end of this subsection.

2.4.2.1 Geocaching and its Origin

Previous work has addressed the uniqueness of the geocaching activity in several ways. Chaves [36; 37] was one of the first researchers to look at motivations and interest of geocachers. O'Hara [159] investigated similar questions from a qualitative perspective with a rather small user set. The basic contribution of O'Hara's book is to underline the importance of the geocaching community by presenting certain unique attributes of the geocaching activity. It is also worthwhile mentioning the work of Kelley [110], who presents interesting anecdotes that raise a couple of questions related to the findings in this paper. Matherson et al. [136] present studies on the first use of geocaching in the classroom and share perspectives from a teacher and her students on its value.

Origin of Geocaching On May 1st, 2000, the amount of error of the Selective Availability (SA) GPS signal was "set to zero" [168]. This allowed much more precise localization for civil and scientific purpose using GPS. The next day was the birthday of geocaching. On May 2nd, Dave Ulmer posted the main idea of geocaching in an internet newsgroup[1]: "it should be easy to find someone's stash from waypoint information. Waypoints of secret stashes could be shared on the Internet, people could navigate to the stashes and get some stuff. The only rule for stashes is: Get some Stuff, Leave

[1]The group can be found under: `sci.geo.satellite-nav`.

Figure 2.12: The first *geocache* was placed by Dave Ulmer on May 3, 2000. The event has been honored with a metal plate describing the first cache?s contents (a). Geocache sizes: A micro cache (b), a small cache (c), a regular cache (d) and a large cache (e).

some Stuff !!!" Ulmer's post is still available online in the newsgroup (Internet forums and websites remain the key areas in which the geocaching community stores and communicates the locations of caches.) On the next day, Ulmer buried a first "stash" (containing a log book, a map of the area, a tape recorder and a can of baked beans), (see figure 2.12) in the south of Oregon (USA) and posted the coordinates in the newsgroup. Soon, the first website appeared for storing and communicating the locations of caches and the community rapidly grew. Interestingly, the activity was originally referred to as "GPS stash hunt" or "gpsstashing", however, the name was changed to *geocaching* after a discussion in the gpsstash discussion group, because "stash" can have negative connotations. Today, nearly one million caches are "stashed" worldwide. Many different kinds of caches were hidden by all sorts of users. In the following, we describe the some of these different cache types and sizes.

Types and Sizes of Geocaches Currently there are about two dozen different cache types[1]. In the following, we describe the most important and most popular ones [244]

Traditional Caches: Very similar to the first "stash" by Dave Ulmer, this is the original cache type comprising, at a minimum, a container and a log book.

Multi-caches: A multi-cache involves two or more locations, the final location being a physical container. The most common multi-caches provide a hint at the first location to guide the geocacher to the next one.

Mystery or Puzzle Caches: This "catch-all" of cache types usually involves quite complex caches. Theses caches can involve brainteasers or puzzles that need to be solved correctly to determine the coordinates of the next or final cache.

[1] `http://www.groundspeak.com/`

Letterbox (or Hybrid) Caches: Similar to the "catch-all" type, the letterbox (or hybrid) cache includes clues that are used instead of coordinates to describe the next position.

Other Caches types: There are many other different types of caches (or similar activities like geohashing) listed on different geocaching websites, some of which are no longer available for creation and some of which are in their nascent stages. These cache types include [124], (Mega-) Event Cache, Cache-In-Trash-Out, Virtual Cache, Webcam Cache, Locationless-Cache and Wherigo.

More information: More information about the various different cache types can be found in [170].

Four main sizes of geocaches can be found:

1. Micro caches are tiny and often just have a logbook and a pen. Often film canisters are used for caches of this size.

2. Small caches are bigger than the micro ones and can have a volume of up to 1.5 litres. These caches can contain travelbugs or small bartering objects.

3. Regular caches have a size of 1.5 to 20 litres and can contain different types of bartering objects.

4. Large caches are caches of more than 20 litres. Typical examples include: rain barrels, boxes, suitcases or even closets.

Again [53; 170; 218] provide additional geocaching details (see figure 2.12).

2.4.2.2 Intelligent Narrative Technologies

Parts of the work described in chapter four as well as in chapter five is rooted in the field of intelligent narrative technologies (INT). Although there are many technological approaches to INT, the field is firmly based on the narrative theory developed in [24]. Our work in the area of INT is also motivated by semantic relatedness measures, such as those described in [31].

2.4.3 Related Work in the Area of Board Sized Interfaces

In this section, work pertinent to our research into board-sized interfaces and to the our application areas for these interfaces is summarized.

2.4.3.1 Multi-touch Interaction & Hardware

Multi-touch refers to a set of interaction paradigms, which allow to control applications with several simultaneously performed touches. Multi-touch technology presents a wide range of new opportunities for interaction with graphical user interfaces because it allows expressive gestural control and fluid multi-user collaboration through relatively simple and inexpensive hardware and software configurations. The technology itself has been available in different forms since the late 1970s. Several patents demonstrate how camera/sensor-based touch surfaces can be constructed [108; 135; 150; 241]. More technical details will be presented later in this thesis. Bill Buxton's multi-touch webpage[1] provides a thorough overview of the underlying technologies as well as the history of multi-touch surfaces and interaction. Multi-touch surfaces can be realized by using different technologies, ranging from capacitive sensing to video analysis of infrared or full color video images [49; 79; 134; 169].

However, it was Han's 2005 [79] presentation of a low cost camera-based multi-touch sensing technique based upon Frustrated Total Internal Reflection (FTIR) that truly highlighted the potential role multi-touch could play in the development of the next generation of human computer interfaces. Han's system was both cheap and easy to build, and was used to illustrate a range of creatively applied interaction techniques – his YouTube demonstration captured the imagination of experts and laypeople alike. In 2007, interest in multi-touch grew as Apple released details of the *iPhone*[2], a mobile phone with a multi-touch screen as a user interface. The interface and interaction techniques of the *iPhone* received considerable media attention and brought multi-touch to the forefront of the consumer electronics market. Later in 2007, Microsoft announced their *Surface* multi-touch table[3]. The *Surface* has the appearance of a coffee table with an embedded multi-touch interactive screen. In manner similar to the HoloWall [137], the *Surface* has a diffuser attached to the projection surface and

[1] `http://www.billbuxton.com/multitouchOverview.html`
[2] `http://www.apple.com/iphone`
[3] `http://www.microsoft.com/surface`

is illuminated from below with infrared light. Reflections of hands and objects are captured by cameras inside the table in an approach described as diffused illumination (DI). By utilizing a grid of multiple cameras, the *Surface* has a sensing resolution sufficient to track objects augmented with visual markers. Considerable research has explored the benefits of multi-touch interaction [49; 103; 147; 148; 188; 206; 230] and multi-touch surfaces have found their way into the futuristic visions of human-computer interaction seen in TV shows and movies (e.g. "James Bond – Quantum Of Solace" and "The Day the Earth Stood Still" [212]). With today's technology it is now possible to apply the basic advantages of bi-manual interaction [34; 60; 121; 137; 249] to any suitable domain. Multi-touch surfaces can be easily integrated into systems supporting interaction of multiple users with two-dimensional data sets.

2.4.3.2 Multi-Touch & GIS

The GIS wallboard [66] is a conceptual example of an electronic white board envisioned to support sketch-based gestures to interact with geospatial data. Other systems, designed especially for tablet PCs were implemented by Oviatt [162] and Egenhofer [56]. Sharma et al. [217] concentrate on multi-modal interaction (speech and gesture) with a large dynamic map display. They evaluated their system in a crisis response scenario with real users. McEachren [132] et al. provide a good overview of large format map applications that support collaborative visualization and decision-making. Often virtual globes are used on large interactive multi-touch screens. NASA World Wind[1] is the second biggest player in the virtual globes market behind Google Earth[2]. While Google Earth is targeted at a general audience, NASA World Wind can be more easily customized for specialized groups of users. As a browser plug-in, Microsoft's Windows Live Local Earth 3D[3] also provides an interface to a variety of high-resolution satellite images and maps. ArcGIS Explorer[4] from ESRI, the leader in the professional GIS market, is a client for ArcGIS Server and supports WMS (Web Map Service). Mice and keyboards can be used to navigate a virtual globe, but are not optimal devices for this purpose. Special 3D-space mice[5], however, do exist to operate Google Earth

[1] `http://worldwind.arc.nasa.gov/java/`
[2] `http://earth.google.com/`
[3] `http://www.bing.com/maps/`
[4] `http://www.esri.com/software/arcgis/explorer/index.html`
[5] `http://www.3dconnexion.com/3dmouse/spacenavigator.php.`

efficiently. Multi-touch has been shown to work well with large-screen virtual globes due to its support of multi-finger and bi-manual operation.

2.4.3.3 Multi-touch with 3D content

VR systems using tracking technologies and stereoscopic projections of 3D synthetic worlds have a great potential to support better exploration of complex data sets. However, interaction with stereoscopic displays is difficult at times. For instance, in a stereoscopic representation of a 3D scene it may be hard to access distant objects [25]. This applies in particular if the interaction is restricted to a 2D touch surface. Objects might be displayed with different parallax paradigms, i. e., negative, zero, and positive parallax, resulting in different stereoscopic effects. Interaction with objects that are displayed with different parallaxes is still a challenging task in VR-based environments. In addition, while the costs as well as the effort to acquire and maintain VR systems have decreased to a moderate level, these setups are only used in highly specific application scenarios within some VR laboratories. In most human-computer interaction processes, VR systems are only rarely applied by ordinary users or by experts – even when 3D tasks have to be accomplished [5]. One reason for this is the inconvenient instrumentation required to allow immersive interactions in such VR systems, i. e., the user is forced to wear stereo glasses, tracked devices, gloves etc. Furthermore, the most effective ways for humans to interact with synthetic 3D environments have not finally been determined [5; 25]. Even the WIMP desktop metaphor [153] has its limitations when it comes to direct manipulation of 3D data sets [40], e. g., via 3D widgets [41]. And as a matter of fact, 2D interactions are performed best with 2D devices usually supporting only two DoFs [82; 196]. Hence, 3D user interfaces are often the wrong choice in order to accomplish tasks requiring exclusively or mainly two-dimensional control [5; 82]. Most 3D applications also include 2D user interface elements, such as menus, texts and images, in combination with 3D content. While 3D content usually benefits from stereoscopic visualization, 2D GUI items often do not have associated depth information. Therefore, interactions between monoscopic and stereoscopic elements, so-called *interscopic interactions* [219], have not been fully examined with special consideration of the interrelations between the elements.

While multi-touch has shown its usefulness for 2D interaction by providing more natural and intuitive techniques such as 2D translation, scaling and rotation, it has not been considered if and how these concepts can be extended to 3D multi-touch interfaces. In particular, the challenge of 3D interaction with stereoscopically displayed objects using a multi-touch user interface has not been examined at all [220]. A benefit of multi-touch technology in the 3D context is that the user does not have to wear inconvenient devices in order to interact in an intuitive way. In combination with autostereoscopic displays, such a system can avoid any instrumentation of the user, while providing an advanced user experience. However, the DoF are restricted by the physical constraints of the touch screen.

Only some researchers have addressed the problem of 3D interaction on any 2D surface. Grossman et al. [75] presented a suite of gestural interaction techniques for use with a spherical 3D volumetric display. However, in this setup the user can already use 3D inputs on a touch sphere in comparison to available inputs on a two-dimensional plane. To allow interactions in the 3D space Benko et al. [16] introduced the *Balloon Selection*, a 3D interaction technique that is modeled after the real-world metaphor of manipulating a helium balloon attached to a string. Balloon Selection allows for precise 3D selection in the volume above a tabletop surface by using multiple fingers on a multi-touch sensitive surface in conjunction with a 3D tracked glove. Although, this technique enables 2.5D selection, real 3D interaction has not been considered in their work.

3

Pad Sized User Interfaces for Spatial Information

In this chapter, pad-sized interfaces for spatial information are discussed. Our first topic and the main focus of this chapter are interfaces that allow the user to combine the advantages of paper maps with the advantages of mobile devices, a novel interaction principle we call mobile map interaction (MMI). This implementation uses a mobile phone as a magic lens that the user can combine with a paper map to get additional dynamic and personalized information. This interaction schema was first described in [211]. We have shown that these interfaces can support people in their daily life's [192] and that this interaction technique is superior to two other navigation techniques that only use the mobile device display for map visualization [194][1].

Additionally, we informally observed important differences in visual exploration strategies. Some subjects constantly focused on the mobile display to find the next target, others switched their gaze to the background map to identify a target region, moved their device to that area and finally fixated the display again.

In the second part of this chapter, we focus on a similar study done with mobile camera-projector units. These pico-projectors can overcome several problems that we

[1] In a formal user study, the performance of three methods for map navigation with mobile devices was compared. These methods were joystick navigation, the dynamic peephole method without visual context, and the magic lens paradigm, using external visual context like paper maps. We measured user performance and motion patterns and collected subjective preference via questionnaires. The results demonstrate the advantage of dynamic peephole and magic lens interaction over joystick interaction in terms of search time and degree of exploration of the search space. Follow up studies [192] showed this advantages with additional details.

investigated in [192] and therefore cannot be clearly classified in one of Weisers device classes. While the pico-projectors are pad-sized, these new mobile camera-projector units are discussed in this section. Namely, projectors eliminate an important drawback of magic lens interfaces observed above: users have to switch their attention between the magic lens and the information in the background. In this second section, we present solutions to this problem by using a lightweight mobile camera projector unit to augment the paper map (or other media) directly with dynamic digital information.

Finally, in the third part of this chapter, we focus on a second application called Photomap, which uses images of 'You are here' maps taken with a GPS-enhanced mobile camera phone as background maps for on-the-fly navigation tasks. We discuss different implementations of the main challenge, namely helping the user to properly georeference the taken image with sufficient accuracy to support pedestrian navigation tasks.

3.1 Mobile Map Interaction

The general idea of MMI is to use a mobile device like a magic lens over a paper map to augment the map with additional information. The detailed interaction properties of MMI are described in the context of MMI research we have done in the following real-life application domains:

- Risk Management,
- Tourism and
- Games.

All three subsections are structured in the same way. First we present the demonstration scenario and briefly discuss the related work for the particular domain. We then discuss the basic interaction concept for each domain, followed by an evaluation, an implementation section and a discussion.

Before beginning our domain-based discussion, it is important to note that in the first and second domain (Risk Management and Tourism), we also show how MMI can be combined with different modalities (in this case speech input and audio output) to improve the access to and communication of spatial information.

3.1.1 MMI in Crisis Response Scenarios

Efficient and effective communication between mobile units and the central emergency operation center is a key factor to respond successfully to the challenges of emergency management. Nowadays, the only ubiquitously available modality is a voice channel through mobile phones or radio transceivers. This makes it often very difficult to convey exact geographic locations and can lead to misconceptions with severe consequences, such as a fire brigade heading to the right street address but in the wrong city[1]. In this paper we describe a handheld augmented reality approach to support the communication of spatial information in a crisis response scenario. The approach combines mobile camera devices with paper maps to ensure a quick and reliable exchange of spatial information.

Recently a misconception of a location led a German fire brigade to a wrong place (the fire brigade headed to the right street address, but in the wrong city) in this case with terrible consequences. The webpage[2]gives a very similar example of an ambulance that was sent to a wrong address. However, information on locations is not only important when talking about street addresses, but also when mobile units need to enter hazardous areas, such as a flooded area or a forest fire. In such situations graphics and maps are needed in addition to verbal communication between the mobile unit and the control center in order to ensure that the right action is performed in the right location. Currently paper-based maps are used in conjunction with the voice channel (either through phones or radio transceivers) to achieve agreement on locations between mobile units and the control center. Computer-based displays (as provided in tablet PCs, PDAs and mobile phones) are infrequently used, because they are too small and of too low resolution to handle this task successfully. Larger displays are too heavy or not ruggedized enough to suit the task. Even in the advent of electronic paper [27] one cannot expect robust large-scale electronic maps in the near and midterm future.

For these reasons we propose an approach where smaller mobile devices (which can easily be ruggedized) are combined with conventional large scale laminated paper maps of the environment. In particular, we argue that if a modern camera-equipped mobile phone is used, information about locations can be easily conveyed by using the available voice channel and a mobile augmented-reality approach. As said in the introduction

[1] http://www.wdr.de/themen/panorama/brand03/toenisvorst_wohnungsbrand/index.jhtml
[2] http://firegeezer.com/2008/05/02/ambulance-sent-to-wrong-city

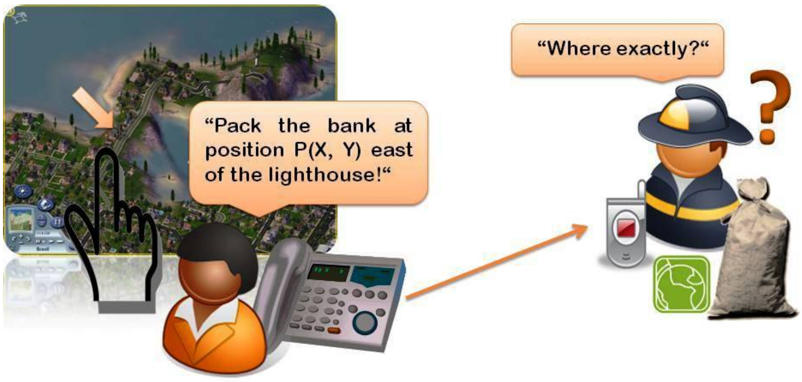

Figure 3.1: Problem of exchanging spatial information via the voice channel.

of this section in this approach a mobile device is moved like a *magic lens* [18] over a printed map and location information is superimposed onto the map in real time. We show that with this additional modality communication errors are heavily reduced in comparison to the voice channel alone. Large scale maps can provide an excellent overview of the whole situation, which would not be possible with the device screen alone.

The integrated cameras can also be used to quickly capture up-to-date photographs of the situation on the spot. These documents can then be linked to positions on the map as multimedia annotations created by the mobile units. We have developed multimodal interaction concepts suited to the emergency tasks at hand, which are presented in the remainder of the paper.

The remainder of this subsection is organized as follows. In Section 3.1.1.1 we briefly describe a scenario which underlines the significance of our approach. We introduce the multimodal interaction concepts in Section 3.1.1.2. Section 3.1.1.3 reports on the implementation and a discussion and summary is provided in Section 3.1.1.4.

3.1.1.1 Scenario

In our hypothetical scenario, floodwaters strike the eastern part of Germany, eventually causing a flood that affects villages and cities, for example the historic downtown of Dresden, along the river Elbe. Evacuation alerts are sent out to the affected populated areas. The Technisches Hilfswerk (THW)[1] constantly inspect all banks and dams to ensure that leaky parts are detected early. On an inspection walkway a THW team gets a call from the Emergency Operation Centre (EOC): A helicopter has reported a leaky spot in the bank on a specific position P with the geographic coordinates *(X, Y)*. The bank is about to break and a THW team has to prevent this by sealing it with sandbags. In this case it is necessary to communicate the exact position P from the EOC to the mobile unit to be sure not to waste valuable time. Of course, this communication channel is not unidirectional. There are situations in which this communication is needed in the other direction as well. Let us assume, for example, that the THW team has detected a leaky spot and needs more sandbags and more manpower to protect the bank. Describing the position verbally or exchanging the coordinates of a position P literally as X and Y via the voice channel is a time consuming and error-prone task [72].

The way dynamic geo features, such as the approaching floodwater, are visualized on the map obviously has a strong effect on the usefulness and understandability of the presented location information. This can be easily done in the EOC by using large dynamic map displays, but it is a challenging task to realize a dynamic map application for THW teams in the field. Tablet PCs still are unhandy and have a short battery runtime. PDAs and mobile devices with network connectivity can provide dynamic maps with the desired content by querying an adequate web service, but these maps suffer from the small display size. Because visual context is lacking, it is often hard to identify locations and landmarks on these maps, rendering them rather useless.

In the following we describe a concrete application of a communication system between mobile units and an EOC trying to support communication of spatial information between them, as well as other critical tasks in a CRS like *Monitoring and Reporting* and *Identification and Notification* [251]. This work differs from the related work in that we apply a combination of two devices typically used by mobile units, namely mobile camera phones and paper maps. Combining both we can provide a robust and intuitive way to easily support the communication of any kind of spatial information.

[1]The Technisches Hilfswerk (THW) is the governmental disaster relief organization of the Federal Republic of Germany.

3.1.1.2 Basic Interaction Concept

As described before, the communication between the EOC and the units in the field typically happens via the voice channel today. This modality makes the rapid and accurate communication of spatial information difficult. In our approach the voice channel is augmented by visual information that is presented on the mobile device screen (see Figure 3.20). Since the device screen is too small to help the user to easily acquire overview knowledge for a spatially extended area, we combine the handheld display with a medium to large size laminated paper map (A5 to A2 size). The paper map provides the static long-term information of the area, which is not affected by the crisis event, such as landscape features or street names. The paper map can still be updated regularly. Since even laminated paper is cheap, a new map could be handed out once a day, which includes recent changes induced by the crisis (see Figure 3.3). The paper map is easily transportable, robust, does not require any power supply, and in effect constitutes a high-resolution display and can be used as a fall-back solution. The mobile device with an integrated camera is used as a magic lens for the paper map. It analyzes the part of the paper map that is visible in the camera view and determines the current focus position on the map [193]. The device shows dynamic overlay graphics and textual information that reflect the most up-to-date state of the crisis. As the situation changes or as new requirements and tasks are communicated by the EOC, the map is updated in real time. During a phone conversation, virtual annotations of map areas can be created in the EOC and the result becomes immediately visible on the display of the first responder (or vice versa). This concept achieves maximum transfer of spatial information at minimum reliance on special-purpose technology and is therefore expected to be quite robust. The camera phone as well as the paper (or cardboard) map can easily be ruggedized. By combining the voice channel and the visual channel, tasks and needs can effectively and unambiguously be communicated.

1:1 Communication between the control center and the first responder Figure 3.2 illustrates one-to-one communication between the control center and the team member. Figure 3.2, left, is an example of communication from the EOC to the THW team. The EOC issues a command to the leader of a THW team that involves a spatial target to operate on—in this case a bank that needs to be protected. Figure 3.2 (right)

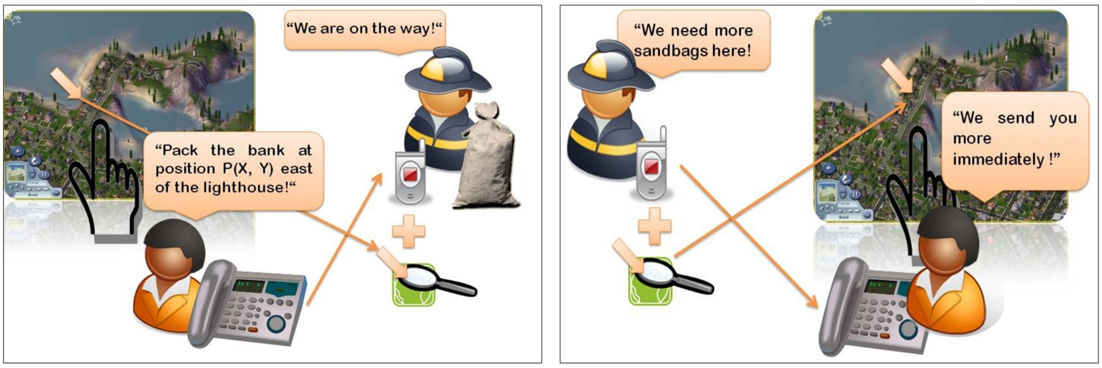

Figure 3.2: Supporting the exchange of spatial information by combining the voice channel and the visualization on the mobile device.

shows, how a specific requirement, again involving a spatial component, is communicated from the first responder to the EOC. In both cases the involved location can be accurately determined by the first responder, when using the *magic lens* approach.

1:n communication between the control center and the first responder Normally, a single task is collaboratively performed by a large number of THW teams in the field. The EOC has to issue out the task to the group as quickly as possible. Again, the *magic lens* approach enables the team members to individually receive voice commands and review the graphical annotations of the map using their personal displays. Even a single large paper map can be collaboratively shared by multiple THW team members by using multiple camera phones.

This one-to-many communication is not limited to the EOC. One team member might need to quickly send information about a spatial object to the others nearby. In this case he or she would create a graphical annotation or select an icon from a library of icons, place it at the intended location, and attach a textual or voice annotation to this icon. This annotation would then be visible to the other first responders in the vicinity.

3.1.1.3 Implementation

Figure 3.4 shows the prototype architecture. The mobile device is tracked over the map using the algorithm described in [193], which relies on a tiny dot grid that is printed

Figure 3.3: 1:n Communication between the control center and the first responder (left). Showing an augmented (or alternative) view of the printed map (right).

on the paper map. This dot grid allows pixel precise augmentation of the paper map on the mobile device screen in a robust way (see Figure 3.3 left). Given the different test maps with about 500 patches, the current implementation (with 12×12 samples per patch) processes 8 to 12 frames per second on a Nokia N95, depending on the number of patches found in the camera frame. The map can be rumbled to a certain degree without severely affecting the recognition rate. The spotlight integrated next to the camera lens on many camera phones can be used to illuminate the map in low light conditions. The voice channel is established via a normal phone call (see Figure 3.4, green arrow, labeled *voice channel*). Spatial information displayed on the large map display, like the actual flood water level, in the EOC can be easily transferred via a specific XML protocol using GSM or, if available, via UMTS. (see Figure 3.4, red arrows, labeled *position data*). Off-screen information can be visualized with the *Halo* technique [10] or the *Wedge* technique [76]. The former draws a ring around off-screen objects that reaches into the border of the device display. From the position and curvature of the ring segment the user can infer the approximate direction and distance of the off-screen object. The Wedge technique increases scalability by drawing "pie Wedges" instead of circles and lays them out such that overlap is avoided if possible. In user studies it was found that both techniques provided equally good cues about the distance of the off-screen object.

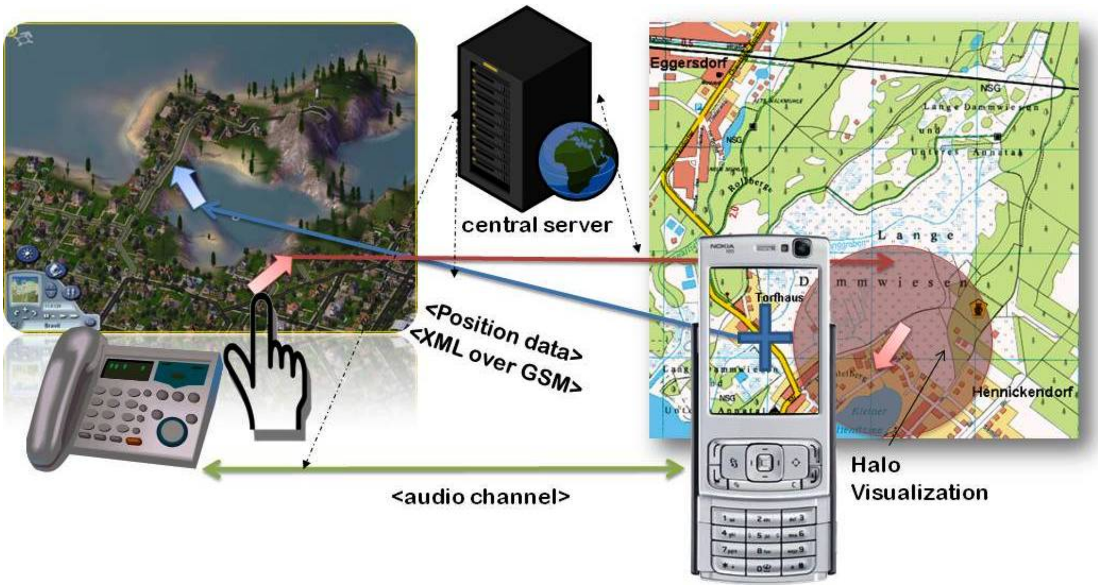

Figure 3.4: System architecture of the application.

3.1.1.4 Discussion & Future Work

We described a method to support the communication of spatial information in crisis response situations. The method combines standard mobile camera devices with printed maps to ensure a quick and reliable exchange of spatial information. We use a robust mobile augmented reality tracking approach to generate overlay graphics at interactive rates. The tracking technique has been successfully deployed in a number of applications [211], [88], [208], [214] and [180]. It was mainly used for implementations in this thesis. We are convinced of the applicability and practical value of the proposed concepts and look forward to take this work further. In terms of usability tests we particularly want to focus on users under stress with little time and little attention available for mobile map interaction.

3.1.2 Wikear

Many mobile applications that lead tourists to landmarks and businesses ignore the educational component of tourism. The systems that do satiate the tourist‘s desire for learning about visited places require so much costly custom content development that they can only be implemented at very limited scales. Moreover, these systems quickly

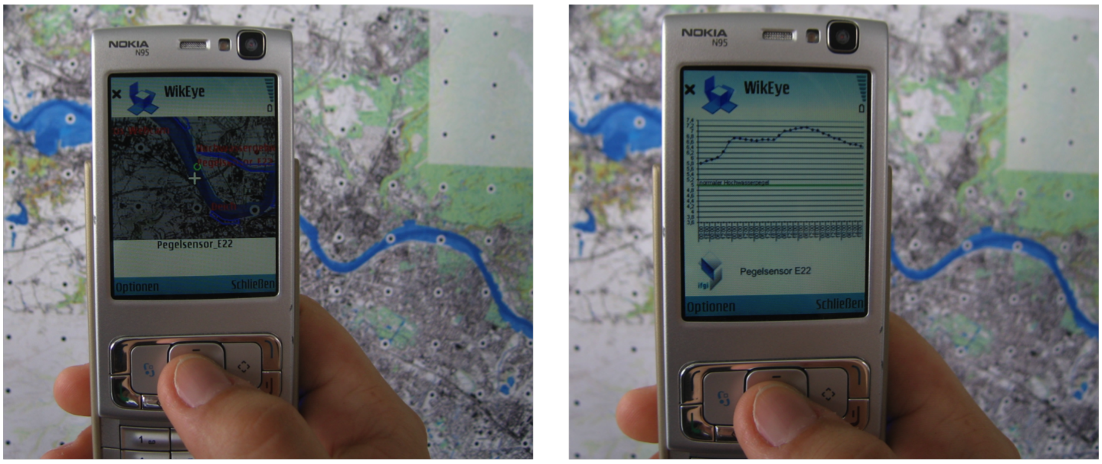

Figure 3.5: Application in use. A user is acquiring additional information about a water level sensor using the mobile phone like a magic lens over a paper map.

fall out-of-date and continually have to be manually updated. In our approach, called WikEar, data mined from Wikipedia is automatically organized according to principles derived from narrative theory to be woven into educational audio tours starting and ending at stationary city maps. The system generates custom, location-based guided tours that are never out-of-date and are ubiquitous even at an international scale. WikEar uses, as described earlier, a magic lens-based interaction scheme for paper maps, which have been shown to be particularly important in the tourist experience. By leveraging on the wide availability of large public city maps, WikEar avoids the costs of GPS and the interaction problems of small screen map programs (see figure 3.6).

3.1.2.1 Scenario

According to a Berlin tour guide with whom we talked about WikEar, many tourists would appreciate at least something of an educational experience out of their expensive efforts to see the world. These tourists need more than the place names and directions provided by Google Maps or a GPS-based device. Currently, they have two options

1. paper guidebooks and

2. customized, highly localized mobile device applications.

Both of which have severe content limitations and are not available for many locations. Writing, editing and post-production of content in these types of tourism tools can be expensive and overwhelming. In addition, the content can quickly become out-of-date. WikEar is an attempt to wed the pervasive and easily-updated content of a mobile map application with the educational capabilities of a guidebook. The content is one of the key factors for such applications. The premise of WikEar is quite simple: The user stands in front of a public city map and selects a spatial feature (such as a building or landmark) using her camera phone as a magic lens, as described in earlier. A guided audio, narrative-based tour between the location of the city map and the destination is then delivered to the user, with the intent that she will listen to the story as she travels to the destination. To track the mobile device relative to the map we again use the magic lens tracking technology [193]. The guided tour comes in the form of a narrative that is automatically mined from Wikipedia by the Minotour system [86; 89]. The output is rendered in audio form using text-to-speech (TTS) technology. As in [86] , we follow the advice of Isbister et al. [99] and others reseraches, and assume that the best tours are those that weave a story as one travels through the landscape. This is our basis for implementing narrative theory methods into our tour generation algorithms, a process that will be described in more detail later. The goal is that users of WikEar will know about historical and current themes present in the regions they have visited. Finally, it is important to note that by combining city maps with guidebook-like content, we meet the call of Brown and Chalmers [28], who suggest that one of the greatest challenges in mobile tourism technology is greater integration of paper maps and electronic guidebooks.

3.1.2.2 Basic Interaction Concept

Imagine you are a tourist in a city or a region and want to learn about the place you are visiting. You find a public city map and hold your camera-enabled cell phone up to the map. You then select your destination, probably one of the city's or region's more famous locations. A story designed to match your start location, destination, and traveling time is then delivered to your phone over the Internet, ready for you to listen to the parts of the story as you head towards your destination, interspersed with directional guidance. The actual guidance is not part of WikEar yet, but could leverage existing tour guide technology [38]. Once you reach your destination you can find

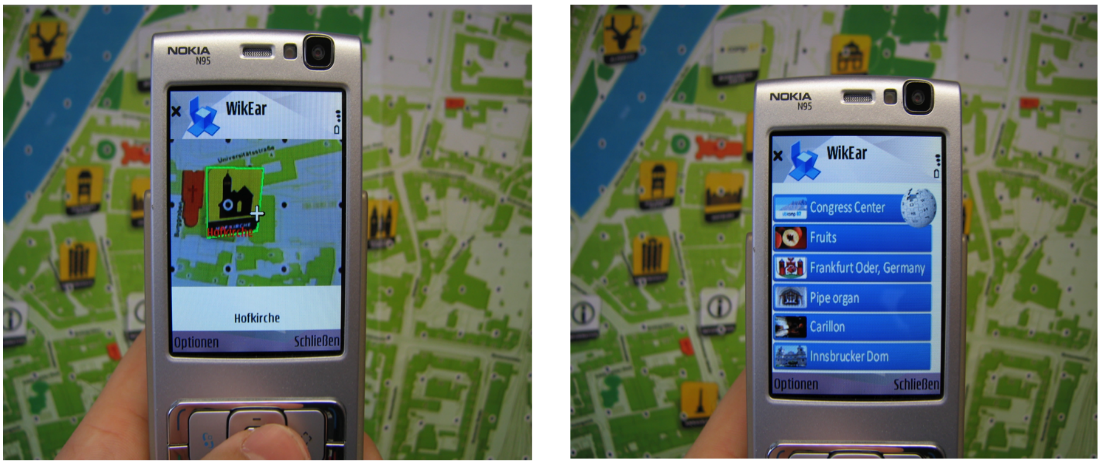

Figure 3.6: User interacting with Wikear. Selection of a destination the system play back an auto generated story.

another city map and repeat the process or access another tour you have already pre-downloaded. A key benefit of this interaction framework is the independence afforded to the tourist. The tourist is not limited to prescribed tour paths or restricted by a paucity of available content. In fact, the only check on the tourist's movement is that, due to the algorithms used in tour generation, the start (location of the map) and end destinations must have associated and geotagged Wikipedia articles which might even have been written prior to the trip by the tourist herself. As such, WikEar, as an instance of mobile tourism technology, is in line with Weiser's vision of Ubicomp [239] in that technology should be much more supportive of spontaneous choices and desire for flexibility.

3.1.2.3 Narrative Theory Approach & Wikipedia Story Mining

As is noted in the introduction, researchers in the field of intelligent narrative technologies state that a successful educational tour is one that weaves a story as one travels through the landscape. As such, the optimal approach to automatic educational tour generation is one that gives narrative a central role. However, before explicating the algorithm used to generate guided tour narratives from Wikipedia it is first necessary to highlight some characteristics of Wikipedia.

Aside from being the largest-ever compilation of user-contributed human knowledge, Wikipedia has three other important features that we exploit in WikEar:

1. From a geospatial viewpoint, Wikipedia can be split up into two types of articles, two of which are vital to the understanding of our algorithmic approach to narrative generation: georeferenced articles and non-georeferenced articles.

2. Because Wikipedia is collaboratively edited, the average of Wikipedia article s are contributed to by at least seven different authors [32]. This fact, in combination with the encyclopedic tone of nearly all of Wikipedia, allows for self-contained paragraphs that are extremely disconnected from each other. This allows us to treat Wikipedia paragraphs as individual entities we have named snippets, and re-order them in any way our algorithm demands.

3. Wikipedia has an elaborate graph structure, a fact that is at the very center of our narrative generation approach. We define the Wikipedia graph W as the set of Wikipedia articles and the associated link structure.

The goal of our algorithm is to find the path through W between the articles about the start and destination spatial features that most resembles the optimal path we have defined according to our approach to narratology (as simplified in figure 3.7). We then take the snippets hosting the links in the nearest-to-optimal path and present them in order to the user as a guided tour. In other words, the algorithm optimally queries and restructures Wikipedia's content to

1. provide location-based information, and

2. to do so in a way that can be perceived as a narrative.

In this re-structuring process, the snippet is treated as atomic. The algorithm determines the number of snippets to include in each tour using an estimate of the travel time between the start and end destinations, thus preventing us from using standard shortest-path solutions to this problem. The snippets are converted to audio using text-to-speech technology, admittedly a drawback in user experience given the current state of that technology. In our prototype, we have used actors to make listening to the tours more enjoyable. How do we define the optimal path through W? According to our adaption of narrative theory, the optimal narrative is the one that best integrates

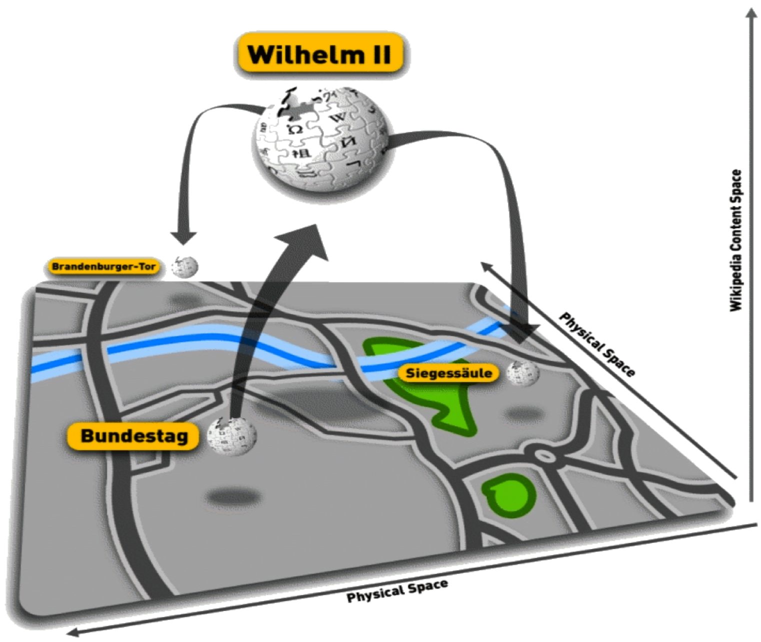

Figure 3.7: Simplified diagram depicting a location based narrative from Wikear.

the structural cues of unity and development into the generated text. In our current prototype version of the algorithm, we have found that such an optimal narrative can be approximated as in figure 3.8 (i), using number of inlinks (indegree in W) as the primary variable.

3.1.2.4 Implementation

Wikear is fully implemented for Nokia mobile camera phones (S60 3rd edition) based on [193]. Mining for the optimal narrative takes place online. This procedure is supported by an extensively parsed database of Wikipedia information, the result of a large offline pre-processing step. The input to this step is one of the frequent Wikipedia dumps, which contain a snapshot of Wikipedia, i.e. all the text of every article in Wikipedia at a given cut-off point in time. With minor modifications, our parser will work with Wikipedia dumps of any Wikipedia-supported language, although at the moment we have built in support for only English and German, Wikipedia's two largest encyclopedias.

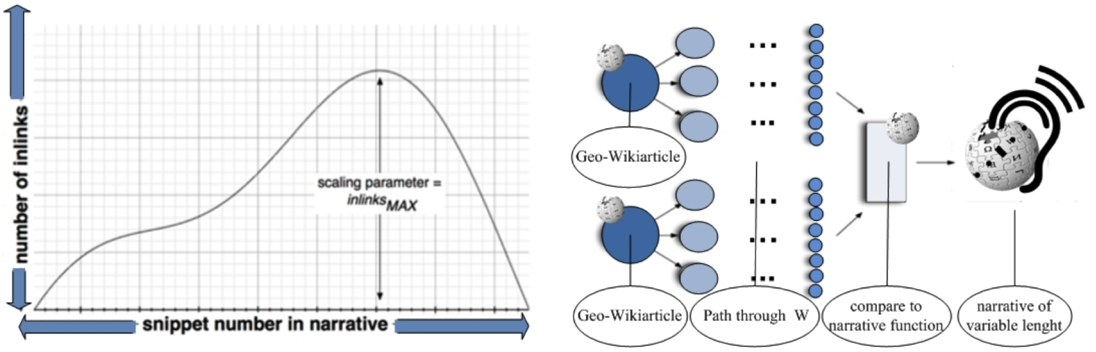

Figure 3.8: (i) The ideal Narrative Function. The narrative algorithm chooses the path through the Wikipedia graph that most resembles this function. (ii) A diagram depicting the operation of the path finding portion of the narrative algorithm. Once found, the paths are evaluated by narrative-theory informed function. Finally, the narrative with the best output (the one most similar to the optimal narrative (i)) is selected.

3.1.2.5 Evaluation

In conjunction with the Ninth International Conference on Ubiquitous Computing that took place in Innsbruck in September 2007, we conducted a semi-formal evaluation of the quality of our generated stories and the new interaction metaphor. We demonstrated our WikEar prototype at the conference and asked interested attendees to give us feedback on the stories and the interaction. We distributed maps (A4 size) and testers were given the option of using one of three Symbian S60 (third edition) phones. We set up a web-based evaluation system where users could listen to two stories again and answer the eight following questions with a six-point ranking system (1=good to 6=bad):

1. How cohesive was this story? Did all the pieces seem to make sense together? (Ranking: from “It was unified” to “It was completely disjointed”. A “1” on this question indicated that the stories were “unified” and a “6” indicated that the stories were “completely disjoint”).

2. Was there a sense of development from beginning to middle to end? Did you feel that the story was progressing to a specific point? (Ranking: from “The story developed as I read. I often thought about what might happen next.” to “I felt lost. The story had no clear direction”. Similarly, a “1” on question two indicated

that the user found that the “story developed as [they] read” and that they often thought about “what might happen next”. A “6” on the same question signaled that the “felt lost” and that the story had no clear direction.)

3. Do you think you learned more about Innsbruck than you would have if just using a paper map? (Ranking: from “Yes” to “No”).
4. If this service were available for your next vacation, how much would you use it? (Ranking: from “Frequently” to “Never”).
5. How well did the interaction between the paper map and the camera phone work? (Ranking: from “Very good” to “Very difficult”).
6. How difficult was icon selection by pointing with the camera phone? (Ranking: from “Very Easy” to “Very difficult”)
7. How did you like the service? What else would you expect from such a service? (Free text field).

We got feedback from 21 conference attendees, three female and 17 male (one participant did not specify gender). All were conference attendees at UbiComp and had a mean age of 34.1 years (SD=6.8). One participant had problems with the evaluation website so we excluded that person from the data that we analyzed. UbiComp attendees are a peculiar population to do an evaluation of UbiComp technology, but we believe this to have had only little effect on our results, because of the very mixed backgrounds of the attendees. The main problem of evaluating auto-generated stories is that these stories take an experimental form. Since most traditional forms of narrative are produced in relationship to cultural, stylistic, and genre conventions, these stories may be evaluated with reference to a norm. Thus when a user encounters a Wikear story, it will likely seem deficient when compared to a traditional literary narrative (if the user even considers it a traditional narrative). This is why we had to test the user directly to see if they used narrative schema in comprehending the material. We isolated two distinct indicators of narrative comprehension and tested for these: unity and development. The overall opinion on the quality of the WikEar stories and the level of narrative comprehension is shown in figure 3.9. Because the standard deviation on both narrative questions (questions one and two) was so high, very few conclusive results can be drawn.

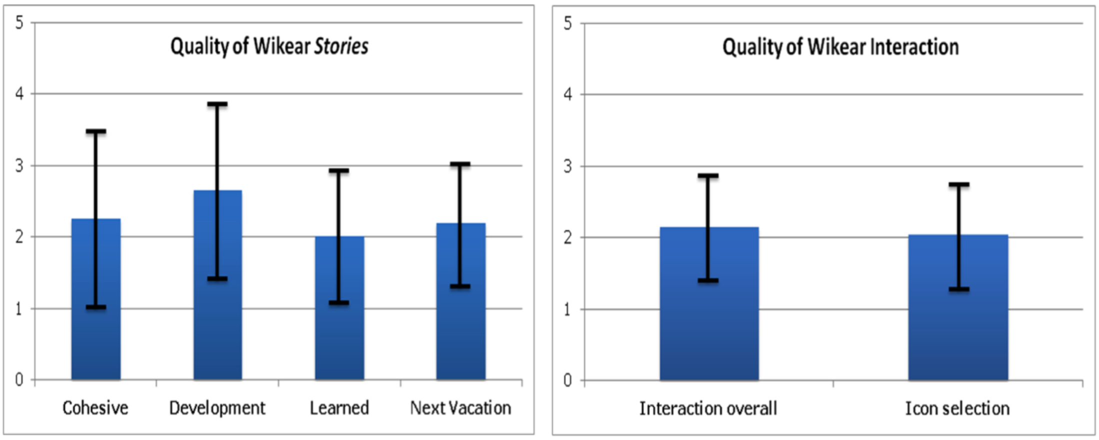

Figure 3.9: User Feedback on Quality of Wikear Stories and Wikear Interaction.

The mean of question one was 2.25, but because the standard deviation was 1.37, the 95 percent margin of error covered nearly the entire range of possible values. The same occurred with question two, which had a mean of 2.65 and a standard deviation of 1.39. We would like to see a more definite result for both unity and development. However, the fact that a large majority of people report positive results of "1s" (3) and "2s" (8) is very encouraging. We conclude that although there are some indicators of success from our survey methodology, we need to make some improvements to better bind the snippets together and to underline the development in the stories. We are currently working on an enhanced version of the narrative generation algorithm (see the future work section) and are also testing interface-based methodologies, such as showing the paragraphs on the screen (see figure 3.9). It is important to note that we got the best rating – a mean of 2.02 – for the question about education, an indication our narratives accomplished our pedagogical goal . Additionally, most people would like to take such a system with them on their next vacation, another encouraging statistic. As is shown in figure 3.9, the user had no problems interacting with the paper map.

3.1.2.6 Discussion & Future Work

We have presented a prototype system that leverages Wikipedia's extensive content and structure in an effort to fill a gap in the mobile tourism technology field. While it will require better text-to-speech for wide-scale adoption, WikEar successfully demonstrates our data and methodology approaches. It is also important to note that we have implemented an application that requires a great deal of spatial context without any GPS technology at all, greatly reducing the cost on the user side. In addition, tourism boards simply need to outfit their existing network of city maps with marker-enhanced maps in order to support WikEar. The total cost is thus much less than that of every user having to buy spatial data and a GPS system, and accessibility is much greater. WikEar incorporates many new concepts that give rise to a plethora of ideas for refinements and further work. First and foremost, we believe that while navigation is a solved problem [114] and thus not a high research priority, it would be a necessary part of a user-ready version of WikEar. One possibility is to implement landmark-based navigation technology, giving directions such as "head toward the big building", which is more conducive to how tourists prefer to navigate than using a street-based system [28]. Also, we have already begun investigating ways to improve our narrative generation process, as well as looking into ways to evaluate generated narratives. Finally, we are also continually working on ways to improve our tracking technology to the point where no obvious modifications are needed on the printed map.

The first item of future work is developing a more formal user study around our narrative evaluation methodology discussed above. Another evaluation idea we will be exploring is the importation of our generated narratives into travel wiki websites. It is our intent to judge the value of the narratives based on the wiki community's reaction to them.

3.1.3 Sight Quest

Sight Quest is a location-based game for mobile camera devices that uses paper maps as game boards played in urban spaces. The players have to solve different quests requiring knowledge about geographic objects on the map as well as real and fictional attributes of these objects. The players move their mobile camera device like a magic lens over the map in order to make the attributes of the objects visible and to control interaction with the game. The game has to be played in different locations to solve all quests. *Sight Quest* can be played at many different locations in the city where interactive stationary maps are available, e.g., at subway stations and public places. *Sight Quest* is aimed for spontaneous interaction, therefore it is designed to attract players in contexts where they are waiting and have some spare time.

3.1.3.1 Introduction & Scenario

The game boards of *Sight Quest* – paper maps – are everywhere around us. In many different places they provide us with important information about our surrounding. This information is not only useful for tourists but also to locals who want to explore unfamiliar places. In former research we have combined the advantages of mobile devices with the advantages of paper maps using the mobile device like a magic lens over the map and thus augmenting the map with additional spatial information. In doing so, we provide a useful dynamic and personalized extension to classical paper based media. In contrast to our former work we apply the map lens approach to a location based map game. While playing the game *Sight Quest* users can not only have fun but also learn more about their current environment. Paper maps in cities are ubiquitous and provide information about our urban environment. These paper maps are usually designed to address the most common questions of average users and therefore contain only general information, such as street names and places of interest. More specific information – such as the locations of automatic teller machines (ATMs), pubs, shops and restaurants – is typically not included for reasons of map complexity. In particular, general public maps are not designed to be game boards. Including printed game-related information is not acceptable, mainly because of the visual complexity reasons mentioned above. In the game we present here, a camera-equipped mobile device is used as a magic lens to display game-related information and for playing the game.

Linking mobile games to public paper maps is particularly appealing, because of their high distributions in urban spaces. With the location based game *Sight Quest* we aim at improving the attractiveness of paper maps and even the knowledge about the surroundings of the user. The kind of game presented here requires no additional infrastructure beyond the user's camera-equipped mobile device and an online server that provides the game content. The *Sight Quest* game is targeted at a broad audience, basically anyone who is the owner of a camera phone and occasionally has to wait next to paper maps. The game interaction should therefore be as simple as possible and the game should require virtually no setup time. In this case, the waiting time for the next subway at a platform can be used to play *Sight Quest* and learn something about the local vicinity. The game client augments the camera image of the physical map with

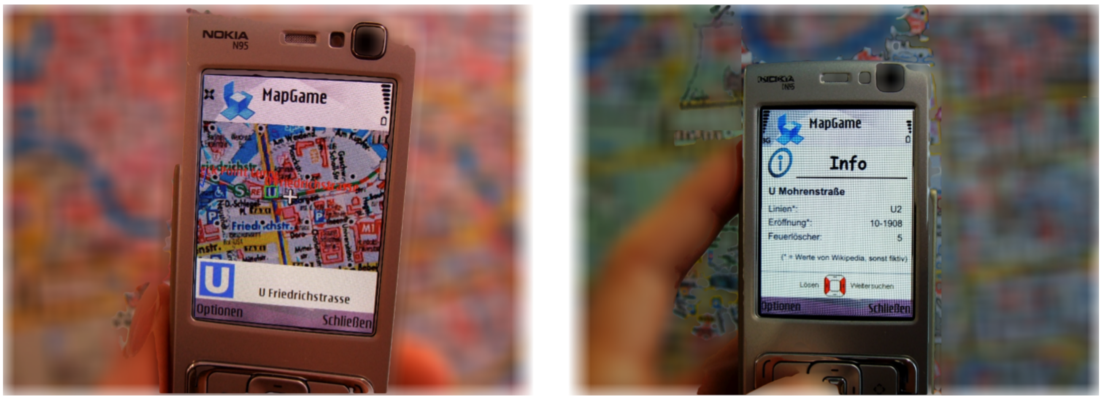

Figure 3.10: Application in use – Exploring the map for possible solutions of a quest. Is that the right subway station you looked for? (Please note, that the map background is artificial blurred).

dynamic content, for instance by displaying additional information about the (fictional) mood of the taxi drivers at a selected taxi stand. This information can then be used to solve a quest of the game. Figure 3.10 shows the camera view with a highlighted sight. Pressing the joystick button shows the attributes of that sight. This interaction allows the player to use the paper map for overview information augmented by their mobile device as a magic map lens.

Users can playfully learn more about their environment with the *Sight Quest* game. Even though some of the attributes are obviously fictional, like the taxi drivers' mood, the relative spatial locations of sites are nonetheless learned implicitly and help players to get a sense of their surroundings. The game concept of *Sight Quest* is quite simple. The main goal is to get as many game points as possible by solving different quests. On every map at different locations in the city, up to 10 quests are "embedded". The quests are depending on the part of the city shown on the map. For example the player has to find the ice cream vendor with a particular kind of ice cream next to a special landmark. When solving this quest the player is awarded with points (the faster she is the more points she gets). The high score could then be uploaded to a central server to attend a lottery. A fundamental requirement of such a magic lens game is to track the mobile camera device over the physical map in real time. For this purpose a grid of tiny black dots is printed on the paper map to simplify the tracking task. Detailed information about this tracking method can be found in [193]. With *Sight Quest* we developed – to our knowledge – the first interactive game combining a paper map and

Figure 3.11: Interaction schema to solve a quest: (1) player chooses a task, (2a) attributes of first sight, (2b) attributes of second sight, (3a) correct selection, and (3b) wrong selection (Please note, that the map background is artificial blurred).

a mobile camera device. Ubiquitous maps provide a variety of different game boards and therefore nearly unlimited gaming opportunities.

3.1.3.2 Basic Interaction Concept

The application *Sight Quest* consists of two parts: the *information* part, where users can roam around the map freely and look at every point of interest they find, and the *game* part, where players have to solve a given quest as quickly as possible. The application starts in the *information* mode and the player can get additional information to any site on the map by sweeping the mobile camera phone over the map. To switch into the *game* mode the player has to press the joystick-up button. A quest is displayed on the mobile device screen (see figure 3.11) to be solved in a maximum of two minutes.

All quests follow a similar scheme: The player has to compare the attributes of different sites and select the right one. An example quest would be to find the taxi driver in the best mood within a distance of 200m around the “Alexanderplatz”. These tasks involve the usage of the map (because the player has to look for the location of“Alexanderplatz” on the map) and the mobile device (to check the moods of the taxis drivers around “Alexanderplatz” with the magic map lens principle). The player has to check all possible candidates for the right attribute and select the correct one to solve the quest as fast as possible. After showing up the attribute window she can not leave with the middle joystick button (as in the information part of the application), but has to decide whether the selected site solves the task or not (see figures 3.11). The quest can be shown again at any time by pressing the joystick-up button again while interacting with the paper map. It can be also cancelled by pressing the joystick-down button.

If the player believes she picked the right sight to solve the quest, she has to press the left joystick button, otherwise the right joystick button (see Figures 3.11)). This is also indicated in the image at the bottom of the display while showing the information of the selected sight (see figure 3.10). If the player pressed the right joystick button, she returns to the map view and can continue searching without penalty; if she pressed the left joystick button, a new window is popping up and informs her about the result. If one ore more sights are on the display of the device, a short description consisting of the name and an icon representing the category of the sight nearest to the cursor is shown at the bottom of the display (see figure 3.10). Further information can be obtained similar to other applications by “clicking” the sight with the center joystick button. Doing so opens a new window with all the information about the sight (i.e. name and attribute-values of the corresponding category). At the bottom of the new window is an image showing the joystick buttons of the mobile with the middle button highlighted in red, which means that by pressing the middle joystick button again the player can close the information window. In our example map of Berlin (showing the quarters “Mitte”, “Schneberg”, “Kreuzberg”, and “Moabit”) 40 sites are placed and grouped in four categories to keep the complexity low (10 sites per category). Not all sites and quests are available on every map, because different maps have different spatial extents and therefore show different parts of the city. Each site category has a maximum number of three attributes that can be easily compared while playing the game:

1. Subway-stations (attributes: lines, opening of station, number of fire-extinguishers),
2. Ice cream parlors (attributes: sorts, price per scoop, size of one scoop),
3. Snack bars (attributes: price of a hamburger, number of patties, amount of chips served to a menu)
4. Cab stands (attributes: number of waiting cabs, mood of drivers (3 possible states).

3.1.3.3 Implementation

Sight Quest is fully implemented for Nokia mobile camera phones (S60 3rd edition). The actual *Sight Quest* prototype is implemented on a Nokia N95 mobile phone. Due the small size of the necessary metadata, it is possibly to download the map description and the game data even with relatively low bandwidth connections like GPRS and use nearly every paper map in a city, without requiring the installation of local access points. The example map of Berlin we use for *Sight Quest* is printed in color on an A1 sheet (84 $\times$ 59 cm). The recognition algorithm processes about 8 frames per second.

3.1.3.4 Discussion & Future Work

With *Sight Quest* we present a new kind location based game using ubiquitous paper maps as game boards and mobile camera phones as input devices. This extension of our former work has the advantages to provide a mobile game playable at different locations. This game can be played spontaneously by users trying to solve just one quest on the go, or others users who want beat the high score for a particular map. The next steps includes to enable users to play the game in the public with real maps and do gather experiences with which to improve the game play. Further extensions could include multi-player games, where several users could play on one map, or even against other users who are playing at remote maps. We believe that these kinds of augmented reality games have some potential, especially if they are designed for short time periods so that users can play the game while waiting for the train or bus. New target groups can addressed by location based raffle games.

3.2 Mobile Projection

With the miniaturization of projection technology the integration of tiny projection units, normally referred to as pico projectors, into mobile devices is no longer fiction. Such integrated projectors in mobile devices could make mobile projection ubiquitous within the next few years. These phones soon will have the ability to project large-scale information onto any surfaces in the real world. By doing so, the interaction space of the mobile device can be considerably expanded. In addition, physical objects in the environment can be augmented with additional information. This can support interaction concepts that are not even possible on modern desktop computers today. We claim that mobile camera-projector units can form a promising interface type for mobile AR applications. In this section we identify different application classes of such interfaces, namely object-adaptive applications, context-adaptive applications, and camera-controlled applications. In addition, we discuss how the different spatial setups of camera and projector units will have an effect on the possible applications and the interaction space with the focus on the augmentation of real word objects in the environment. Furthermore, we present two examples of applications for mobile camera-projector units and present different hardware prototypes that allow augmentation of real world objects.

Mobile phones are used for a wide range of applications and services in today's everyday life, but still they have many limitations. Aside from the lack of working memory and the small display size is one of the major bottlenecks. Digital projectors are shrunken to the size of a mobile phone. The next step is to integrate them directly into the mobile device. Today several prototypes have been presented, and the first series-production device is already on the market. First prototype phones with integrated projectors already exist and are available on the consumers market but both models are currently not available on the mass market. Such phones could overcome the shortcomings of the small screen and make it possible to present large and complex information like maps or web pages without the need for zooming or panning [81]. Considering the anticipated widespread availability of phones with integrated cameras and projectors in just a few months, surprisingly little research has been conducted so far to investigate the potential of such a mobile unit (in the following we use the term mobile camera-projector unit as a synonym for a mobile phone equipped with a camera and a projector). We identify different application types based on the spatial configuration of the camera and the projector. As part of this classification, we derive three different application types using of different spatial layouts of cameras and projectors: congruent setups, partially intersecting setups and disjunct setups. Such a classification is useful to structure the design space of mobile camera-projector systems, because we think other researchers can categorize their applications to focus on specific problems and topics of each type. Our approach needs to be elaborated further and more deeply by others researches, but we think it still gives a good framing of this important problem for the usability of mobile camera-projector units, because mobile projector phones and mobile projection is still a young research field. The remainder of this section is structured as follows. The first subsection describes the different application classes of these interfaces. In this conceptual section we also discuss how the spatial layout of the camera relative to the mobile projection unit can affect the characteristics of applications for this new sort of hardware. Next, two example applications are presented: Map Torchlight and LittleProjectedPlanet. Map Torchlight is an application that combines high resolution paper maps with lightweight mobile projection to augment the paper map directly with additional personal and dynamic information. The mobile camera-projector unit is tracked over a paper map and precisely highlights points of interest, streets, and areas to give directions and other guidance for interacting with the map. With the LittleProjectedPlanet prototype [127] we explore the possibilities of camera projector phones with a mobile adaptation of the Playstation 3 (PS3) game LittleBigPlanet[1]. The camera projector unit is used to augment the hand drawings of a user with an overlay displaying physical interaction of virtual objects with the real world. Players can sketch a 2D world on a sheet of paper or use an existing physical configuration of objects and let the physics engine simulate physical procedures in this world to achieve game goals. In addition, we discuss the technical setups we used in both prototypes. Finally, we provide some concluding remarks and outline different scenarios where mobile camera projector units can be useful in the future.

[1] http://www.littlebigplanet.com/

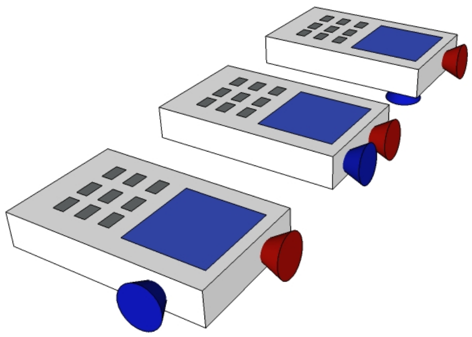

Figure 3.12: Different spatial layouts of integrated projector and camera in a mobile device. The camera is indicated with a red cone. The projector is indicated with a blue cone.

3.2.1 Different Spatial Configurations of Mobile Camera-Projector Units

Today‘s prototypes of mobile (camera) devices with integrated projectors are very limited in the functionality of the projector. Often the projector unit is just used to project media such as pictures or videos onto any surfaces. The projector is therefore mainly used to extend the real estate of the mobile device screen. To fully exploit the potential of mobile projection we discuss the impact of different spatial layouts of the camera relative to the projector unit on the interaction.

3.2.1.1 Spatial Layout of Camera & Projectors

When discussing the spatial arrangement of the camera to the projector we first want to define the terms camera field of view (FoV) and the term field of projection (FoP). The FoV of the camera is defined, as the area the camera is able to ‘see”. The FoP is the area the projector is able to project on. We distinguish between three different spatial layouts: First setups where the FoV and the FoP do not overlap are categorized as disjunct, because the projection goes to a completely other direction than the visual field of the camera. Setups were the the FoV of the camera and the FoP overlap are categorized in two different classes, which have the direction of the projector as well

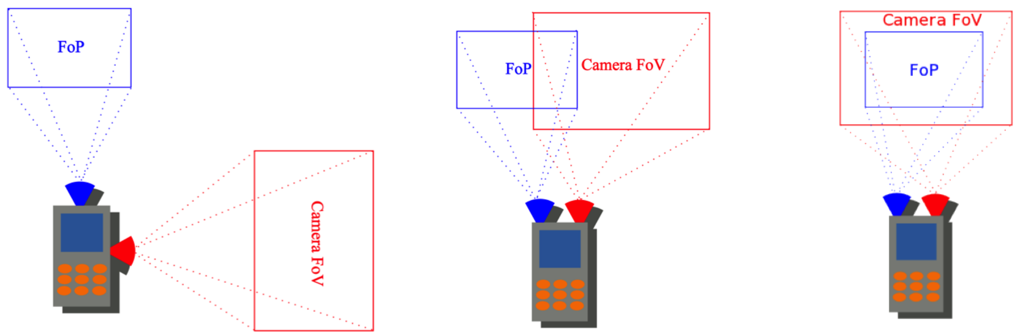

Figure 3.13: Disjunct Setup: The field of view of the camera (FoV) is not overlapping the field of projection (FoP). Partially intersecting Setup: The field of view of the camera (FoV) is partially overlapping the field of projection (FoP). Congruent Setup: The field of view of the camera (FoV) is completely overlapping the field of projection (FoP).

as the direction of the visual field of the camera in common: partially intersecting and congruent. They differ from the configuration of the lens of camera and projector and their distance to each other. If the visual field of the camera overlaps partially with the projected field, then it is categorized as an intersecting projection. In the third category, named congruent, the entire projected field is situated within the image produced by the camera (see 3.12). Due to different hardware specifications of cameras and projectors (different throwing angles, aperture, and others properties) the actual spatial setups could be very different. Today, due to the technical limitations, just disjunct setups exists. We think the partially intersecting and congruent provide a lot of more potential for new interactions as we illustrate in the following paragraphs.

3.2.1.2 Disjunct

In the case of disjunct projection, camera and projector are often attached to two different sides of the mobile device (see figure 3.12 and figure 3.13 or see the Epoq EGP-PP0 prototype). As a result, the visual field of the camera and the projected image are not overlapping. The alignment described is rather unsuitable for the augmentation of physical objects. The tracking systems would identify the objects located in the visual field of the camera, but the projection is directed towards a different angle so directly projecting onto these objects is not possible. Despite that, there are two ways of overcoming this problem. One possibility is determining one's own position in relation to that of an object by means of a spatial model of the environment and subsequently augment the object. Again this approach, however, requires the availability of a spatial model of the environment at all times. Furthermore, this procedure causes a considerable restriction of mobility. Another possibility is adapting objects or taking advantage of the structure of an object in order to augment it. For example, an optical marker, which can be identified and interpreted by the camera, could be attached to the first page of a book, resulting in the projection of additional information onto the open cover of the book. This would enable users to quickly and easily access reviews, summaries, and other services. A benefit of systems that use disjunct projection is that they allow for optical flow tracking, which is not (just with major hardware modification as described below) possible in the other types of projection, as projection within the image produced by the camera would then impair the tracking process. An infrared filter in front of the camera in combination can overcome this problem in combination a projector emitting no infrared light, but we think mobile cameras with build in infrared filters are still far from the current market. Optical flow tracking could be used to navigate websites, as Blaskó et al. [22] have described in their work. The movements of the camera/projector unit could be translated to instructions for the browser by the tracking system and navigation through a website which is projected onto a wall would be possible. Such navigation is very similar to the experience of a torchlight: it illuminates a part of an object at a time and the object stays static. This interaction metaphor is also used in Map Torchlight application (a partially intersecting setup) described below.

3.2.1.3 Partially Intersecting

In the case of partially intersecting projection the visual field of the camera and the projected field are situated on the same level partially overlapping each other as shown on figure 3.13. By knowing the angle of aperture of the camera and projector's lens, the size of the visual field of the camera and the projected field as well as its misalignment can be calculated. This kind of projection is the most suitable for the augmentation of visual objects. The fact that the projected field does not affect or minimally affects the image produced by the camera makes the stable use of visual trackers possible. However, this works only for the augmentation of bigger-sized objects. The field of smaller-sized objects is just too small for the augmentation and the tracking and projection as well as visual tracking would influence each other. The Map Torchlight application uses a partial spatial synchronous projection for the projection of additional POIs on a large paper map. An additional application, which assists someone in fixing e.g. the engine of a car, could be realized in a very similar way. By attaching visual markers to the engine compartment the possibility is given to determine the position of the mobile camera-projector unit relative to the engine, so that it can mark for example the screws which shall be removed in a particular step of a procedure which allows to perform a task. The advantages (or differences) of a partial overlapping projection compared to a congruent projection is that in the field of view of the camera as well as the field of projection of course are areas that are not effected by the camera respectively the projector. For example the non-overlapped area in the camera field of view can be used to allow gesture based interaction (as proposed by Baldauf et al. [6]) without interfering with the projection.

3.2.1.4 Congruent

A congruent setup is given when camera and projector are attached on the same side (or with appropriate hardware) of the mobile phone as it is the case with the partially intersecting projection with the exception that the entire projected image is shown in the camera field of view. (see figure 3.13) A disadvantage of these spatial configurations is that the projection could influence the processing of the camera image. However, the congruent projection enables the user to interact directly with the projection without any limitation. The application LittleProjectedPlanet, as described below, introduces in this spatial configuration, which works on the approach of the direct manipulation and which enables the user to operate the projection through the modification of physical objects. Another domain for a synchronous projection setup could be an OCR, which recognizes and marks spelling mistakes. School children would be able to control their homework by holding their mobile phones with the integrated projector upon their exercise books on which the projector could mark the mistakes and give e.g. additional information about them. However, the realization of such a system as an end product for costumers will take considerable research. The main problem is the robustness of the handwriting recognition process.

3.2.1.5 Other Classifications of Mobile Camera-Projector Applications

Beside the classification based on the spatial configuration of mobile camera-projector systems we could also classify the applications by their type. These are, for example, applications whose projection adapts to objects like it is the case with Map Torchlight (object adaptive applications) or applications on which the content of the projection is navigated by the recognition of an object or a marker and the projected area is independent (context-adaptive applications). The third category contains applications whose projection is mainly navigated by the camera but is independent from context and location (camera navigated applications), an example is WUW [145]. Besides this classification other issues should be taken into account when designing applications for mobile camera projector systems. The related work section provides an overview on the latest research done in this field, e.g. concerning multi-user settings, but some open issues are discussed in the next section. Not only the spatial configuration of the mobile device camera and the projector play a role when discussing the potential and limitations of mobile camera-projector units. Today, hardware issues still hinder the exploitation of the full impact of mobile camera-projector units. The effects of environmental light, energy problems (current projectors need a considerable amount of energy), and depth of focus, are problems that have to be solved by the hardware manufactures as well as further research. Other limitations, such as the limitations by nature of objects, project against object properties are still not discussed or investigated: "Am I allowed to project on a stranger passing by?". Many technical challenges still remain and have to be solved by the hardware engineers. They also have to create and enable different spatial layouts of mobile camera-projector units. Effects of hand shaking and tremor can be overcome utilizing accelerometers. Moreover, camera-tracking methods have to be improved. All these factors currently have a big impact on the user experience and have to be taken into account when designing applications for mobile-camera projector units.

3.2.2 Examples for Concepts & Application Classes

In the following section we present two fully implemented prototypes to illustrate the concept presented earlier.

Figure 3.14: The Map Torchlight prototype in use (Please note, that the map background is artificial blurred).

3.2.2.1 Map Torchlight

The advantages of paper-based maps have been utilized in the field of mobile AR in the last few years as hilighted earlier. From an HCI perspective, the main challenge of magic lens interfaces is that users have to switch their attention between the magic lens and the information in the background. With the Map Torchlight application we attempt to overcome this problem by augmenting the paper map directly with additional information. The Map Torchlight is an example for a partially synchronous projection and is tracked over a paper map and can precisely highlight points of interest, streets, and areas to give directions or other guidance for interacting with the map (see figure 3.14).

Basic Interaction Concept The general advances of a mobile projection system also show up in our Map Torchlight system: The projection area is larger and the mobile projection can overcome the switching cost of magic lens interfaces. The basic interaction pattern is similar to magic lens interfaces. Sweeping the camera projector unit over the map, the projector will, for instance, highlight different POIs on the map. Because the projection is significantly larger than the device display (around 4 times in our setup) more dynamic information can be directly presented on the map (as can be seen in figure 3.14). It also provides a higher resolution compared to a standard mobile device display, if the projector is controlled independently from the device display. As shown in figure 5, larger objects can be highlighted compared to a traditional magic lens interfaces. The projector can also be used to collaboratively interact with a map by using the map as a shared screen. For instance, one user can tell another a route through the city by moving a projected crosshair over the map. The waypoints could then be stored in a Keyhole Markup Language[1] file and transferred via Bluetooth to the second user's mobile device. Again, in all of these examples, there are no switching costs for the users. A downside of projection is that the real-world view cannot completely be blocked out, as is possible with (video see-through) magic lens interfaces.

Implementation The Map Torchlight is fully implemented for Nokia mobile camera phones (S60 3rd edition). The actual prototype is a Nokia N95 mobile phone with an AIPTEK V10 Mobile Projector (640x480 pixel) attached to the phone using a standard AV cable. The whole setup weighs about 360 grams. Due to technical limitations the mobile phone screen can only be mirrored and not be extended on the projector. Due to this issue, the projector always shows the mobile screen content, even if detailed information is presented on the mobile device screen. The focus and projection size needs to be calibrated manually, because the focus of the projector can only be adjusted manually. The tracking algorithm processes about 12 frames per second.

3.2.2.2 LittleProjectedPlanet

With the LittleProjectedPlanet prototype we explore the possibilities of camera projector phones with a mobile adaption of the Playstation 3(PS3) game LittleBigPlanet. The camera projector unit is used to augment the hand drawings of a user with an overlay displaying physical interaction of virtual objects with the real world. Therefore a spatial synchronous projection setup is needed. Players can sketch a 2D world on a sheet of paper or use an existing physical configuration of objects and let the physics engine simulate physical procedures in this world to achieve game goals. We propose a mobile game combining hand drawn sketches of a user in combination with objects following a physics engine to achieve game goals. Enriching sketching in combination

[1]KML is a is an XML-based language schema for expressing geographic annotation and visualization.

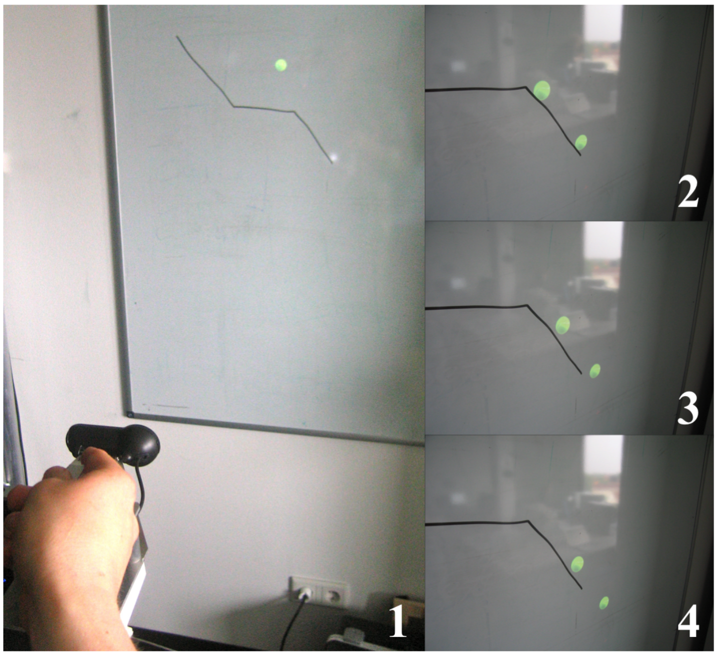

Figure 3.15: The LittleProjectedPlanet prototype in use.

with physical simulation was presented by Davis et al. [2; 47]. The ASSIST system, was a sketch understanding system that allows e.g. an engineer to sketch a mechanical system as she would on paper, and then allows her to interact with the design as a mechanical system, for example by seeing a simulation of her drawing. Interestingly the ASSIST system was bought by the creators of the game LittleBigPlanet and parts of ASSIST were integrated into the game play. In contrast to the ASSIST system we present a game that is designed for mobile projector phones combining real world objects and projected ones utilizing a physics engine. We think that this kind of mobile projection camera unit can been utilized to improve the learning and collaboration in small groups of pupils (cause of the mobile setup of our prototype) in contrast to more teacher-centered teaching e.g. one interactive white board (as shown by Davis et al. [2; 47]).

Game Concept The slogan of the popular game LittleBigPlanet is “play with everything” and that can be taken literally. The player controls a little character that can run, jump and manipulate objects in several ways. A large diversity of pre-build objects is in the game to interact with, and each modification on such an item let them act in a manner physically similar to those they represent. The goal of each level is to bring the character from a starting point to the finish. Therefore it has to overcome several barriers by triggering physical actions. But the main fascination and potential of the game is the feasibility to customize and create levels. Creating new objects is done by starting with a number of basic shapes, such as circles, stars and squares, modify them and then place them in the level. Having done so, the user can decide on how these objects should be connected mechanically. We took this designing approach as an entry point for a mobile augmented reality game using a mobile camera projector unit. It allows the user to design a 2D world in reality, which is then detected by a camera. Out of this detection a physical model is being calculated. Into this model the user can place several virtual objects representing items like tennis balls or bowling balls. These virtual objects then get projected into the real world by the mobile projector. When starting the physic engine, the application simulates the interaction of the virtual and the real world objects and projects the results of the virtual objects onto the real world surface. Just like in LittleBigPlanet our application offers the user different ways of playing: One is like the level designer in LittleBigPlanet; the user can freely manipulate the 2D World within the projected area and place virtual objects in it. Similar to children building tracks for marbles in a sandpit, the player can specify a route and then let the virtual marbles run along it. A different gaming mode is a level based modus, but instead of steering a character as in LittleBigPlanet, the user designs the world. As a goal the user has to steer a virtual object e.g. a tennis ball from its starting point to a given finish. The game concept uses a direct manipulation approach. Enabling the player to modify the world at runtime let the real world objects become the users tangible interface. But not only the objects are used for the interface, by changing the orientation and position of the projector the user can also modify the physical procedures (e.g. gravity by turning the mobile camera projector unit).

Basic Interaction Concept For designing a 2D world the players can use several methods. Basically they have to generate enough contrast that can be detected by using a standard edge recognition algorithm (utilizing the Sobel operator). Sketching on a piece of paper or a white board for example can do this, but simply every corner or edge of a real world object could generate a useful representation in the physics engine. So there is no need for an extra sketching device or other for example IR based input methods. Just requiring the camera projector unit itself the game is playable nearly anywhere with nearly everything and it is easy to set up. Figure 3.15 shows how a user is projecting virtual marbles on a track she sketched on a whiteboard. An important problem to allow a smooth and seamless interaction for the user is that the gravity in the projection is aligned with the real worlds gravity. For that a Nintendo Wii is attached under the camera-projection unit. Also gravity can be utilized in the game to control some action. A user can take control of the gravity by changing the orientation of the projector. Doing this the user can let virtual objects fly through the levels.

Implementation Due to the unavailability of sophisticated projector phones (with an optimal alignment of camera and built-in projector and e.g. a GPU that is able to process the physics simulation) we used, in contrast to the Map Torchlight application, a Dell M109S, a mobile projector with a maximum resolution of 800 by 600 pixels and a weight of 360g, in combination with a Logitech QuickCam 9000 Pro. All together our prototype weighs around 500g and is therefore okay to handle (e.g. compared to the prototype used in Map Torchlight (see above) our prototype is 240g heavier, but the projector has 50 lumen instead of just 10 and also has a higher resolution). We think this prototype provides a good trade-off between mobility and sophisticated projection quality. In contrast to the few mobile devices with built in projectors, our projector and camera are mounted in such a way that the camera field of view fits the projected area (spatial synchronous projection). But because of the different focal lengths of camera and projector in this setup the camera image is always wider than the projected image. Therefore the camera is installed in front of the projector as can be seen in figure 3.15. For controlling the application and to determine the orientation (to set the gravity) a Nintendo Wii remote is attached to the camera projector unit. Most actual Smart Phones are already equipped with an accelerometer or an electronic compass, so the functionality of the Wii remote can easily be covered using a mobile phone. The application is fully implemented in Java using the QuickTime API to obtain a camera image. As a physics engine Phys2D[1], an open source, Java based engine, is used. The communication with the Wii remote is handled by WiiRemoteJ[2]. Connected to a standard laptop or PC the camera projector unit has a refresh rate

[1] `http://www.cokeandcode.com/phys2d/`

[2] `http://www.world-of-cha0s.hostrocket.com/WiiRemoteJ/`

of approximately 15fps when running the application. The area of the camera image containing the projected image is processed via an edge recognition algorithm. Every pixel of a detected edge gets a representation as a fixed block in the physics engine. That gives the user total freedom in designing the world. Such a physic world update is done every 300ms but it can be stopped by the users, for example for editing the sketch. Adapting the gravity of the physical model to the actual orientation of the camera projector unit is done through calculating the roll (this denotes the angular deviation along the longest axis of the Wii remote) of the Wii remote.

3.2.3 Discussion & Future Work

We have presented different application classes of interfaces utilizing a mobile camera-projector unit. The interfaces all focus on the augmentation of real word objects in the environment. We showed how the different spatial setups of camera and projector units effect the possible applications and the physical interaction space. This classification can help to structure the design space of mobile projection applications. Of course many open issues still remain. As discussed earlier not only the spatial configuration of the mobile device camera and the projector play a role when discussing the potential and limitations of mobile camera-projector units. Today, hardware issues still hinder the exploitation of the full impact of mobile camera-projector units. The effects of environmental light, energy problems (current projectors need a considerable amount of energy), and depth of focus, are problems that have to be solved by the hardware manufactures as well as further research. Our research tries to make a contribution into the direction that by assuming we will have better hardware of mobile-camera projector units, we will have more powerful applications such as Map Torchlight and LittleProjectedPlanet. Both implementations show how researchers can overcome the current hardware problems and investigate the area of mobile camera-projector systems more deeply. With our categorization using different classes based on the spatial configuration we want to establish a first initial framework for mobile camera-projection systems. With our example applications we highlighted the potential of mobile camera-projector units. We think that they have a big potential to enrich the usability of mobile devices. They enable larger presentation sizes and are well suited for multi-user settings. In future work we intend to evaluate how mobile projection interfaces are superior to classical AR interfaces such as HMDs and magic lens interfaces.

Figure 3.16: The PhotoMap application in use. The location of the user is highlighted (with a red marker) on a photo of a public map.

3.3 PhotoMap

In many mid- to large-sized cities public maps are ubiquitous. One can also find a great number of maps in parks or near hiking trails. Public maps help to facilitate orientation and provide special information to not only tourists but also to locals who just want to look up an unfamiliar place while on the go. These maps offer many advantages compared to mobile maps from services like Google Maps Mobile or Nokia Maps. They often show local landmarks and sights that are not shown on standard digital maps. Often these 'YOU ARE HERE' (YAH) maps are adapted to a special use case, e.g. a zoo map or a hiking map of a certain area. Being designed for a fashioned purpose these maps are often aesthetically well designed and their usage is therefore more pleasant. In this section we present a novel technique and application called PHOTOMAP that uses images of 'YOU ARE HERE' maps taken with a GPS-enhanced mobile camera phone as background maps for on-the-fly navigation tasks. We discuss different implementations of the main challenge, namely helping the user to properly georeference the taken image with sufficient accuracy to support pedestrian navigation tasks. We present a study that discusses the suitability of various public maps for this task and we evaluate if these georeferenced photos can be used for navigation on GPS-enabled devices. From a slightly different perspective the approach described here can be considered, in part, as a form of Note taking [14] whereby a mobile user walks away with information that they believe will be of later utility In contrast to other map alignment software such as Microsoft's Map Cruncher [59] or HPs Map Aligner, PhotoMap is designed to support spontaneous interaction with 'You are here' maps in the environment without the need of downloading map data on the mobile device beforehand. Following the ideas of Weiser [239], of *spontaneous interaction*, the PhotoMap application can support users, because one will often see such specialised maps on public signs situated where they may be required (in a just-in-time sense) on the go. Microsofts Map Cruncher [59] is a Microsoft Research project which uses the Virtual Earth API to import supplemental maps into Virtual Earth as a desktop tool. The Map Aligner from HP is a map aligned tool for a PDA. It requires that the user downloads map material beforehand on the device to georeference it in the field.

According to a recent BBC article[1], Nokia predicts that in the coming year it will sell 35 million GPS-enabled phones and this reflects how personal navigation has become the latest feature to be assimilated into the mobile phone. Similarly, the same article states that Symbian's operating system was shipped on 188 million phones last year and a third of those came with GPS. In line with these developments, a range of "free to download" mobile map applications are currently available. The most popular being Google Maps Mobile or Nokia Maps. Indeed a compelling scenario for the use of GPS (Global Positioning System) enabled phones is support for pedestrian navigation, e.g. enabling a user to glance down at the screen of her mobile phone in order to be reassured that she is indeed located where she thinks she is (see figure 3.16). While service based approaches to support such navigation tasks are becoming increasingly available – whereby users download relevant maps of their current area onto their GPS enabled phone the approach is often far from ideal [189].

Typically such maps are highly generalised and may not match the user's current activity and needs. For example, rather than requiring a standard map on a mobile device, e.g. using Google Maps Mobile or Nokia Maps, of the area, the user may simply require a map of a university campus showing all departments or a map showing footpaths around the area in which she is currently trekking. Indeed, one will often see

[1] `http://news.bbc.co.uk/2/hi/technology/7250465.stm`

Figure 3.17: Comparison of a locale map of the campus of Lancaster, UK (middle, right) and a map by Google Maps (right). © 2009 Google - Map data © 2009 Tele Atlas.

such specialised maps on public signs situated where they may be required (in a just–in–time sense). It is interesting to consider how one might enable users to walk up to such situated signs and use their mobile phone to 'take away' the map presented in order to use it to assist their on-going navigation activity. Figure 3.17 shows a comparison of a map by Google Maps (right) and a local map of the campus of Lancaster, UK (left, middle). The local map (left) shows local landmarks omitted from Google Maps and other web map services, such as departments and info points and detailed information such as footpaths supporting navigation on the campus.

In this section, we expand upon the concept introduced in [39] of enabling users to walk up to such situated signs and use their GPS-enabled mobile phones to 'take away' the map. This happens by taking a photo of it, in order to use it to assist their on–going navigation activity (see figure 3.16). The contributions presented in this paper are the analysis of the design and characteristics of public maps we collected, a description of different approaches and implementations for georeferencing photos of maps on the fly, and proof-of-concept evaluation, providing validation of the feasibility of the concept using images of public maps in the context of pedestrian navigation tasks.

The remainder of this section is structured as follows. Section 3.3.1 describes the scenario for the PhotoMap application and in Section 3.3.2 we analyze a corpus of 93 'You are here' maps we have taken. In Section 3.3.3 we show three different methods for enabling users to georeference the maps and we describe the implementation and limitations of our georeferencing approaches. Next, Section 3.3.4 presents a proof-of-concept, exploring the possibility to use photos of public city maps for mobile navigation purposes. Finally, Section 3.3.5 provides some concluding remarks.

Figure 3.18: The "Smart Alignment" approach: User takes a photo of a public map and connects the 'You are here' dot position to his actual location (the left images, the 'You are here' is marked with a yellow circle). After that the photo is overlayed on a digital map and the user can adjust it by scaling it (right). This can be done by using the standard 4-direction joystick (as indicated with the yellow arrows). After completion of this appraoch the user can use the the photo for navigation - the user's position is indicated with a blue marker (marked with a yellow circle right).

3.3.1 Scenario

The following scenario illustrates the concepts more fully: Maximilian is walking through the city of Münster, Germany. He approaches the botanic garden behind the castle of the city. Standard maps of this area just show a grey (not even a green) area (see Figure 3.20; 2^{nd} from left), but fortunately there is a paper 'You are here' map at the entrance of the garden containing all details and POIs of the garden. Maximilian takes a picture of the map of the botanic garden with his GPS-enabled phone and performs some additional actions (three different methods are described in detail in Section 3.3.3) to do the georeferencing in an easy and appealing way. For that we utilize the current GPS positions where Maximilian took the photo in order to establish the map's scale and extent (and northing). This can be verified by the current lat/lon coordinates coming from the phone's GPS unit. The map remains as an image on Maximilian's phone (see figure 3.16), because it contains height information and additional useful local information, but is now georeferenced and can be used for navigation. For instance, it might provide information on the different flowers or vegetation that is not available in standard map applications. Note that the 'You are here' dot shown in the screenshots reflects that shown in the physical map and serves as a useful reassurance to Maximilian regarding his current location. For this approach we *only* have to tackle one main challenge. Namely that the users have to georeference (also referred to as rectifying) the map photo on-the-fly on their mobile device. This is described in detail in section 3.3.3.

Figure 3.19: The usual interaction scheme of a PhotoMap user. A user comes to a public paper map, takes a photo and uses this photo for navigation.

3.3.2 PhotoMap Corpus

In order to explore the variety and characteristics of 'You are here' maps we collected 93 maps in 21 cities (in 8 countries in central Europe and North America). The cameras used for this process comprised the built in cameras of the Nokia 5500 (with 2 megapixel), the Nokia N95 (with 5 megapixel) and a Nokia E71 (with 3.2 megapixel camera). Figure 3.20 gives an impression of the variety and density of these maps and their superiority against standard map services. 36% of the maps we collected were city maps showing POIs and important sights in the inner city areas. 31% of the maps were showing local areas at a large-scale, e.g. campus maps, surrounding maps at train stations or bus stops, and maps showing local shops, or sports areas. 16% of the maps in our corpus were hiking maps showing different hikes through national parks or mountain areas. 14% of the maps showed gardens, parks and zoos. 2% of the 'You are here' maps showed larger regions (e.g. 1% were historic maps of a local area. It was interesting that about 96% of the maps were spatially correct, meaning they were designed using underlying map data and had scales ranging from 1:1500 to 1:2500. 81% had the correct northing and 62% of the maps had a 'You are here' dot. As can

Figure 3.20: Different local 'You are here' maps (top) compared against Google Maps (bottom). From left to right: Centre of education of the finance school in Gievenbeck, Germany; Botanic garden in Münster, Germay; Zoo of Münster, Germay; Central Station Amsterdam, Netherlands; Hofgarden in Innsbruck, Austria. All maps at the bottom © 2009 Google.

been seen in figure 3.21 the 'You are here' dot was mainly positioned in the middle of the map. We analyzed the (x, y) positions of all 'You are here' dots in all maps with a 'You are here' dot and normalized the position to a scale from 0 to 100 in the x- and y-direction. Also figure 3.21 shows the area of distribution of the 'You are here' dots. As noticed above most of the 'You are here' dots were positioned in the middle. Another "hot zone" was the lower middle edge. Having the dot at that position has the advantage that the user is in front of the area she can explore. Most of the other 'You are here' dots were positioned on the upper left quarter of the map, probably because of the reading direction in central Europe and North America.

3.3.3 Georeferencing approaches

The georeferencing step needs to be supported by the users themselves. For that we developed two approaches explaining a trade off between user interaction and system complexity, namely the "Two Point Referencing" and the "Smart Alignment" approach. In the first version, the map is rectified in two steps and the user has to move physically between these steps before being able to use the map. Our second implementation allows the user to adjust his photo of a YAH-Map to an already referenced map on her GPS trace and then use it as a foundation for navigation.

Generally the projection of the map and the image distortion (caused by the tilting angle between map and mobile device) should not be too extreme. An optimal photo

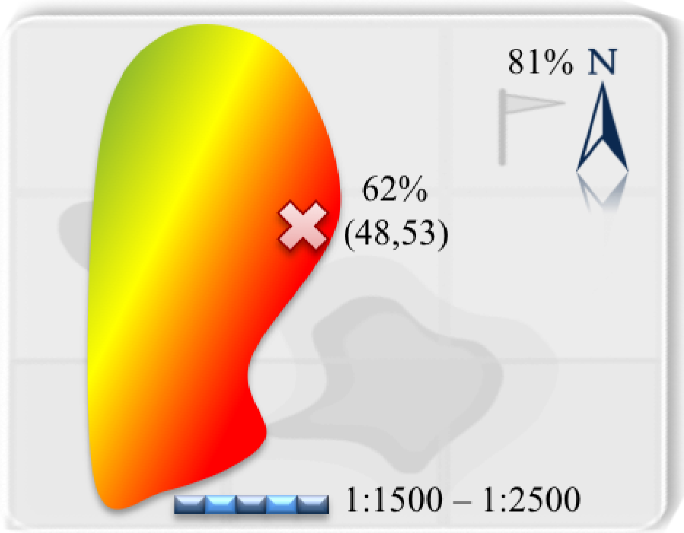

Figure 3.21: 'You are here' maps properties. From our collection of 93 maps 81% had the correct northing; 62% had a 'You are here' dot mostly positioned in the middle. The colored area indicates the most frequently used positions for 'You are here' dots on a map. The color indicates the frequency of the dot placement from green (low) to red (high). In average the maps had a scale from 1:1500 to 1:2500.

would be taken by a mobile device held parallel to the map (see figure 3.19). Reference points (assigning real world coordinates to the image pixel) need to be given by the user and need to be combined with the current GPS-coordinates provided from the GPS-module of the mobile phone. Currently we have implemented the following approaches for our PhotoMap prototype.

3.3.3.1 "Two Point Referencing"

Directly after the user has taken the image of the map she should either indicate where she is on the map or mark the 'You are here' dot to determine her actual position. So, she positions a crosshair over the dot using the standard 4-direction joystick of the mobile device. Then she has to wait for a GPS-Signal, at which point her actual position is connected to the x-, y-coordinates of the 'You are here' dot in the image. After she has moved a sufficient distance the phone should request another indication of position from her. In this second step she must repeat the process (see figure 3.22) again. After that PhotoMap rectifies the map and allows her to look up her location

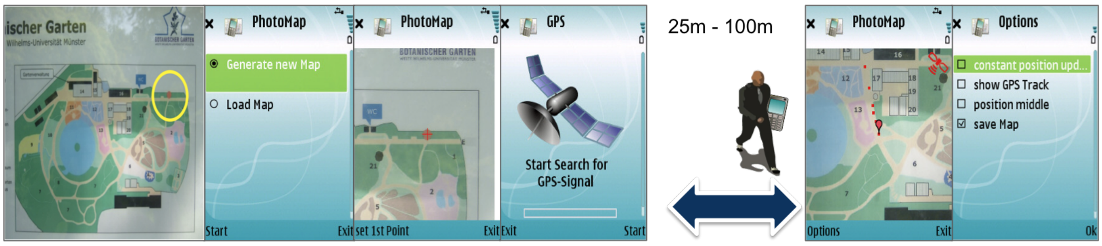

Figure 3.22: The "Two Point Referencing" approach: The user takes a photo of a public map (left) (the 'You are here' dot is marked with a yellow circle). Marks the dot with a cross-hair and connects it with current GPS position. The user walks away from the map (about 25m to 100m) and repeats the process after. After that he is able to use the photo of the map for navigation tasks. The screenshots (right) show the working application and the options menu.

or create a GPS trace or record a GPS-Track (see figure 3.22 right). She is able to pan around the map and can either obtain her actual position and/or her GPS trace. In the actual implementation the screen of the mobile device is about four times larger than the photo of the map. This is a trade off between photo resolution and map scale. A world file[1] is stored for the taken photo so it can be read by a GIS. The "Two Point Referencing" method is implemented in Java ME, so it is compatible with most mobile devices. For testing purposes we mainly use the Nokia N95 and E71 due to their build-in GPS and their reasonable quality cameras.

3.3.3.2 "Smart Alignment"

To significantly reduce user interventions we developed a second georeferenzing method called "smart alignment". After taking a snapshot of a YAH-Map the user has to indicate again where she is on the map or mark the 'You are here' dot to determine the actual user's position. After that the photo of the map is displayed semi-transparently over the map data. The 'You are here' dot is anchored to one corresponding position on the map. From our map corpus we utilise the knowledge that the scale is normally about 1 : 1500 to 1 : 2500 and that about 81% of the maps have the correct northing. Hence the photo is roughly adjusted depending on the scale. The user has to perform the fine adjustment of the photo on the map. This can be done by scaling the image in

[1] A world file is a plain text computer data file used by geographic information systems to coordinate raster map images introduced by ESRI, one of the leading GIS companies. These world files are six-line ASCII files with decimal numbers on each line [247].

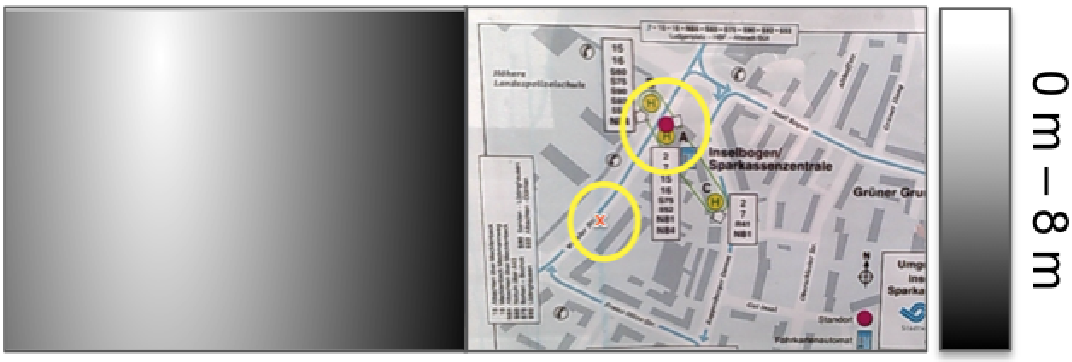

Figure 3.23: Mean error introduced by the "Two Point Referencing" approach. The error range (from 0m - 8m) was mapped on 255 grey-scale value (white pixel 0m error; black pixel 8m). The mean error was about 5m.

every direction using the touch screen, rotating (if necessary in 12% of all cases) can be done by pressing additional keys on the keyboard. Because of its seamless integration of Google Maps Mobile we chose the Android platform using the developer version of Google G1 for our implementation. After adjustment, the photo should be rectified to high accuracy after which it can be used for precise navigation. Again the photo is stored with the corresponding world file for further use in GIS or lightweight systems like virtual globes.

A slightly modified version of the "smart alignment" method utilizes the GPS trace of the user. Instead of overlaying the photo of the map over a digital map, the GPS trace is displayed. If the device is equipped with a touch screen, such as the Google One, the user can directly draw on the PhotoMap the path on which she approaches the map, otherwise she needs to indicate way-points with the help of a crosshair. Of course, here it is assumed that the user's trace overlaps with the spatial extend of the photo of the map. As we can conclude from our map corpus this is the case in nearly half of the cases (about 48%). Problems will occur if the 'You are here' dot is positioned at an edge of the map.

3.3.3.3 Comparison of both Approaches

Both georefercing approaches have several advantages and disadvantages. A general major problem of this method is the inaccuracy of the GPS-Receiver, which potentially leads to a wrong or imprecise rectification. As reported from the literature [97] a

Figure 3.24: The photo of the hospital area that was used for the navigation for all users (left with a dashed red border). Example of photos users took using the PhotoMap application. Out of 20 photos the users took just one photo was not usable for navigation (right).

standard GPS can achieve an absolute accuracy of 5m to 10m. Of course, the GPS signal can be improved by using Assisted GPS (A-GPS) [248] with the disadvantages of requiring a data connection to a server.

In addition the user can introduce additional error sources: they can introduce an offset between the 'You are here' dot and their position indicated by the crosshair (see figure 3.3.4.1). In the first version they can also indicate a second wrong position. In the second "smart alignment" version they can choose a wrong scale, but the error sources in the first approach are significantly more critical.

To estimate the error introduced by the user we let 5 users select a 'You are here' dot and a second point on a bus stop map near our former institute. The map can be seen in figure 3.3.4.1 (middle). Both locations are marked with a yellow circle. Figure 3.3.4.1 (left) also shows the resulting error maps. The rectified maps were overlayed on cadastre data and the pixelwise offset was calculated for each pixel. The error range (from 0m - 8m) was mapped onto 255 grey-scale value (white pixel 0m error; black pixel 8m). The mean error was about 5m. So it doubles the inaccuracy of the GPS in the worst-case scenario. Other disadvantages of the first version are:

- Only northed maps are supported (81% of the maps we collected had the correct northing).
- It may take up to the second point until the georeferencing process produces results. Until then, users have no navigational support.
- At remote locations or in urban canyons the user needs to indicate her position on the PhotoMap at least three times at sufficiently remote locations to obtain

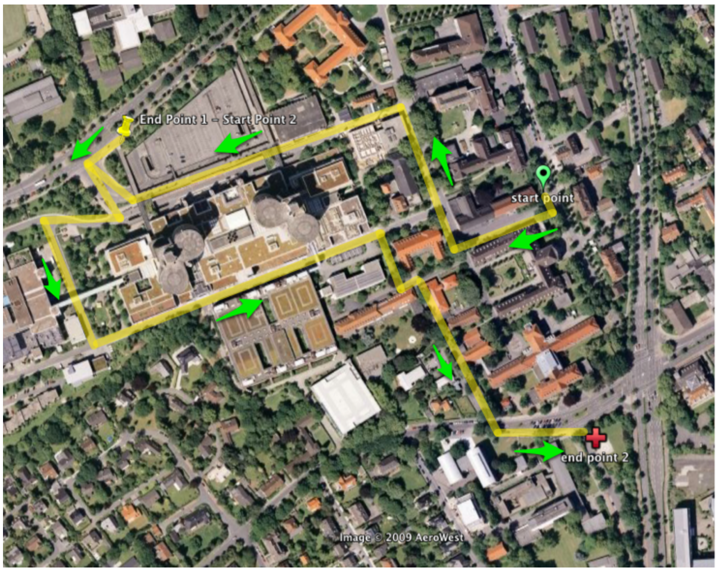

Figure 3.25: Map of the start and end Points for the navigation task. First the user had to navigate from the green marker (right) side to the yellow pushpin (0.8 km) (middle) and down back to the red cross shaped marker (1.3 km). © 2009 Google and Areal Image © 2009 AeroWest.

good results for the two-point version. Of course the user can improve the results by adding more than the required reference points.

To overcome these limitations, we developed the second georeferencing version. We see many more possibilities how the georeferencing could be improved with minimalizing user interaction, for example, the user could easily limit the degrees of freedom by indicating the scale or the northing of the map (using our knowledge gained from the map corpus).

The big advantage here is that georeferencing starts immediately after the interaction and it is expected that the quality of the georeferencing process increases because the user just has to adjust the scale in most cases (in some cases also the rotation). But the downside of this technique is the additional requirement for a digital map and the costs that are caused by either downloading or storing it. Using the GPS trace for this purposes can overcome these problems (as present in the modified version of the "smart alignment" method), but of course the GPS trace is again reintroducing the error from the GPS device. The optimal solution of course would be to modify one of the presented georeferencing approaches to reduce the user interaction to a minimum.

3.3.4 Proof-of-Concept & Evaluation

Pre-tests revealed that the usage of a standard map client, such as Google Maps Mobile, did not support the users. Users had to memorize the paper map and the mobile map client was of little use. The PhotoMap application outperformed the mobile map client. Therefore we decided not to compare a mobile map client against our PhotoMap application.

Instead for the initial field evaluation, our goal was to validate the feasibility of the concept using images of public maps for pedestrian navigation tasks. Ease of navigation and task performance with mobile maps is influenced by map alignment to the orientation of the user [252]. To eliminate these side effects we chose a north oriented map that centred the user in the middle of the map (photo). We were interested how a photo taken by a user could be used for navigation tasks and exploring any issues raised by inaccuracies caused by the different ways of georefercing the photos on-the-fly.

3.3.4.1 Participants & Apparatus

The study took place in the hospital area in the city of Münster, Germany with 10 participants, 5 males and 5 females with an average age of 24.3. The participants where undergraduates and graduates from the local university. The study was conducted over a period of one week in early January 2009. All of the user had used a mobile camera device before. 45% of the users used a GPS navigation system (e.g. a car navigation system or an ordinary GPS for pedestrian navigation of geocaching activities). The test was performed on a Nokia E71 Symbian GPS enabled camera phone with a 3.2 megapixel camera with autofocus and flash running a J2ME PhotoMap version modified for the user test. To eliminate side effects of inaccuracy caused by the different ways of georefercing the photos on-the-fly we did the following: The photo the user took was stored on the mobile device and replaced with

(1) a correct georeferenced photo of the map of that area and

(2) a photo with an inaccuracy normally introduced through our application (see section 3.3.3.3 and section 3.3.4.1).

The test consisted of two navigation tasks. All participants performed the test with the precise photo of the map (1) and with the inaccurate georeferenced version of the photo (2) (we introduced an inaccuracy of 5m on average (see figure). Half of the users used the map (1) for the first navigation part and half of the user the map (2). The order was reversed for the second navigation task. As participants walked the route the experimenter shadowed them. Their route and the duration were automatically logged by the system. Also the photo the user actually took were stored on the mobile device (see figure 3.3.4.1). If the user took a wrong turn, the experimenter did not correct the user.

After the actual test users were asked to rate the map navigation techniques by filling out a modified version of the "user interface evaluation questionnaire" of ISO 9241-9 [98] with only a single Fatigue category. The ISO questionnaire is a seven-point rating evaluation. Higher scores denote a better rating. The total time each participant took for the whole study was about 50 minutes.

All users had to navigate by foot to two POIs on the hospital area. The shortest path for the navigation was about 2.1 km long and consisted of a route with 15 turning points. The participants were introduced to the PhotoMap system at a public paper map not showing the hospital area. After that the experimenter walked with the user to the starting point and showed the first POI on a public paper map. The user took a photo of that map and started to navigate to the first POI (about 0.8 km and 6 turning points, in average every 133m a turning point). The main goal was to investigate the general suitability of photos of maps for navigation and not to compare the different georeferencing approaches or investigate the effect of the inaccuracy introduced by the users in the georeferencing process. Arriving at the first POI the same procedure was repeated with the same map of the area (obvisouly with antother 'You are here' dot position) and the user had to navigate to the second POI (about 1.3 km and 9 turning points, in average every 144m a turning point). The second route was longer but had the same characteristic than the first one with the same ration of route length/ turning points. After completion of the test the users had to fill out the "user interface evaluation questionnaire" in the field.

3.3.4.2 Results

All participants were able to complete all navigation tasks. The users took 20 photos of the same 'You are here' map in the hospital area. Figure 3.24 show the original map (left) and four of the 20 maps the users made. It is interesting to observe, that

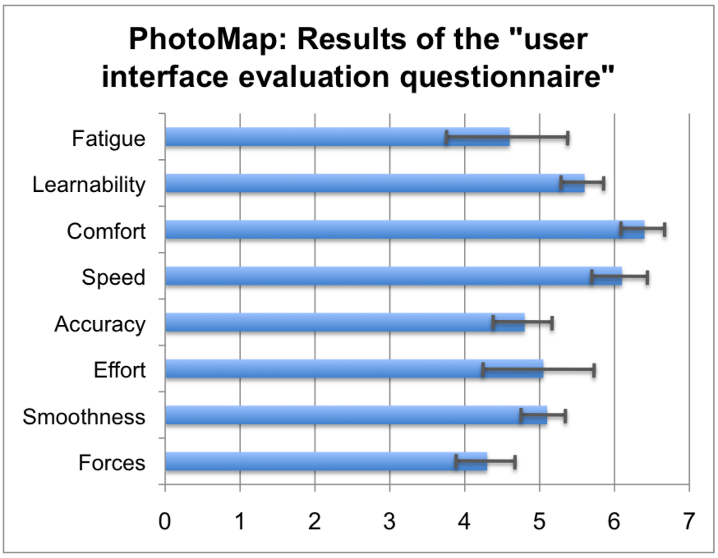

Figure 3.26: Results of the "user interface evaluation questionnaire". Users rate the factors comfort, speed and learnability best.

the users nearly took a photo of the same area. Just one photo was unusable for the navigation task (see figure 3.24 right). None of the users noticed that the photo they took of the map was replaced with a pre-georeferenced photo. The main performance measures taken were trial time and error rate. Trial time is the time from the start of a navigation task until the destination was reached. Overall we collected 20 trails (10 user $\times$ 2 navigation tasks). Both trial times and errors were derived from the recorded GPS traces. One trial of one user was removed for calculating the error rate and trial time, because this user had used a building entrance to shorten his path to the destination. The building entrances were marked on the map and this user took advantage of these shortcuts.

The error rate indicates the number of wrong turns taken by a user. The overall average trial time was 1745 sec. (95% confidence interval: 1568 - 1922 sec.) and the overall average error rate was 2.1% (95% confidence interval: 1.7 -2.5%). There was no significant difference in trial time and error rate comparing the map condition (1) and (2) (trial time $F(1,19) = 56.42$, $p \geq 0.05$; error rate: $F(1,19) = 54.11$, $p \geq 0.05$). Both differences are within the limits of the 95% confidence interval and thus not significant at the 5% level. So the map condition (2) with the inaccuracy of up to 5m did not affect user performance.

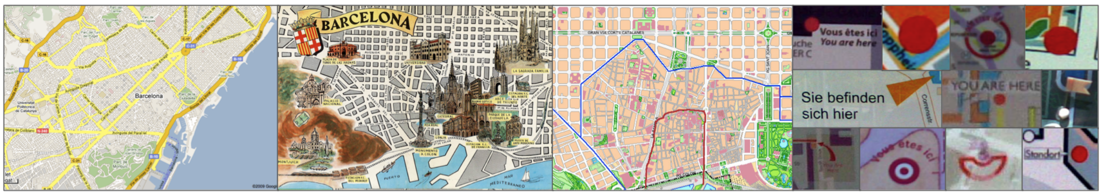

Figure 3.27: 'You are here' maps of Barcelona, Spain slightly rotated compared against Google Maps (left), so that the coastline was at the bottom of the map and the streets were lined up as a more or less regular grid. © 2009 Google. A collection of 'You are here' dots (right).

The results of the "user interface evaluation questionnaire" were very promising. The users gave high grades for the factors: comfort, speed and learnability. Furthermore, some users pointed out that the system was very easy to use compared to other navigation systems. "Just a couple of button presses are needed and I have the right map on my mobile - I hate searching the right map portion on a digital map". " It is great that you can even use it wearing gloves, because you mainly have to press the middle button." Users also pointed out that the barrier to entry was significantly lowered compared to other mobile maps services. The inaccuracy in the (2) condition was not noticed by any user and subsumes under general GPS inaccuracy. This is also reflected in the equal error rate of both conditions.

3.3.5 Discussion & Future work

In this section, we have presented our initial explorations into the technical feasibility and associated usability implications of allowing GPS-enabled mobile phones to support the capture, georeferencing and subsequent display of traditional 'You are here' map signage. In particular we have made the following novel contributions:

1. **Analysis of 'You are here' photo corpus:** we have investigated 93 maps in 21 cities, taken with different cameras of off-the-shelf mobile phones. By doing so we could infer that most 'You are here' maps can indeed be in most cases easily georeferenced by non-expert users.

2. **On-the-fly georeferencing methods:** we have presented two on-the-fly georeferencing methods that can be carried out on the mobile device, namely the "Two Point Referencing" and the "Smart Alignment" method. Both methods

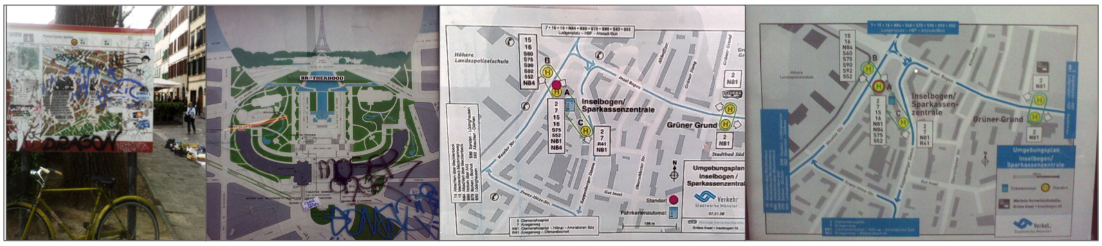

Figure 3.28: Example of vandalism to a 'You are here' type map in Florence, Italy and in the city of Paris, France (left). Dynamic character of these public maps shown with a map at a bus stop in the city of Münster, Germany photographed in late 2008 and early 2009.

require only little intervention from the user. While the first method just needs two clicks and a physical displacement of the user, the second method can be performed on the spot without the need of additional movement by the user.

3. **Conceptual evaluation**: we have presented results of a first user trial, where we collected evidence of the general applicability of 'You are here' map photos to pedestrian navigation tasks, given the expected mobile georeferencing error.

4. **PhotoMap application**: we have presented a first technical implementation of the PhotoMap concept. The "Two Point Referencing" approach was implemented in J2ME [221]. The "Smart Alignment" was developed for the Android platform.

Especially the user studies with the working prototype provided encouraging feedback (as well as revealed interesting insights into the advantages and drawbacks of the different geo-referencing methods which we employed),which encourages us to further explore this novel approach to exploit traditional map signage with mobile devices. Since this we just presented the first steps to investigate this interesting combination of traditional and digital map usage, each of the above mentioned contributions lead to a couple of interesting research questions:

The first contribution directly leads to the question how to archive and collaboratively manage and collect YAH maps. One specific avenue, which we intend to explore, is the potential utility of collaborative approaches, whereby users could profit from the work of other users who have already carried out the geo-referencing process. This could yield, in a novel way, to provide coverage of urban areas by customized maps.

One idea would be to apply a web 2.0 approach allowing users to upload maps to a web site. This applies to photos that have been georeferenced by the user as well as to photos that have not been georeferenced. One could imagine that, if the same 'You are here' map has already been uploaded by another user it could be found (by an image similarity match) and used instead of the original photo. One could even try to use the matching function to perform the georeferencing of the newly take picture by using the information from the already uploaded version. This can also help us to collect a richer corpus with maps uniform distributed all over the world, because we noticed differences of 'You are here' map properties between different countries (e.g. in Germany more than 81% of the maps had the correct northing (nearly 92%) compared to the maps of Spain (there we just had 5 sample maps from Barcelona and Palma de Mallorca)). In addition such a photomap web 2.0 portal can also be used to get more knowledge about local reference system. We noticed that all 'You are here' maps of Barcelona, Spain we took were not correctly northed, but slightly rotated so that the coastline was at the bottom of the map and the streets were lined up as a more or less regular grid (see figure 3.27 (left)). This phenomenon is also known in the literature [85]. A bigger photomap corpus would help to investigate such effects. In addition a set of 'You are here' dots can be used to automatically detect the 'You are here' dot using computer vision algorithms (see figure 3.27 (right)). It is interesting to see that public maps are certainly not static. Here, we would like to investigate the dynamic character of these public maps. In Figure 3.28 (right) two maps from the exact location are shown. One photographed end of 2008 and the other was photographed in early 2009 (right). For example the POIs indicating telephone booths are no longer shown in the map of 2009.

In deepening of the second contribution, we also plan to apply more sophisticated georeferencing approaches. The most promising direction to follow is to use the GPS-trace of the user recorded by the device before the picture has been taken. By matching structural properties of the trace (e.g. turns and distance travelled) with structural properties of the map (e.g. pathways and streets) it could be possible to perform the georeferencing automatically in the background. Of course, we intend to carry out additional field studies to explore the concept of PhotoMap from the user perspective and extend our findings of contribution 3. Here we plan to contrast the different referencing methods to understand their advantages and drawbacks in terms of general

Figure 3.29: A hackneyed 'You are here' map.

usability. Again we think that is was very important to first investigate the feasibility of the concept using images of public maps for pedestrian navigation tasks. A pre-test showed that navigating with the PhotoMap application outperformed a mobile map application such as Google Maps mobile for the navigation on a hospital area. The implementation of PhotoMap (contribution 4), although already fully functional, can be further improved. We plan for example to allow users to access additional georeferenced information (such as Points of Interest) directly from the PhotoMap. This would integrate even further the paper based map with digital information available online. Unfortunately, an additional challenge for the approaches described in this paper is the vandalism that can occur to public signage. Figure 3.28 (left) shows an example of such damage to a 'You are here' type maps in the city of Florence, Italy or in the city of Paris, France and these examples are certainly not rare. With a web 2.0 styled online map collection tool different users can collect, share and merge their PhotoMap. With that we can address the problem of vandalism by looking for *older* and undamaged PhotoMap in the online map library. This tool can also be used to retrieve the georefering of other users' photos applying algorithms like SIFT [129] to compare a taken photo with already georeferenced images in the web application. Finally, it would be interesting to include additional information on the 'You are here' map usage available from the physical environment. For example information about where people have physically touched the map could be utilized. As can be seen in figure 3.29 certain areas of the map are hackneyed. Of course the actual position, but also other areas of the map were touched by many people. This information could be extracted with computer vision methods to get information about the importance of a point.

4

Tab Sized User Interfaces for Spatial Information

While the cornerstone of this thesis is the pad-sized interface, in this chapter we investigate the use of the tab-size interface for spatial information. A greater number of pad-sized devices exist today, but more and more tab-sized interfaces are coming to the mass market. Netbooks, eBook readers, and tablet PCs are typical examples of this device size class.

This chapter covers two research projects on tab-sized interfaces for spatial information. First, we present our work developing tab-sized interfaces designed for the geocaching community. An extensive study of the geocaching community was completed to better understand the requirements for an interface to service the community. We then analyzed the results of this study in the context of earlier findings from geocaching studies. From a technological point of view, it is interesting to see which media are currently used by geocachers in the field, especially with regard to their use of a combination of traditional and modern media. We then developed a new tab-sized interface for geocaching that is based on a GPS-enabled eBook reader. In our second project described in this chapter, we show how this GPS-enabled eBook reader can also be used to improve the reading experience. Using the digital and dynamic characteristics afforded by eBook readers, we describe the iBookmark system. In iBookmark, authors create stories that change in response to the location of the eBook itself. They can do so by setting context variables based on current and past locations of the eBook reader and using these in the rule-based generation of text and illustrations. In sum, with iBookmark we have developed a new rhetorical device for writers that extends the expressive range of eBook-delivered stories.

4.1 Geocaching

The activity of geocaching is attracting more and more attention both in terms of the number of geocachers as well as the number of researchers that have used geocaching as an interesting example of a popular location-based service [159] (and see table 4.1). To date, location-based and pervasive games have mostly been studied in rather artificial conditions in which the participants have been asked to use research prototypes [8; 15; 101; 140]. Geocaching is an emergent cultural behaviour that has not been designed from any central authority. Additionally, geocaching is even more interesting to researchers because it relies only on the most basic location-based service of all: a simple positioning service provided by a Global Positioning System (GPS) device.

4.1.1 Introduction & Motivation

According to O'Hara [159] Geocaching is a interesting study area, because "as a technology-enabled location-based activity, it has various attributes that make it significant to understand both in itself but also to inform our more general understanding of location-based computing practices" [159]. O'Hara identifies three interesting properties of geocaching:

1. First, it is a location-based experience that has established and sustained itself over several years (currently, about 10 years).

2. It involves both the *consumption* and the *creation* of experiences.

3. It is a game that uses extensively virtual and physical representations and involves a mobile hand set.

While O'Hara's [159] is the first investigation that tries to address the motivations and interests of geocachers, it could not address the community as a whole, given the qualitative nature and low number of participants in the study. In this section, we exploit the three attributes mentioned by O'Hara with two user studies.

Plattform	Worldwide	Germany
Geocaching.com	754.295	51.162
Opencaching.de	23.153	21.432
Navicache.com	11.389	6.253
Terracaching.com	5.560	87

Table 4.1: Main geocaching platforms, the number of caches world wide, and caches in Germany.

Table 4.1 displays the numbers of caches that were registered in early 2009 on various Internet platforms of the geocaching community[1]. As can be seen from the figures, approximately 10% of them are located in Germany. By attaching questionnaires both by placing them very prominently on various web pages and discussion forums (general study) and to the digital representations of caches (regional study), we have managed to reach out to a large part of Germany's geocaching community. In the two studies we have conducted, we have managed to obtain more than 2200 valid online questionnaires, which have helped us to draw representative conclusions on the structure of the German geocaching community. To our knowledge this is the first attempt to analyze a large community in the area of location-based services and in the spirit of criticisms of Greenberg and Buxton [74] we believe that our work provides data to verify and deepen findings of the sort that O'Hara [159] and Chaves [36; 37] have presented previously. In addition, this data set can provide insights into how next generation geocaching interfaces may look and how the research community can make them more integrated with existing technologies such as paper documents and how the community can increase the "calmness" of the technology can be increased, following the ideas of Weiser [239].

The next sections investigating the geocaching community in detail. The study results are mainly based on the thesis of Telaar [225]. We use the extract of it to support the implementation of our Geocaching interface. The first subsection describes the two users studies we conducted. A discussion of the several findings is given in the next section. For all results please refer to Telaar [225]. Based on these results we developed a new geocaching interfaces called "GeoBook" based on a modified eBook reader. We

[1]Please note that all platforms are rapidly growing, e.g. Geocaching.com now has about 901,395 registered geocaches (online retrieved September 2009).

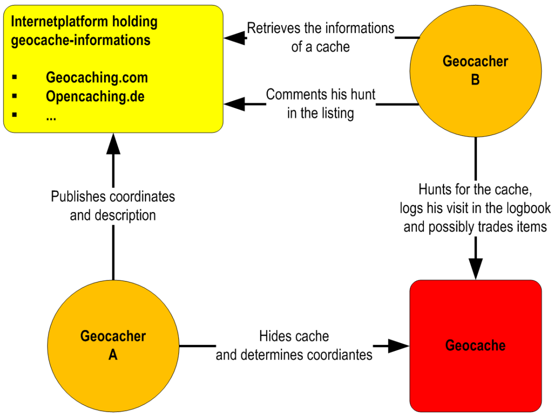

Figure 4.1: The geocache interaction model.

describe the implementation as well as the evaluation. Finally, we present our conclusions and ideas for future work in section seven. In contrast to the related work so far, no large-scale quantitative study has been carried out on an entire geocaching community. This study is necessary to empirically verify the findings of the prior pieces of work. In the following, we investigate more extensively the structure of the geocaching community and draw conclusions. We also present a novel tab-sized interface based on our conclusions.

4.1.2 Geocache Interaction Model

To understand the general procedure of geocachers, as well as the involved actors, we briefly present the geocache interaction model. In general the geocaching process can be subdivided into consumption and creation. Consumption is the more common activity, but the hiding of caches are the fundamental basis of this activity [201]. Both activities can be done by geocachers with GPS enabled mobile devices. Different Internet platforms serve as a backbone for the geocachers and is used to store and exchange all needed information. Figure 4.1 illustrates the involved actors and shows the general actives [156] and processes that are involved in the geocaching activity.

4.1.3 User Studies

In this section we will present the goals, methodology and the results of the two studies we have conducted. Since we were interested in a large data set we decided to use online questionnaires, instead of conducting interviews, which would have been too time consuming and thus not feasible to conduct with the number of participants we had in mind.

4.1.3.1 Study Goal & Methods

The first questionnaire was designed to assess the structure of the German geocaching community and was addressed to the general German geocacher. The goal was to collect demographic data on sex, age, income, to identify behavioural patterns and the motivation of geocachers as well as information on the tools they usually use. In the following we will refer to this study as the *general study*. The second questionnaire was targeted at a specific region in Germany. For the length of the study we observed all successful geocaching attempts in a radius of 20 km around the city of Münster for 45 selected caches observing the digital logbooks of theses caches on the web. An email questionnaire was send to the finders of those particular caches. We were interested to learn how a certain region is perceived by geocachers and how this perception changed after the activity. In the reminder of the section part of the study will be called the *regional study*. Both questionnaires have been distributed through Internet platforms, which administrate geocaches. Every geocaching attempt requires the player to look up the coordinates and additional information on the cache online and is usually concluded by reporting the successful geocaching activity, and which items have been removed and added to a particular cache. The use of these platforms allowed us to automate the process of data collection through questionnaires and resulted in high return rates. Instead of sending questionnaires via email (which can be blocked by filters or ignored by users), we decided to place the questionnaire on a dedicated server and placed a link to the survey on geocaching community web sites. Both questionnaires underwent a pre-test with 30 subjects to ensure that all the questions were understandable and that the time to fill out the questionnaires was reasonable. The feedback was also used to refine the questions. In order to motivate participants, they were given the opportunity to participate (after completion of the questionnaire) in a lottery and win small prizes, like Geocoins[1] and other small geocaching artefacts.

[1] A geocoin is a metal coin minted in similar fashion to a medallion, token coin, military challenge coin or wooden nickel, for use in geocaching. Some geocoins are trackable on the internet using a serial number and website address engraved on the coin [242].

4.1.3.2 General Study

The general study took place from May 21, 2007 to June 18, 2007. During these 28 days we had 3416 users visiting the start page of the questionnaire of which 2759 users started it. All of the submitted questionnaires were verified manually to identify partially filled-out questionnaires or non-serious attempts. Consequently, we used 1982 valid questionnaires for our further analysis. This equates to 71.83% of those that have started the questionnaire and 52.02% of those who visited the start page. In turn 21.16% of the questionnaires were incomplete or obviously filled out at random. These results make us confident that the questionnaire was easy enough to handle by most of the participants and that the design of the questionnaire had no major flaws. The return rate can be considered excellent, especially if compared to other Internet studies [42] or traditional "paper and mail" return rates [140]. The questionnaires have been advertised by different means.

- Firstly we have opened an email-thread on one of the major German discussion forums Münster discussing the advantages of such a study.
- Secondly we were able to place links on several web sites related to geocaching as much as on the start page of one of the major German geocaching-platforms[1].
- Furthermore we have distributed the link to the questionnaire through various mailing- and news-lists.
- Finally we selected on a random basis two geocaches in each federal state of Germany (32 caches in total) and asked (via email) all 1300 finders during the study period to answer the questionnaire.

The questionnaire consisted of 37 questions, a mix of closed and half-open questions, such as multiple-choice and ratings on a Likert-scale. A few open questions allowed the participants to provide comments and free thoughts. The questionnaire could be completed easily in 10 to 15 minutes. The feedback we have received by many of the participants was very positive, expressing the high interest of the German geocaching community in the goals of the study.

[1] http://opencaching.de/

4.1.3.3 Regional study

For the regional study we selected 45 geocaches near the city of Münster, Germany. The selection was driven by the results of the pre-test, which indicated a strong preference towards three types of geocaches: Traditional, Multi and Mystery cache (see related work chapter) of which 15 for each type have been selected on various geo-caching platforms[1]. We tried to balance the age, the spatial distribution and the vicinity to urban structure of the caches as well as possible. We used the email contact utility of the websites[2] to notify successful finders and ask them via email to participate in the study. We also visited all caches beforehand and placed little paper notes informing potential finders about the study goals. The study was conducted in the months of April and May 2007. During that period the caches in question were found 462 times. Finders have sent us 310 valid questionnaires. Some came from the same persons, since it was not very unlikely that one person found several of these caches during the study period. The basic structure of the questionnaire was similar to the one used in the general study. However it was slightly shorter (28 questions) and was targeted explicitly at the finder of that geocache and was addressed to get more information about how a certain region is perceived by geocachers and how this perception changed after the activity. In consequence the questionnaire could be completed much faster (5 to 10 minutes). Each questionnaire contained questions related to the particular geocache and the surroundings. The feedback of participants to the regional study was also very positive and expressed the high self-motivation and interest of the participants.

4.1.3.4 Results

This section reports the results of both studies. Given the high amount of returned questionnaires in relation to the assumed size of the overall community (approximately 10.000 geocachers in Germany in 2007) we consider the results as being highly valid and representative. We will concentrate on those results, which we believe could have an impact on the interaction of users with devices and the environment.

[1] for example `http://www.geocaching.com/`, `http://opencaching.de/`, and `http://navicache.com/`.

[2] `http://www.geocaching.com/` and `http://opencaching.de/`

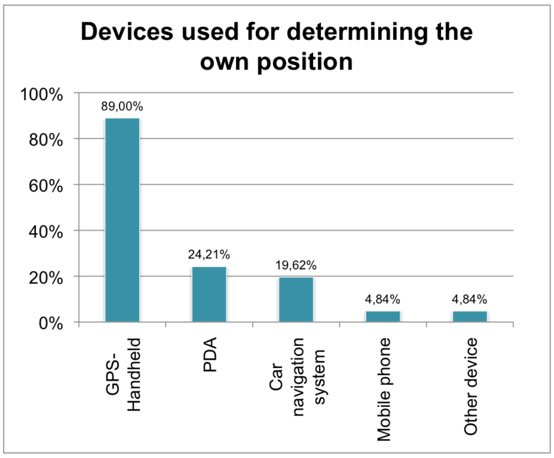

Figure 4.2: Devices used for determining the users postion.

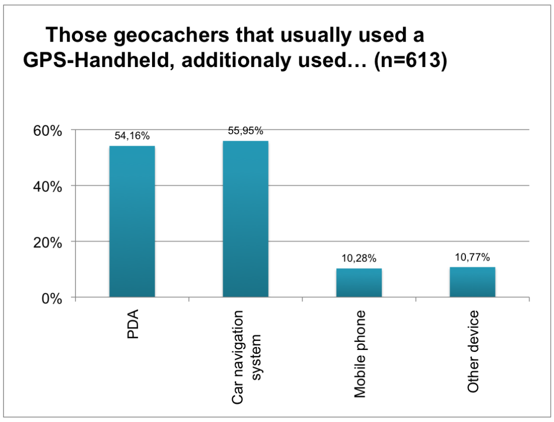

Figure 4.3: Additional used devices and artefacts by the cachers.

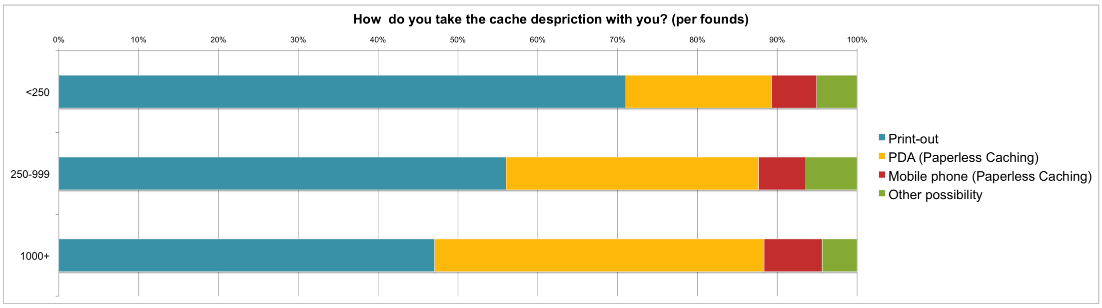

Figure 4.4: Ways of carrying around the cache description.

Results of the General Study In the following we will discuss the main implications that can be drawn from the general study. We focus on the used devices and the ways of carrying around the cache description to draw conclusion how to close this gap between analog and digital media. Instead we will elaborate more on findings related to *used devices and tools, motivation and interests* of geocachers as well as on the *context and properties of the caches.*

Used Devices The activity of geocaching involves the use of a localization device, usually based on the GPS. There exists a broad spectrum of such devices on the market ranging from simple GPS handheld receivers (which provide users with a geographic coordinate expressed in latitude and longitude) to personal digital assistant (PDA) and mobile phones that in principle can also provide cartographic information on the screen. We were interested to learn from geocachers which devices they prefer for their activities. As figure 4.2 indicates, the large majority of geocachers use a simple GPS-device to localize themselves. 1764 participants (89%) indicated this option. Roughly a quarter use a PDA (24.21%) and 19.62% a mobile car navigation system. Not very surprisingly most of the geocachers used GPS-devices, since they are often ruggedized and some of the newer models (for example the Garmin GPS 60 device series) provide special support for geocachers and run long on batteries. The responses revealed that 34.75% of GPS-device users used additional devices: 54.16% of those used a PDA, 55.95% a mobile car navigation system and 10.28% a mobile phone. Participants were also allowed to name additional devices and tools and approximately 5% of the geocachers made use of a compass and many indicated paper-based material such as traditional maps as well as print-outs from web mapping services (such as

Google Maps, Google Earth or Open Street Map). We asked participants whether they used digital mobile maps while geocaching, given that many devices support digital map formats. This was confirmed by 70.33% of participants. Of those geocachers the rate of digital map usage was highest with PDA users (87.30%), followed by mobile car navigation systems users (81.20%) and mobile phone users (76.00%). GPS devices were less used with digital maps (68.60%). These results are very much in line with what we expected given the technological capabilities of each device category.

Media types of cache descriptions One important preparation step of geocaching is to lookup basic information about the cache (e.g. coordinates and degree of difficulty). Multi-caches are often described by missions that have to be completed and mystery caches involve riddles that have to be solved in the field. Geocachers therefore need to bring this information along. When asked about the used media types, 90.72% of the participants reported to use paper print-outs, 33.10% used their PDA, 8.17% mobile phones and 7.52% reported to use other media types[1]. The type of media was dependent on the cache type. Multi caches often require mathematical calculations, which are better supported by analogue paper printouts than digital media. It appears that more active geocachers (more than 1000 caches found) use much more digital media than casual geocachers (less than 250 caches found). This relationship is expressed in figure 4.4. 71.03% of the less frequent cachers used paper printouts while only 40.09% of the most frequent geocachers rely on analog media.

Social context Being a game that can be played together and simultaneously we were interested how often participants were conducting the activity with friends and family. The responses reveal that 23.16% of the participants usually find geocaches on their own, 47.3% reported that they would equally geocache on their own as well as with others and 29.41% stated that they would carry out this activity only together with others. The 1523 participants of the two latter categories (all geocachers that like to be active in a group) were additionally asked with whom they usually find caches. 60.93% reported to go out very frequently with their partners, 20.75% said that they would sometimes do it. 26.00% of the participants reported to frequently take the whole family on a geocache search. Also of interest was the answer to the question of

[1]Participants were allowed to give multiple answers.

how often non-geocaching friends have been taken on a trip (64.28%). Overall it can be stated that geocaching is a social activity that brings together family members, friends and other geocachers.

Selection Principles We were interested to learn how geocachers select a cache that they would like to visit. To investigate this question we asked the participants to chose amongst three statements: a) "I usually stick to my favorite cache-type and avoid other types", b) "I make a choice depending on my current interests and needs" and c) "I have no dedicated preference, I just want to find all caches in my surroundings". The results splits the geocachers into two categories: 47.78% were purely interested to find all caches in their regional proximity (option c), and 34.66% agreed to option b), indicating that they make an informed decision based on their current interests and needs. Only 9.33% were focusing on one particular cache type (option c). 8.53% reported completely different selection principles in the free text section of the questionnaire. For example participants reported that the decision, which cache to select, depends on the company and the group. For example, with family members they would carefully reflect the degree of difficulty (option b), while when on their own they would just try to find as many caches as possible (option c).

Motivation We tried also to shed light on the motivations of geocachers and offered participants nine statements that they were asked to rate on a four-point Likert-scale (see figure 4.5). According to the results, the main motivation for the activity of geocaching is the possibility to be in nature (nearly everybody either agreed or totally agreed with the corresponding statement). Although, as we discussed earlier, most of the participants enjoy geocaching with others, this seems to be less important. Getting more familiar with the regional environment as well as being able to discover new places is another important motivation.

Results of the Regional Study The results of the regional study are also not discussed. The properties and characteristic of a geocache that attract cachers and their impact in their enviroment are not in the focus of this thesis. As said earlier, more detailed can be found in [225].

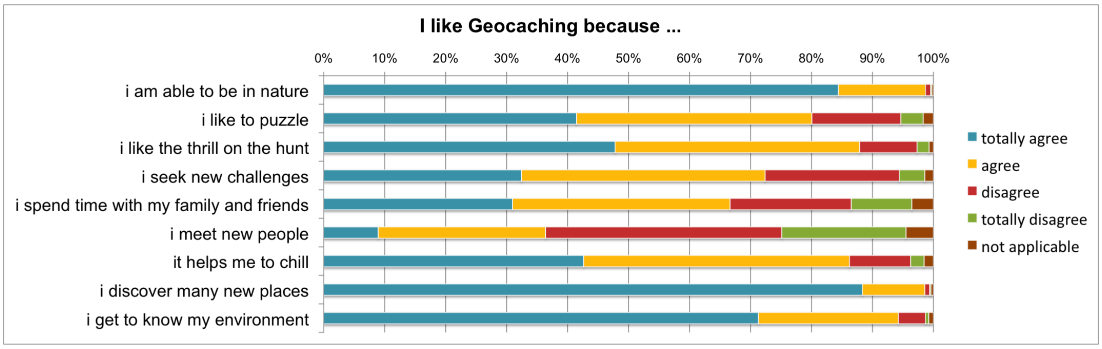

Figure 4.5: Motivation of geocachers.

Discussion of the User Study In this section we discuss some of the results of our representative study and compare our findings to previous work, i.e. the work of Chavez [36; 37] and O'Hara [159]. We can mostly confirm the demographic data of Chavez, although our study area is twice as big and the amount of returned questionnaires 15 times as large. Our data indicates slightly more male geocachers than Chavez has reported. We have looked at a variety of additional factors including the educational level of participants and we found that geocachers have an above average level of education. We have made use of the community platforms to spread our questionnaire and developed a new method to investigate caches within a certain region. By addressing successful geocachers in a certain region via email, we achieved an extremely high return rate of more than 60% with very relevant responses. In line with O'Hara and Chavez, we can therefore confirm, that one of the main motivations of geocachers is to be out in the nature and to discover new places and locations. We could also confirm that geocachers usually perform their activities together with others and could, for the first time, quantify this quite precisely. In line with O'Hara we assumed that this would be also one of the major motivations for geocachers in general. Surprisingly, it played a less important role than expected. One explanation could be that the attractiveness of geocaching is the fact that it is a socially more acceptable use of technology and therefore gives highly active geocachers the possibility for frequent geocaching trips together with friends and family. This hypothesis is supported by the fact that many geocachers are accompanied by non-geocachers during their activity. We could also show that there are two major categories of geocachers that split the community: the *collector* and the *gourmet*. Geocachers of the first category are mainly interested in making as many caches of a certain region as possible; O'Hara [159] had identified a similar category. The second category consists of geocachers that choose on a very individualistic basis their next caches. These findings are significant for informing the design and placement of new caches. For collectors, new caches need to be placed in sparse cache areas, while gourmets ask for exciting locations or difficult riddles. Beside consumption, O'Hara highlights the importance of creation of geocaches for the 14 participants of his study. We could not find much evidence for that particular claim. Our data shows that only one third of the participants have ever created a geocache on their own. Similar to other communities (such as the Wikipedia community) consumption is clearly the more commonly observed behaviour. Beside the demographic data and the social dimension using GPS for the geocache activity we also examined the devices used. Although simpler GPS-devices were mostly used it was surprising to see how many users still used PDAs and mobile phones as well as mobile car navigation systems. In addition geocaching novices often took a printout description of the cache with them. Next generation of geocache interfaces should allow the combination of paper-based media with new devices such as GPS equipped mobile phone. Applications such as Wikitude[1], Layer[2] or Wikeye [194] would be ideal to be used as geocaching interfaces. The fact that a large fraction of geocachers uses more than one device also surprised us and next generation interfaces should combine the all different functions necessary for a successful geocaching experience. A first approach for a new interface is presented below with the GeoBook interface. Moreover rural areas can profit from geocachers and try to attract wealthy geocachers targeting the geocache websites and Internet communities. Additionally geocaching, as a technology driven activity, can support help people to corporealize.

4.1.4 The GeoBook Interface.

The results from the studies were used to develop a new user interface for geocaching activities. We focus on aspects how to combine the different media types normally used by the geocachers. As stated in the result section often (especially for multi caches) analogue paper printouts was used in combination with digital media (GPS, PDA, or similar device). With the "GeoBook" interface we want to combine the digital and

[1] `http://www.wikitude.org/`
[2] `http://layar.com/`

analog worlds better and improve the geocaching experiences. The following section will shortly introduce the interface, the implementation and report on a preliminary evaluation. The new user interface mainly consists of a modified cache description. The cache description is crawled on demand from the website and than enriched with the current location. This makes the whole geocache process easier, because the users do not have to switch between the analog and the digital media. As noted above multi caches often include some calculation or small quizzes. For example a instructions could be like: "The final cache, a bison container, is located at 47 38.(a+d+e)d(c-a) and 122 20.(b+e)a(d+e-1)." (where a to e are coordinates at other locations of this cache.) Normally the users had to write this coordinates down on a sheet of paper and than make the calculation by hand. Than the had to type in the calculated coordinates again by hand into the GPS device. Currently the user interfaces (UI) of standard GPS device are still quite complex and not well suited for beginners. With our interface the locations a to e were stored on our modified eBook reader and than presented to the users at the moment they had to solve the task. The users can manually calculate the next coordinates or use a build in calculations function. A screenshot of a user with the ebook reader can be seen in figure 4.6. The "Geobook" interfaces is based on an iLiad eBook reader by iRex Technologies. The same technology was also used in the iBookmark project [204]. The reader connects to local wireless networks, using a web browser to render the story contents and allow user navigation through a geocache portal. In order to trace the location of the reader, a GPS/GPRS module based on the Telit GM862-GPS chip is attached to the eBook reader. This module posts back its current position to the server at regular intervals over GPRS. The cache descriptions of the portals are than adapted to the current location by a standard Linux-Apache-MySQL-PHP (LAMP) server, which produces a new page with the cache description adapted to the location, e.g. filling in current coordinates into calculation fields. We conducted an initial user study to evaluate the new GeoBook interface with 3 groups of 2-4 users and to compare subjects' self-reported experiences. The first group consists of two male geocaching experts (more than 25 geocaches). The second and the third group consist of 3 respectively 4 female geocaching beginners ($\leq$ 2 caches). With all three groups a multi cache was solved (about 1h-2h cache time). Overall, both novice user groups, really like the GeoBook interface and they gave us comments like: "It is cool - I do not have to use the bulky GPS anymore" or "Everything is on one screen.

Figure 4.6: The geobook interface. The interfaces is based on an iLiad eBook reader by iRex Technologies in combination with a GPS/GPRS module.

Really easy to use compared to my Garmin GPS device". Also the experts gave us promising comments. They like the integration of the description and the 'Tunes like option to download the different cache description on the fly". They said us that this is very helpful, e.g. if there is time for another cache in that area. With the GeoBook interface they are able to ad hoc do a cache without great preparation. Both statements underline the argument that the user believes the GeoBook interface could be a next generation interface.

4.1.5 Discussion & Future Work

In this section, we have presented a large scale quantitive survey of the German geocaching community. We have partly confirmed the studies of O'Hara [159] and Chaves [36; 37]. The geocaching community is still rapidly growing and compared to O'Hara [159] we found that the consumption of caches is more important for most of the users than the creation. The geocache activity is an interesting research field, because the activity was created by early adopters of a new technologies and the research community can help to design next generation interfaces to keep the geocache community growing. Compared to Chaves [36; 37] we examined a larger area, but it would be interesting to investigate cultural differences on a world wide scale or investigate the geocache activities in developing countries, where internet and mobile devices are not so commonly used. This leads us to another point: What are the geocache devices of the future? We found out, that people often use ordinary GPS devices in combination with other electronic devices and analogue artefacts (such as paper print-outs and maps). This can help to open the current geocaching community (mostly well educated male experienced internet users) to currently under-represented communities by creating new kinds of devices especially designed for people unfamiliar with the geocaching interaction model (see figure 4.1). Finally, we are investigating if we can transfer the knowledge and dynamics of the geocache community to inform and stimulate the design of other location-based services. On the one hand it would be interesting to see if new geocache interfaces can amplify the current geocaching hype. With the GeoBook interface we make a first step into that direction. On the other hand it would be interesting to see if the "hype" will seep away, because more and more geocachers become pure consumers or the cache density saturates in nature [116; 156]. We are interested to compare the act of geocache creation to those of other web 2.0 creation processes, such as authoring Wikipedia articles [143] and to investigate the balance of creators and consumers. While the main contribution of the section is the presentation of a rich quantitative survey, we think that our analysis grounds related work effectively and places certain earlier findings from geocaching studies into the right perspective. Two observations from our data set are particularly relevant to this regard. Firstly, the part of our data set relating to consumption (of the experience) versus creation (of a geocache) seems to conform to Nielson's 90-9-1 theory relating towards the general pattern of user generated content creation [156] which stipulates that only 1% of online community users will contribute regularly, 9% will contribute intermittently and the remainder will "lurk", consuming content without contributing. Secondly, the data shows some interesting facts, such that geocachers are often accompanied by non-geocachers and family members, but being with others seems not to be the primary motivation for the activity. From a technological point of view it is interesting to see which media are currently used to bring the cache description into the field. We think that the use of a combination of traditional and modern media is especially interesting. It is our plan to make the whole data set available, allowing others to draw their own conclusions and perform their own analysis that may go beyond what we have found in the data (after we published our results). We believe that other researchers should be aware of this data set. And concluding according to Greenburg and Buxton [74]: Why not repeat the study in four years again? And do it again in another four years?

4.2 iBookmark

With the recent developments in ePaper technology, consumer eBook readers have display qualities and form factors that are approaching that of traditional books. These eBook readers are already replacing paper in some commercial domains, but the potential of eBooks to extend forms of writing and storytelling has not been significantly explored. Using the digital and dynamic characteristics afforded by eBook readers, we are developing iBookmark, a GPS-enabled eBook reader. In iBookmark, writers can create stories that change in response to the location of the eBook itself. By setting context variables based on current and past locations of the eBook reader and using these in the rule-based generation of text and illustrations. We are developing new rhetorical device for writers that extend the expressive range of eBook delivered stories.

4.2.1 Introduction & Motivation

The first eBooks were written for specialist domains and for a small-devoted group of readers, for whom the equivalent paper documentation was prohibitively large. The scope of the subject matter of these e-books included technical manuals for hardware, manufacturing techniques, and other subjects. Current consumer eBook readers, such as Slick (Foxit), Kindle (Amazon), Sony Reader (Sony), iLiad (iRex Technologies) and iPad (Apple) are allowing publishers to distribute books that are either in the public domain or specifically targeted at the eBook market. Interestingly, as with the Internet-based distribution of popular music, eBooks have the potential to be a disruptive publishing technology, allowing authors to distribute text outside the confines of mainstream publishing and conventional book production. The reproduction of the form factor of the paper book has been the principal motivating force in the development of eBook readers. However, eBook readers afford the practical delivery of new forms of writing that lie at the intersection of digital technologies and traditional storytelling so called interactive stories [152]. Our goal is to explore a new form interactive storytelling that is afforded by the book-like nature of eBook readers themselves. These locative texts are stories that respond to the place at which a book resides and is read. This work is related to the field of Alternate Reality Gaming [139; 224], but the gaming component is not the main focus. It could be defined as Alternate Reality Experience. As our motivation, we have considered the relationship between a text and the place

Figure 4.7: The iBookmark eBook reader. A GPS unit is attached to the reader for actual sensing various context variables based on the location of the users. Stories read by users on our modified eBook reader evolve and change every time a user reads at a different location. Here, the title adapts based on the current location. A photograph of the location is used on title page and, a result of the story, an "alien spaceship" is composited in the image.

at which it is read [89; 99]. On one hand, the traditional role of storytelling can be thought of as to remove the reader from their time and place and immerse them in them in an alternative world. Alternatively, reading can be cast as a device for causing a user to reflect on their current situation. In contrast to traditional interactive storytelling research [24], we embrace the passive quality of reading (in user interaction terms) and instead our aim is to explore new forms of writing that use an awareness of the place-of-reading to adapt texts dynamically. Locative texts have the potential to adapt according to their locale, either allowing aspects of the real world to leak into the text, or force a juxtaposition of the real and the written (much as the theory of "suspension of disbelief" aids theatre [216]). In practice, we envisage a more subtle and creative space than these extremes suggest. Our first step has been to equip an eBook reader with a GPS/GSM unit for sensing various context variables based on the location of the users.

4.2.2 Basic Interaction Concept

One advantage of being able to contextually change the story is altering it to represent the opposite of the current context. For example, the weather in the story may be sunny and hot, unlike the stormy conditions where the reader is. This likely aids readers in loosing themselves in the story, enhancing the suspension of disbelief of the reader. By altering subtleties in the story, the emotional connection the reader makes as they read will be different depending not just on the reader themselves, but also on their surroundings. This allows for different interpretations by the same reader, not just by different readers. Various forms of information are used as context variables, including the user's current location, past locations, current time, twitter feed, and Gmail account. From location information, for instance, we can infer a multitude of other information, including current weather conditions, local time of day, place names, photos of surrounding areas and country specific information such as government names. From email data, we can produce character names, combining subtleties from emails with name extensions to produce both standard and fantasy names. Also, writers can create different storylines for different day times. A story read on the weekend would be different from a story read during the week, for instance. Of course, one can think of many different authoring variations using the context variables.

4.2.2.1 Context Variables

The iBookmark framework allows for any pre-published story to be used, by providing a large number of simplified variables, which can be added to each page in the story. These variables are specific location dependent properties abstracted from the location of the iBookmark reader, and can be accessed directly from the page in the story, for example the author could simply ask for a new female character name, or current weather condition. In addition to the author including a context variable in the sentence, a parameter is included that allows the variable to be persistent, so important character names once set will always be the same, whereas place names and weather may change. Along with these helper functions, the author may include their own images overlayed with a location dependent background image, which can be specified by

Figure 4.8: The iBookmark system architecture. The eBook reader sends its current location to a server over GPRS to create stories that change in response to various context variables derived from the location.

using keywords according to the context of the story. Story variables can be thought of much like programming language constructs, with local variables represented by a non-persistent story context and global variables representing names and important places. By introducing the element of passing of time (this can be obtained from the GPS timestamps), variables can be changed over time, or according to distinct events such as sunset. In essence, the process of converting a traditional story for use with the iBookmark is a simple one, consisting of replacing context aware parts of sentences with helper variables defined in the framework. This process can be carried out by either the author or a third party, although it is noted that by allowing the author full reign over the context may result in new types and styles of writing, much more transient and emotional that traditional story building.

In Figure 4.9 snippets from a sample story are presented. The story "Last days" (adapted from a pre-existing electronic story from the web called "Second Thoughts")

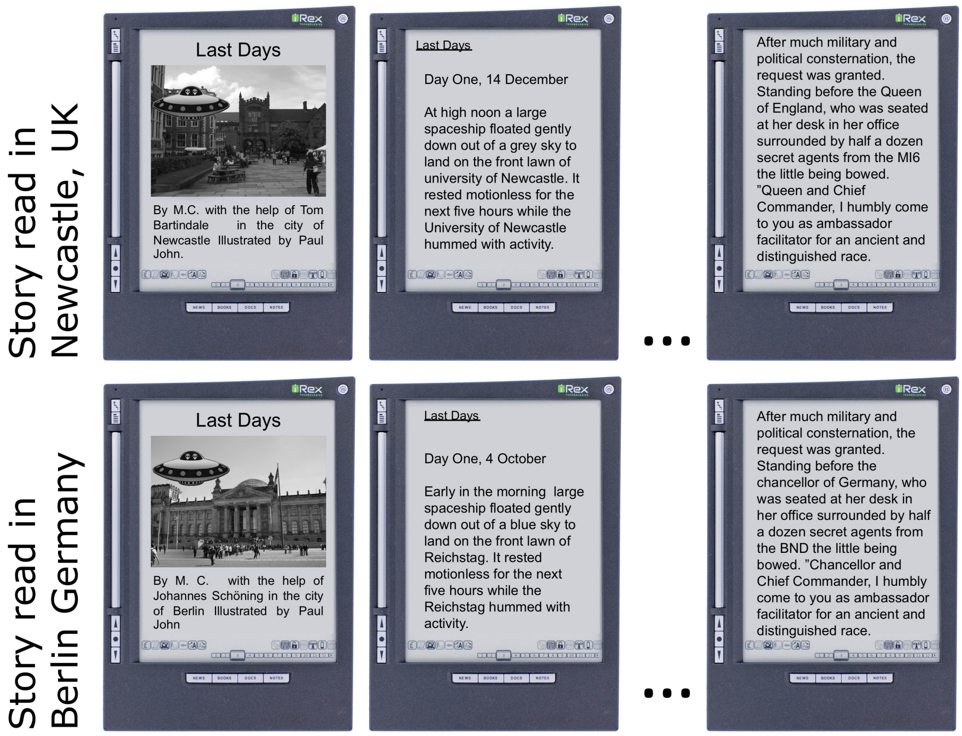

Figure 4.9: Snippets from a sample story "Last days". The story (adapted from a pre-existing electronic story from the web called "Second Thoughts") is about an alien invasion on the earth.

about an alien invasion on the earth. The story on the top of Figure 4.9 is the story like it would be appear on the user's eBook reader reading the story in Newcastle, UK. The story on the bottom is the same story read by a user in Berlin, Germany. In the story, various forms of information are used as context variables, including the user's current location, current time, and Gmail account. Derived from the actual location information of the reader we infer the current weather conditions, time of day, place names, photos of surrounding areas and country specific information (e.g. the government names and the names of the secret service) to use these variables to adapt the story.

4.2.3 Implementation

The system is based on an iLiad eBook reader by iRex Technologies. The reader connects to local wireless networks, using a web browser to render the story contents and allow user navigation through the story. In order to trace the location of the reader, a GPS/GPRS module based on the Telit GM862-GPS chip is attached to the eBook reader. This module posts back its current position to the server at regular intervals over GPRS. The story content is produced and rearranged by a standard Linux-Apache-MySQL-PHP (LAMP) server, which produces the next requested page in the story using a pre-defined templates, filling in appropriate variables with location, time and other context data. Stories are written such that variables that hold importance to the story will not change once set. For our first prototype, we generated a story from a pre-existing electronic story from the web, simply substituting place names, character names and location information. This proved far less interactive for readers than the next version, in which we commissioned a professional writer to create a story that fully utilized all the features of the iBookmark. Knowing what was possible inspired the writer to produce a story in a different manner than he would normally write. We think that this new wave of creativity gives readers a new and enhanced reading experience.

4.2.4 Discussion & Future Work

We have presented the first combination of an eBook reader with a GPS Device to create stories that change in response to the location of the eBook itself, as well as other context variables. The current prototype is only a first implementation of our ongoing work on this concept. We would like to make the GPS/GPRS more unobtrusive, designing it like a bookmark for the eBook reader. In addition, we are currently building an authoring tool to help professional writers to create locative texts and stories. Finally, we are interested in the users reaction to the strong relationship between a text and the place at which it is read.

5

Board Sized Interface for Spatial Information

In this chapter, we discuss board-sized interfaces for spatial information. Our focus in particular is on multi-touch-enabled large surfaces. Interest in multi-touch interaction with large- and small-display surfaces has increased enormously over the last few years. The technical realization is described first. While this is independent from the presentation of spatial information is it still an important cornerstone, while the "do-it-yourself" character of these interfaces is unique. In addition it is also not connected to a specifc size of the interfaces, but most of the home grown multi-touch tables have board size. Our goal is to enable graphics and interaction design practitioners to embrace multi-touch by providing the basic knowledge required to *build your own* multi-touch surface. The second part describe novel multi-touch gestures for spatial tasks basic to a GIS, e.g. pan, zoom and selection operations. We also are interested in developing new forms of data that more fully leverage the potential of multi-touch in the spatial domain (and outside of it). We investigated the combination of this new interaction paradigm with novel spatial data schemas. A secondary focus of this chapter is describing first approaches how these new multi-touch paradigm that combine traditional 2D interaction performed in monoscopic mode with 3D interaction and stereoscopic projection. We refer to systems that support these paradigms as interscopic multi-touch surfaces (iMUTS) and briefly discuss them in the second half of this chapter. In the final section, we present a pad-sized interface that allows spatial authentication on large interactive multi-touch surfaces, thus combining pad-sized and

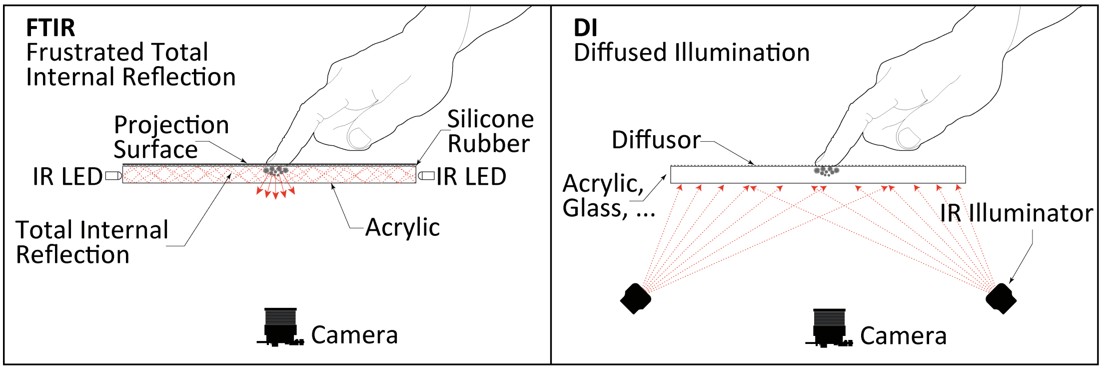

Figure 5.1: General set-up of a FTIR system (left). General set-up of a Diffuse Illumination system (right).

board-size devices.

5.1 Technical Realization of Multi-touch Surfaces

Many techniques, such as Resistance Based-, Capacitance Based-, or Surface Wave-Touch screens, generally require industrial fabrication facilities. By contrast we focus exclusively on optical approaches to multi-touch sensing as these can be built quickly and easily integrated into graphical user interfaces. As said in the introduction the "do-it-yourself" character of these interfaces is unique. Therefore this section was actively developed with other members in community during different workshop starting in 2007 (Multi-touch workshop in Münster, Germany and Multi-touch Bootcamp in Amsterdam, Netherlands).

5.1.1 Optical Based Touch Surfaces

Optical approaches to multi-touch use image processing to determine the location of interactions with the surface. These systems typically use infrared illumination, and due to their simple set-up have the potential to be very robust. In addition to FTIR and DI we discuss two other related, but distinct, approaches: Laser Light Plane (LLP) and Diffused Screen Illumination (DSI) [205]. This technology is not new, having been available in different forms since the 1970s. The rediscovery of the FTIR principle [79] has greatly accelerated the development of new multi-touch applications. Han's YouTube demonstration captured the imagination of experts and laymen alike. His system was both cheap and easy to build, and illustrated a range of creatively applied multi-touch interaction techniques. In 2007, Apple presented their new mobile phone, the iPhone[1]. Where other touch based cellular phones only allow single point interaction, the iPhone uses multi-touch technology. The resulting interaction techniques and interfaces have received considerable media attention and brought multi-touch interaction to the consumer electronics market. Later in 2007 Microsoft presented their version of a multi-touch table, the MS Surface. The table has the appearance of a coffee table with an interactive surface. The sensing technique used in MS Surface is similar to the HoloWall [137] exploiting a diffuser which is attached to the screen material. The table surface is illuminated from behind with infrared light and when a user touches the table reflected infrared light is captured by cameras inside the table. Because of the use of multiple cameras, the input resolution is high enough to detect objects.

The remainder of this section is structured as follows: We start with a brief overview of existing (multi-)touch technologies. Section 5.1.3 focuses on the technical challenges users face when constructing a **"build your own"** multi-touch surface, namely (as described in section 5.1.3.1) infrared illumination, camera set-ups, filters, projectors, silicone compliant surfaces, projection screens and the integration of all this hardware into the final multi-touch surface. Section 5.1.3.3 focuses on existing software libraries which allow deployment of multi-touch applications upon optical multi-touch surfaces. Finally, we describe a selection of interesting projects currently utilizing the technologies that we describe.

5.1.2 Touch Technologies

Before describing FTIR and DI, it is important to note that there are a number of alternative technologies that can be used to construct multi-touch surfaces:

- Resistance Based Touch Surfaces
- Capacitance Based Touch Surfaces
- Surface Wave Touch Surfaces (SAW)

[1] `http://www.apple.com/iphone/`

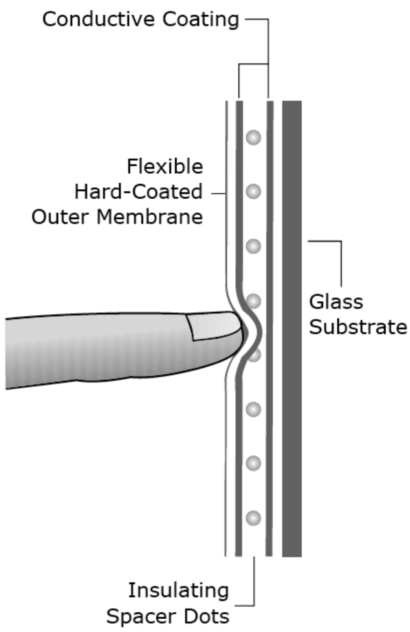

Figure 5.2: Schematic construction of a touch screen based on resistive technology.

Unfortunately, these technologies require industrial quality fabrication facilities to construct and are therefore not suitable for self-built surfaces. However, we discuss them here for the sake of completeness.

5.1.2.1 Resistance Based Touch Surfaces

Resistance based touch surfaces generally consist of two conductive layers that are coated with substances such as indium tin oxide. These layers are separated by an insulating layer, usually made of tiny silicon dots (see figure 5.2).

The front of the panel is typically made of a flexible hard coated outer membrane while the back panel is often a glass substrate. A controller alternates between the layers, driving one with a specific (electric) current and measuring the current of the other. When users touch the display, the conductive layers are connected, establishing an electric current that is measured once horizontally and vertically by the controller in order to determine the exact position of a touch. Such touch surfaces have the advantage of low power consumption and are used in mobile devices such as the Nintendo DS[1],

[1] http://www.nintendo.co.uk/NOE/en_GB/nintendo_ds_1023.html

PDAs and digital cameras and can be operated both with fingers or a stylus.

However, resistance based surfaces provide a low clarity interactive surface (about 75%–85%) and additional screen protection cannot be applied without impacting on their functionality. More detailed information about classical resistance based (multi-) touch surfaces can be found in [51].

5.1.2.2 Capacitance Based Touch Surfaces

In general capacitance based (multi-) touch surfaces can be subdivided into two classes:

- Surface Capacitance
- Projected Capacitance

Both techniques were primarily developed for single touch interaction. One advantage of capacitive touch surfaces in comparison to other technologies is their high clarity, which makes them very suitable for use in many kinds of touch displays beyond simple touch pads. Capacitive touch screens can also be operated by any conductive device and are hence not limited to finger based interaction. However, capacitive touch panels are relatively expensive to produce, although they do exhibit high durability and reliability. Consequently, capacitive based systems are often preferred for use in rough environments such as public displays and industrial applications.

It is possible to use capacitive based systems for multi-touch surfaces, but typically the number of simultaneous touches is limited by firmware or by the design of the controller. Finally, accuracy decreases when performing touches with more than one object. Having noted these general limitations, capacitance based technologies developed e.g. by MERL overcame many of these restrictions in order to allow many simultaneous touches. These are briefly described below.

Surface Capacitive Touch Surfaces Surface capacitive touch panels consist of a uniform conductive coating on a glass layer. Compared to resistive technologies, a much higher clarity can be achieved by using indium tin oxide [243] as the conducting material (it is transparent as well as colourless when used in very thin layers). From each side of the touch panel, electrodes maintain a precisely controlled of store or electrons in the horizontal and vertical directions which sets up a uniform electric field across the conductive layer. As human fingers (or other conductive objects) are also

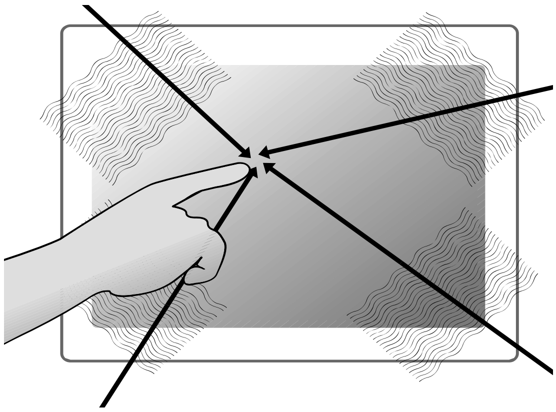

Figure 5.3: Surface Capacitive (Multi-) Touch Surfaces: Electrodes around the edges distribute voltage across the conductive layer creating an electric field. Touching the panel results in current drawn from each corner which is measured to define the position.

electrical devices capable of storing charge and exhibiting electric fields, touching the panel results in a small transport of charge from the electric field of the panel to the field of the touching object. Current is drawn from each corner of the panel; this process is measured with sensors located in the corners, and an microprocessor interpolates an exact position of the touch based on the values measured (see figure 5.3). Panels based on surface capacitive technology can provide a high positional accuracy.

Projected Capacitive Touch Surfaces Of the technologies we describe, projected capacitive touch devices are the most expensive to produce. Their performance is worse than many of the other approaches. However, they afford superb mechanical resilience. Projected capacitive surfaces can also be covered by a non-conductive material (with a maximum thickness of around 20mm) without negatively impacting on their functionality. When used for (multi-) touch displays, as described by Rekimoto [188]) a very thin grid of microphone wires is installed between two protective glass layers (see figure 5.5). When touched, capacitance forms between the finger and the sensor grid and the touch location can be computed based on the measured electrical characteristics of the grid layer. The accuracy of projected capacitive technology is similar to surface capacitive technology, although light transmission is superior because the wire grid can be constructed such that it is nearly transparent. The technology is also highly suitable for rugged environments such as public installations, as a protective layer (such as thick

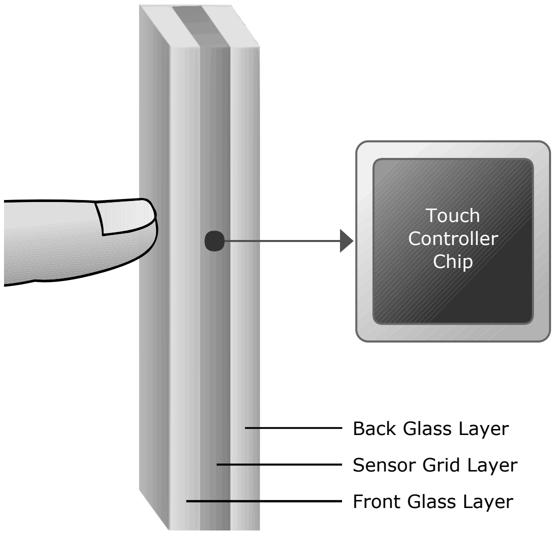

Figure 5.4: A thin grid layer is protected by two glass layers. Capacitance forms between the finger and the grid during a touch. The change of electrical properties is measured to determine the touching according position.

glass) may be added without drastically decreasing the sensitivity. Finally, multiple simultaneous touches can be more easily interpreted compared to surface capacitive based technology.

MERL Capacitive Diamond Touch In 2003, *DiamondTouch* was developed in the Mitsubishi Electric Research Laboratories (MERL). *DiamondTouch* was designed to support multiple touches, be tolerant of objects placed on the table, be durable, and be inexpensive to manufacture [49]. The system has a number of additional distinctive characteristics:

- the ability to handle many touch points and users (only limited by the size of the table and the available space around it);
- the ability to identify which users are interacting with the surface;
- it is affected by debris objects like cups placed upon its surface
- it does not require additional devices for interaction (such as special pens).

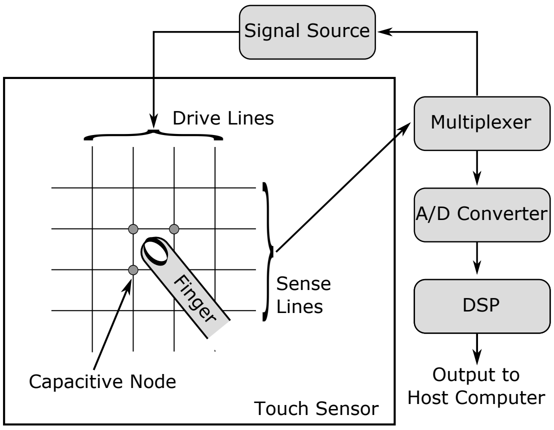

Figure 5.5: A simplified view of projected-capacitive touch screen adapted from Rekimoto.

Furthermore, the isolating layer between the antenna array and the user can be manufactured from a wide array of materials. Hence, it can be made to be very robust. For example, when a special fibre glass laminate is used, alcohol may be ignited on the surface without causing damage. Using capacitive coupling *DiamondTouch* is composed of a table with integrated antennas transmitting unique signals, a ceiling-mounted projector presents a display onto the table, one conductive chair connected with a receiver for each user and a computer. "When a user touches the table, a capacitively coupled circuit is completed. The circuit runs from the transmitter, through the touch point on the table surface, and through the user to the user's receiver back to the transmitter.“ [49]. *DiamondTouch* works by transmitting signals through antennas in the table; these signals are used to identify the parts of the table each user is touching. This information can then be used to calculate [49] the finger's position. Usually a user touches several antennas at once. For this reason the signals have to be separable (in technical terms orthogonal). This can be achieved by frequency-division multiplexing, time-division multiplexing or code-division multiplexing. The antenna pattern consists of two layers similar in design, but with one rotated by ninety degrees. The rows/columns (antennas) of each layer are composed of diamond shapes connected in one direction and isolated in the other. In this way, the covered surface is maximised and the shielding effect minimised. Usually there is an antenna every five millimetres (which is resulting the minimum pointing accuracy). Due to image projection from above the only obstructions are shadows cast on the table by objects (i.e. hands or arms) inserted into the projector's light beam.

Surface Wave Touch Surfaces (SAW) Systems that use surface wave technology are similar to those that use infrared grid technology. Transmitting and receiving piezoelectric transducers, for both the X- and Y-axes, are mounted on a faceplate and ultra-sonic waves on a glass surface are created and directed by reflectors. By processing these two electronic signals and observing the changes when the faceplate is touched, it is possible to calculate the position of that interaction. Most SAW systems can support dual-touch.

5.1.2.3 Optical Based Touch Surfaces

Both optical and camera based-approaches share the same concept of processing and filtering captured images on patterns. As already discussed, a number of systems are based on infrared illumination and as a result can suffer interference from ambient light in the environment. However, due to their simple configuration, optical approaches have the potential to be very robust.

5.1.2.4 Frustrated Total Internal Reflection (FTIR)

Han's work in 2005 [79], which utilised the principle of FTIR in multi-touch interaction, can be seen as the critical point in the development of optical systems. The FTIR approach is based on optical total internal reflection within an interactive surface. Electromagnetic waves transmitted within an inner material are completely reflected at its boundary if: (1) the inner material has a higher refractive index than the outer material; and (2) the angle of incidence at the boundary between the materials is small enough. The most common FTIR set-up has a transparent acrylic pane, with a frame of LEDs around its edges, into which infrared light is injected (see figure 5.1 (left)). When the user touches the acrylic, the light escapes and is reflected at the finger's point of contact due to its higher refractive index; an infrared-sensitive camera can then clearly see these reflections. A basic set of computer vision algorithms (see section 5.1.3.3) is applied to the camera image to determine the location of the contact point. As the acrylic is transparent, a projector can be located behind the surface (near to the camera) yielding a back-projected multi-touch sensitive display (see figure 5.1 (left)).

Diffuse Illumination (DI) systems are similarly configured, with both a projector and an infrared sensitive camera placed behind the projection surface. However, for DI, infrared lighting is placed behind the projection surface; causing the area in front of the surface to be brightly lit in the infrared spectrum. Consequently, the camera is capable of detecting the reflection of fingers and objects on, or in close proximity to, the surface (see figure 5.1 (right)). Touch detection exploits the fact that the projection surface diffuses light, blurring objects at a distance. The main advantage of FTIR is that it allows very robust tracking of fingers, however, diffuse illumination additionally allows tracking of physical objects which can be identified either by their shape or the use of fiducial markers [43] (easily recognizable markers) on the base of the objects. Furthermore, hovering gestures can also be recognized, and any transparent surface (such as safety glass) can be placed between the projection screen and the user since sensing does not rely on surface contact.

5.1.3 BYO Multi-Touch Surface

When designing and constructing an optical multi-touch surface a number of challenges need to be addressed. In this section, we divide up these issues as they relate to both hardware and software, and provide practical advice based on our own experiences of developing robust tabletop systems.

5.1.3.1 Hardware

The hardware of an optical multi-touch system comprises: infrared illumination sources, silicone compliant surfaces, projection screens (or the use of LCDs), cameras, filters, and projectors.

Infrared Illumination Both FTIR and DI require an infrared light source. Achieving the right infrared illumination can be challenging and requires a knowledge of both the different methods of illuminating a surface and different the types of IR LEDs (5mm, 3mm, SMD) that are available commercially. Almost all existing IR-based set-ups employ light-emitting diodes (LEDs) as light sources. Two commonly used types of IR LEDs are Osram SFH4250 (SMD) and Osram SFH485 (5 mm). Whether SMD devices or standard LEDs are more appropriate depends on a number of factors, for example, if the LEDs have to be mounted to the rim of an acrylic glass plate, this is easier with

SMD, as it is possible to simply attach them to the rim with instant glue. After hardening, instant glue is chemically identical to acrylic glass and is therefore able to create a very strong, transparent bond. Mounting standard LEDs requires holes to be drilled into the material, which can be a time-consuming and error-prone process, and should be undertaken with care. One major problem for both FTIR and IR systems is their sensitivity to ambient IR light from the external environment. This can be mitigated by adding a small electronic circuit to the set-up which supplies short high-current pulses instead of a continuous low current. The pulse current is usually set high enough such that under sustained operation, the LEDs would be likely to suffer permanent damage after a few seconds. Typically, these pulses are given a duration of between a hundred microseconds and a few milliseconds. The high current level, which is possible during the short pulses, results in a much higher light output. The pulse duration and the following cool down period should be kept as close to the manufacturer's specification as possible to prevent overheating of the LEDs. As modern computers are usually not equipped with the hardware or software to undertake such real-time control tasks, we suggest using a simple microcontroller (e.g., PIC or AVR) or the venerable 555 timer for pulse generation. A second-level switching element is also necessary, to handle the high currents which flow through the LEDs. Field-effect transistors (FETs), such as the IRF512 logic-level FET, are particularly easy to integrate with logic circuits and we suggest using these as second-level switches. A final precaution against LED damage is an ordinary fuse. A fuse with a *lower* rating than the expected pulse current should be inserted in series with the LEDs. Although more current will flow through the fuse than it is rated for, it is unlikely to blow during pulsed operation. Pulsing the LEDs significantly increases total light output, but this in itself does not produce enough contrast with ambient light levels. Instead, the pulses need to be synchronized with the camera in such a way that: (1) one pulse is emitted for each camera frame, and (2) each pulse's duration is equivalent to the camera's exposure time. As the LEDs are usually brighter by approximately one order of magnitude during the pulse, the contrast ratio with respect to environmental light is also significantly higher. If the camera exposure time is longer than a single pulse, stray light from the environment is accumulated during the cool down period between pulses, decreasing the contrast ratio. However, in the continuous mode, the brightness of the background is approximately 160 (when the LED is displayed with a maximum brightness (255 in 8-bit mode)), whereas in the pulsed mode, the background values are approximately 20, an eight-fold difference. To realise the pulsed operation mode, the camera needs to have configurable trigger output and exposure duration. These are standard features incorporated in almost all industrial-grade cameras. Some camera models even allow the generation of the entire control pulse with the trigger output, thereby reducing the external circuitry to 2 components (FET and fuse). For illustrative purposes we can consider how to calculate the correct pulse/exposure duration for a specific camera and LED combination (Pointgrey Firefly MV and Osram SFH4250 LEDs)[1]. If we assume a frame rate of $f = 60Hz$ then one full pulse/cool down cycle must have a duration of $D_{max} = \frac{1}{f} = 16.67ms$. If we are operating the LEDs at a voltage of 2.4 V (12 V divided by 5 LEDs) then the current is 1 A. We now have to calculate the total cycle duration, based on the duty cycle for each curve and the allowed pulse duration at a current of 1 A. For example, at a duty cycle of 3.3%, the pulse duration is approximately $t_P = 120\mu s$ for a total cycle duration of $D = 3.6ms$. At a ratio of 1% with a pulse duration of $t_P = 250\mu s$, the total duration already rises to $D = 25ms > D_{max}$, which is too long. We must therefore select a duty cycle of 2%, resulting in a pulse duration of $t_P = 200\mu s$ with a total duration of $D = 10ms$, which still offers a comfortable safety margin. Of course, the camera must be able to provide such short exposure times (as is the case for the Pointgrey Firefly MV). More information can be found in [55].

Cameras, Lenses, Filters and Projectors

Cameras FTIR and DI rely on cameras to detect fingers touching the surface. To create a functional surface a camera set-up must be found that is capable of sensing light in the near-IR spectrum; this must be coupled with a configuration of special filters that are designed to cut off interference from visible light. Although this can be challenging, the correct choice of camera and filter is essential to gaining the high camera signal quality required of a responsive multi-touch surface. Camera sensors that are capable of detecting IR light are required; however, the sensitivity of CMOS/CCD image sensors to infrared light varies considerably. When choosing a camera it is important to find out which sensor is used and determine (from the datasheet) its sensitivity to specific wavelengths of IR light. In many cases illuminators that have a

[1] `http://www.osram-os.com`

wavelength of 880 nm are used. For low cost initial prototypes, a USB web camera such as the Philips SPC900NC which uses a Sony CCD image sensor (type: ICX098BQ) is ideal. Web cameras often contain an infrared filter to block ambient infrared light. This filter layer must be removed. In some cases it is detachable, although often it is either glued on to the lens or applied as a coating on the camera sensor itself. The Philips camera, for example, has an infrared blocking filter glued onto the lens; therefore it is necessary to replace the original lens. Whilst high-end consumer USB cameras are capable of transmitting images of VGA resolution (640×480 pixels) at high frame rates, they often introduce significant latency. Any latency will reduce the responsiveness of the multi-touch interface. Therefore, FireWire based cameras are generally preferred, e.g. the Unibrain Fire-i board colour camera. This camera uses the same sensor (Sony ICX098BQ) as the Philips web camera but has a much lower latency. Depending on the size of the display and the projected image, cameras should normally be run at VGA resolution or higher (so as to achieve a reasonable precision) and smooth interaction requires a frame rate that is at least 30 fps. Because the camera only needs to be sensitive to infrared illuminated objects, it is advisable to mount an IR band pass filter to prevent interference from light in the visible spectrum (for example, from an image projected on your multi-touch surface). For optimal performance this should be a (relatively expensive) band pass filter which blocks all light other than the IR wavelength of the LEDs you are using; an alternative (cheaper) solution is to use an overexposed developed negative which acts as a (less specific) IR band pass filter.

Lenses, Exposure & Gain Once a camera has been chosen it must also be correctly configured so as to provide a highly sensitive camera image at a low latency. The *exposure time* controls how long a camera's shutter is held open for and thus how much light reaches its sensor. Setting the exposure appropriately is important for high quality tracking as even although a longer exposure time increases the camera's sensitivity, it can negatively impact upon the camera's frame rate. The camera's *gain* brightens images and increases contrast, but too much gain can lead to unwanted noise in an image. Another important choice is the type of camera lens. Integrating a wide-angle lens in a system allows smaller distances between the camera and the surface. However, lens correction and image rectification, waste pixels and so reduce tracking accuracy (especially toward the edge of the camera image).

Projectors Rear projection is commonly used to present the actual image upon the surface; but a number of factors must be considered when deciding upon an appropriate projector for a multi-touch surface. One important factor is the required display resolution. The necessary projection resolution is strongly application dependant, however, a resolution of at least 1024×768 pixels (XGA) is usually sufficient. Additionally, when choosing a type of projector – usually Digital Light Processing (DLP) or Liquid Crystal Display (LCD) – it is important to consider both the contrast ratio and the brightness (in lumens). Rear projection generally means that a lower brightness can be tolerated. In most cases, standard office projectors are not appropriate because of their long throw (the distance between the projector and projection surface required to produce a clear focused image). It is possible to use mirrors to reduce this distance, but this usually reduces the quality of the image and significantly complicates the physical design. Where necessary, a front surface mirror should be used to remove the double projection (ghosting) that can occur due to reflections from the glass front of a conventional mirror. In practice we have explored the suitability of several commercially available short throw projectors and recommend the 3M DMS 700 which is capable of projecting a screen size with a diagonal of 102 cm from a distance of 50 cm.

Compliant Surfaces and Projection Screens The FTIR set-up comprises a layer of polycarbonate augmented with a frame of infrared LEDs. When a finger is in contact with this layer, light from the LEDs which is internally reflecting within the polycarbonate is frustrated and produces a bright intensity region that can be tracked by a camera.

Compliant Layer A plain polycarbonate surface requires the user to apply significant pressure to achieve the frustrated light levels necessary for a responsive tracking. The use of a compliant surface can overcome this problem. Applying an additional layer on top of the polycarbonate material can greatly improve the sensitivity of the surface. These compliant surfaces are typically composed of a soft and transparent material. Figure 5.6 (left) highlights the relevant layers of a commonly used composition. When pressure is applied on the surface, the coupling of the diffuse layer and the polycarbonate surface triggers the FTIR effect. Use of the correct material for a compliant surface is critical as different materials can give rise to two common problems: (i) a sufficiently

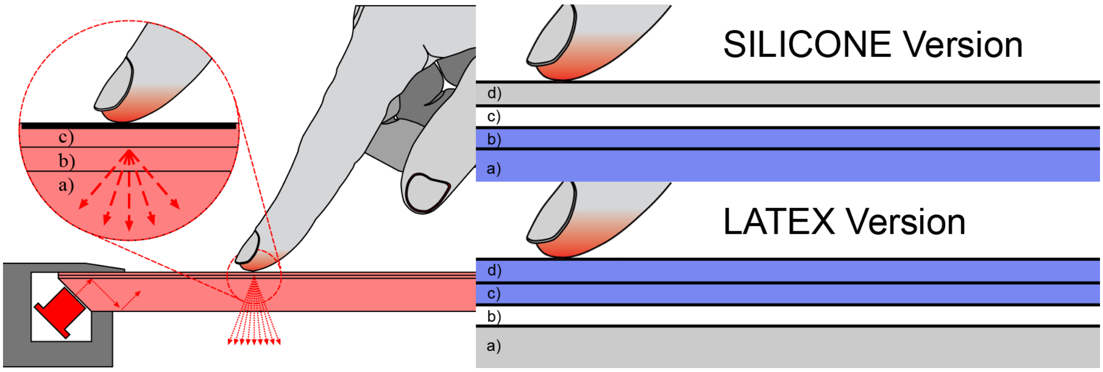

Figure 5.6: The three layers needed to track the finger touches: the polycarbonate plate (a) is covered with a compliant surface layer (b) and a diffuse layer (c) on top (left). Silicone compliant layer: the gap (c) is between the projection surface (d) and the combined silicone (b) polycarbonate (a) layer (right top). Latex compliant layer: the projection (d) and the latex layer (c) must be combined; the gap (b) is between these two and the polycarbonate plate (right bottom).

strong contact is not made with the FTIR layer (see Figure 5.1.3.1 (b)); or (ii) the material sticks to the surface, constantly triggering the FTIR effect even after a finger has been removed (see Figure 5.1.3.1 (d)). In our experiments the best results for the compliant surface were achieved with SORTA-Clear 40[1] and ELASTOSIL RT 60[2] silicone, both materials being relatively hard (Hardness Shore $A >= 40$), non tacky and very clear. Once hardened, both silicone layers can easily be removed from, and re-attached to, the polycarbonate surface. However, using silicone as a compliant surface poses a construction problem as the material comes as a gel, which must be poured evenly over the surface (a relatively difficult and messy task). ELASTOSIL RT 601 is less viscous and hence easier to pour, resulting in fewer air bubbles in the vulcanized layer. As an alternative to silicone, we found that a thin layer of latex also works well. This also has the significant advantage of not having to be poured, reducing the construction time for the combined layer significantly. Furthermore, latex is easier to handle, cheaper to produce, and more readily accessible as an off-the-shelf component. The order in which the compliant surface is combined with the other projection and polycarbonate layers is important in creating a functional surface; this varies depending on the material

[1] http://tb.smodev.com/tb/uploads/SORTA_CLEAR_40_32707.pdf

[2] http://www.wacker.com/internet/webcache/en_US/PTM/TM/Elastosil/Elastosil_RT_Addition/ELASTOSIL_RT_601.pdf

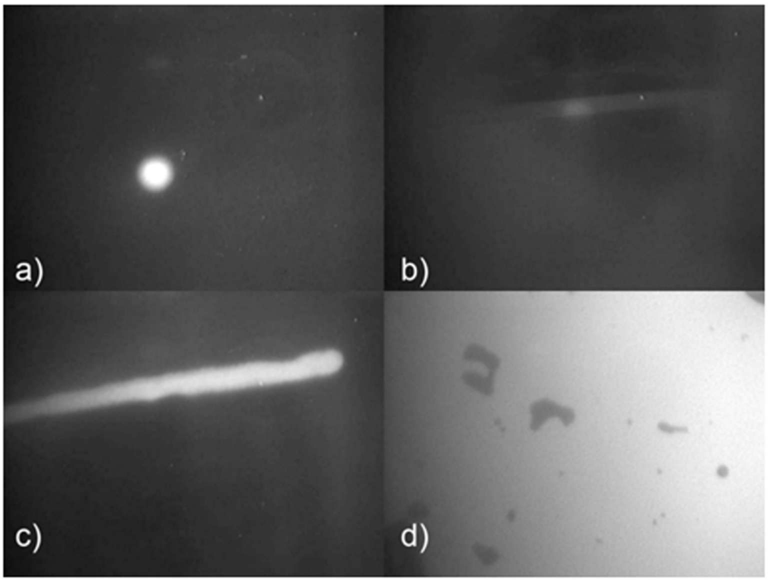

Figure 5.7: (a) Rigid PVC (backlit) (b) Rosco translucent [4] (c) Sihl polyester film 100 at, and (d) HP backlit UV. Using the wrong combination of materials can results in two main problems: (b) either the FTIR effect is not strong enough; or (d) the layers stick together.

used. Latex must be combined with the projection layer, with an air gap between the latex and the polycarbonate base plate. In the silicone version we have exactly the opposite requirements. Figures 5.6 (right figure) show this difference between the latex and silicone layer construction. More information can be found in [78].

Projection Layer As mentioned in the previous section, the configuration of surface layers varies with the choice of compliant surface material. Depending on whether silicone or latex is used a different projection screen must be chosen. The main factors to base this choice upon are that an air gap should be achievable between the two layers, and that when the screen chosen is pressed against the compliant surface the FTIR effect is triggered. Not all materials meet these requirements. Figure shows different results for projection materials on top of silicone.

Figure 5.1.3.1 (a) shows an optimal result for FTIR with a high contrast touch point. Materials that resulted in too dark touch points (b) or showed permanent traces on the silicone as well as materials that completely stuck to the silicone (c) are not suitable for

FTIR. Rigid PVC and tracing paper appear to be a good solution in combination with silicone. They do not stick to the silicone but trigger the FTIR effect quite well. When using latex, we found HP Colorlucent Backlit UV (a material originally designed for use in backlit signs) to be a good choice. Similar to rear-projection screens, it yields a diffuse image without any hotspots from the projector, making it a good rear-projection surface. Because of its glossy backside, it cannot be used with the silicone, as it adheres to the silicone as shown in Figure 5.1.3.1 (d). Rosco screens can also be combined with latex.

5.1.3.2 LCD Enabled MT Surfaces with Optical-Based Sensors

The use of an LCD monitor to display an image on a surface affords several key advantages, over projector-based systems. Generally, LCD monitors provide a higher display resolution than projectors (often for a lower price). For instance, a screen with full 1080p HD resolution will cost several thousands of dollars less than a projector with a similar pixel output. Additionally, the slim profile of LCD monitors makes them easy to house. This is especially important for those wishing to embed a multi-touch surface into the structure of a tabletop. Lastly, unlike, LCD screens do not have issues with key-stoning and throw-distance. These unique properties of LCDs make them a compelling option for multi-touch interfaces, especially for applications that demand a high degree of visual fidelity and resolution. There are however challenges that need to be overcome for the multi-touch developer wishing to utilize an LCD screen as a display technology. Firstly, it is imperative to have knowledge of how LCD technology works and be familiar with their manufacturing and assembly. The first issue that must be understood is that each pixel of an LCD monitor is comprised of three electronically controlled filters (red, green, and blue) which modulate over a backlight to emit a desired color. Essentially, the LCD glass panel is transparent when no current is running through the screen. Next, on the front and back side of the glass panel are criss-crossing polarizing filters. The polarizing filters give an LCD its black appearance since their opposing orientation blocks visible light. However, polarizing filters do not polarize light within the IR spectrum. So while a LCD panel looks opaque to our eyes, IR light can be transmitted through the screen unperturbed. This concept is crucial for the use of LCD screens in optical multi-touch systems. The next part of the LCD assembly is the back-light and filter-chain. A backlight is necessary in order to illuminate the LCD pixels. The backlight for monitors that are less than 23" in size consist of a long thin fluorescent light bulb, which lines the length of the monitor. Attached to the bulb is an acrylic sheet (called the light guide), which has a honey-comb pattern of white dots. Based on the principle of total-internal reflection, the light from the fluorescent tube travels inside the acrylic sheet until it reflects off one of these white dots. This method for back-lighting allows for thin displays. For LCD monitors that are 27" and larger, the acrylic-guide method is replaced by a rail of lights which are placed behind the screen to provide the backlighting. However, because of the polarizing filters and the method in which the crystals distort and filter light, having only a back-light is somewhat ineffective for illuminating the display. This can be understood if one imagines adjusting their laptop screen in order to achieve the best viewing angle, which is orthogonal to their line of sight. Tilt the laptop screen too much and the display image loses much of its color and appearance. To improve the lighting conditions of the display, LCD manufacturers include a layering of several different filters, which modulate and affect the backlight in various ways. The most common filters include: (1) Diffuser: this filter diffuses the backlight to disperse in every direction. (2) Brightness Enhancement Film (BEF): this filter can magnify light with a shorter focal length in different directions. This filter is used to disperse light in 180 degrees. (3) White Reflector: this is an opaque white filter, which reflects any light that may have escaped the filters. These are the three basic filters for LCD monitors; however some manufacturers may use additional filters to improve their product quality; for instance, using different types of diffusers, or more than one BEF to improve the viewing angle. Of these filters, the only one that impedes IR light, and is therefore of concern when developing optical multi-touch surfaces is the last white opaque filter. This filter is totally white and hence needs to be removed; the rest can and should remain to keep optimal viewing performance. There are two broad methods, which so far have been successful for creating interactive LCD surfaces with optical sensing. The first, and easiest, is the side-illuminated method where IR LEDs are installed around the bezel of the LCD. The LEDs shine IR light across the top surface of the screen; when a finger touches the LCD screen, light reflects off the finger and traverses through the monitor, which is then captured by an IR sensitive camera. The illumination hardware required for this approach is very similar to the FTIR method. Therefore it is often possible to install an FTIR panel on top of an LCD screen and then remove the acrylic; keeping the LEDs intact. With this method, it is recommended to identify IR LEDs with a small package (3mm or SMD) and with a small viewing angle (typical angles for LEDs are 30 degrees, but shorter angles are available at around 15-18 degrees). Choosing a smaller angle will focus the more of the light to across the screen. Finally, emerging IR Laser LEDs promise to be the ideal choice for this method because these light sources ensure that the IR light beam is small and hence focused as a blanket over the surface [149]. The second method for enabling multi-touch with LCD screens is to create a matrix of IR transceivers behind the LCD panel as described in [92]. Each transceiver consists of an IR emitter, and an IR detector. The emitter pulses IR light at a certain frequency, which the sensor can detect (similar in theory to IR remote controls except here the light is not encoded to pulse information). When a finger or an object touches the screen, the finger reflects back the light, which is detected by the sensor. By creating a matrix that consists of many of these transceivers, it is possible to cover the entire surface area of the LCD screen. The number of transceivers, their size and pitch (distance between sensors) determines the accuracy and resolution of the touch surface. This approach allows for a thin form factor display as the sensors can be placed directly against the surface, in contrast to the cameras used in other approaches. However this approach is not simple to construct and requires expert knowledge of electronics, circuit design, and digital-signal processing (DSP). Additionally the approach is not scalable, as larger surfaces require more sensors, which increases cost and latency. More information can be found in [54; 149].

5.1.3.3 Software

Once the hardware is in place the next major challenge is the software processing of the camera image to interpret the users interactions. To achieve this, a pipeline of image processing operators that transform a camera image into user interface events must be set up.

FTIR Tracking Pipeline Figure 5.8 shows the typical imaging pipeline of an FTIR set-up. In an initial step, images captured by a camera are preprocessed. Preprocessing consists of first rectifying the camera image so that the image pixels and display pixels match up. This transformation can also be done during postprocessing, where only

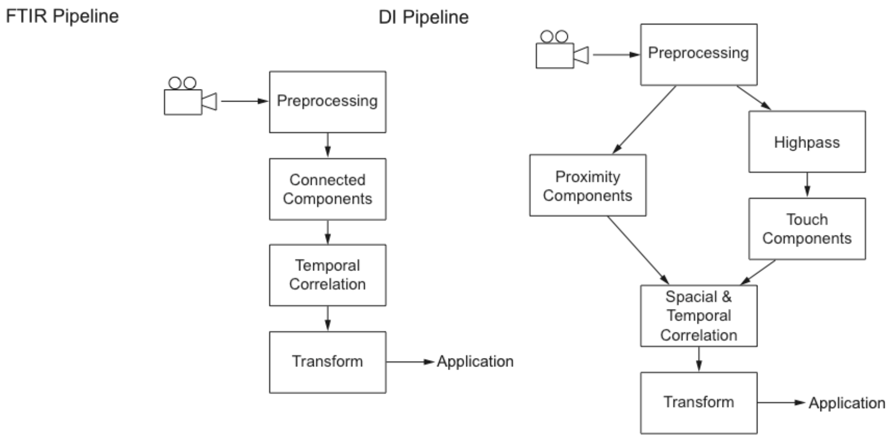

Figure 5.8: FTIR Tracking Pipeline (left). DI Tracking Pipeline (right).

the actual blob coordinates have to be calculated. Done at the beginning, it has the advantage that intermediate images can be displayed on the surface without distortion.

Following the rectification, history subtraction is used to remove any unchanging parts. If the camera image is noisy, an appropriate noise reduction filter (opening/closing, lowpass or similar) can be added to the pipeline.

Simple threshold-based segmentation using a connected components algorithm (described e.g. in [80]) finds bright regions - so-called 'blobs' - in the pre-processed image. These are the areas where something is touching the surface. Principal Components Analysis (PCA) can be used to calculate statistical data (size, eccentricity, etc.) for the blobs. Using this data, it is possible to distinguish touches with fingers from other objects and from noise.

Post-processing involves finding corresponding touches in different camera frames (temporal correlation). Finding an algorithm that consistently detects the movements of touches from one frame to another turns out to be surprisingly hard. A simple greedy algorithm that goes through all new blobs and picks the closest old blob for each one is less than optimal. Blobs that split and merge confuse it. Also, the greedy algorithm often incorrectly exchanges blobs when many fingers are quickly moving over the surface.

There are several improvements that can be made over the naive algorithm:

1. A robust solution first calculates all distances between blob pairs and sorts the

results by distance. In a second step, the blob pairs with the closest distances are correlated. The second step is repeated until all new blobs have been accounted for.

2. Dead reckoning can be used to extrapolate the position of the old blob using its previous speed before calculating the distance.

3. The distance function need not be the Euclidian distance between the blob centers. Statistical data from the PCA performed earlier (e.g. size and eccentricity of the blobs) can be factored into the distance function.

DI Tracking Pipeline DI tracking is a more complex process but allows for proximity as well as touch to be sensed. DI Touch detection exploits the fact that objects at a distance from the surface appear blurred. reacTable [105] does this by adaptive thresholding based on the curvature of the luminance surface (see [44] for a detailed description of the algorithm). The multimedia platform *libavg*[1] used in the *c-base MTC* pioneered the use of a high-pass filter to achieve the same effect. Note that a full high-pass is computationally expensive, so *libavg* uses fragment shaders to implement the filter.

As can be seen, the image pipeline is split and the end the connected components algorithm is run twice, once each for touch and once for proximity sensing. Touch sensing involves an additional high-pass filter to isolate areas very close to the surface. After the regions have been found, touch and proximity information can be correlated. With appropriate thresholds, hand (proximity) blobs reliably enclose the finger (touch) blobs, so a correlation is easy to establish. Additionally, the vector from hand center to finger center is a very good approximation of the direction the finger is pointing to. The bottom right image in Figure 5.8 shows the result of this process: Fingers touching the surface have been identified and associated with hands.

Interface Considerations The tracking pipeline provides higher-level software layers with information about finger and hand positions. TUIO [106] uses Open Sound Control over UDP to transmit this information in a format which can be interpreted easily by a wide variety of tools and languages. By default Touchlib and many other

[1] `http://www.libavg.de/`

libraries come with a wrapper which sends TUIO events over the commonly used Open-Sound Control[1] protocol. For many modern programming languages such as C#, Adobe Flash (Actionscript 3), Java, Max/DSP, Processing, Pure Data, Python and Visual Basic, OSC libraries are available. When using Flash it is required to convert UDP packages to TCP. This can be done by using the tool Flosc which acts as a proxy. Work is in progress to provide higher-level interfaces (*libavg*, *libtisch*[2]. *libavg* which includes event processing that correlates touches to a hierarchy of on-screen widgets[3]). This corresponds to the mouse event handling that window systems provide and hence affords the basis for robust implementation of classical GUI widgets like buttons and scrollbars. Both libraries support emerging gesture standards that allow for dragging, rotating and scaling of GUI elements. When an application uses the OSC protocol, it is only be able to receive events containing properties of the detected blobs. It is not possible to adjust the settings of Touchlib from the application. However, since OSC uses the UDP network protocol to transfer data it makes it possible to create a set-up in which a dedicated system provides blob tracking and transfers the data to another system that provides the visualization. At higher levels, window-system-like event processing, classical GUI widgets (buttons etc.) and emerging gesture standards (dragging, rotating and scaling elements, for instance) are supported by some libraries. The most commonly used tracking libraries that are currently available are touchlib, tbeta, libavg, multi-touch lib T-Labs, OpenFTIR, VVVV, and OpenTouch. For a more detailed overview please refer to [205].

5.2 Multi-touch Interaction with Virtual Globes

While in the last section the technical realization of multi-touch surfaces is describe we now focus how such board sized multi-touch interfaces can be used to act as user interfaces for spatial information. Virtual globes have progressed from little-known technology to broadly popular software in only a few years. We investigated this phenomenon through a survey and discovered that, while virtual globes are en vogue, their use is restricted to a small set of tasks so simple that they do not involve any spatial thinking. Spatial thinking requires that users ask “what is where” and “why”; the

[1] OSC `http://www.cnmat.berkeley.edu/OpenSoundControl/`
[2] `http://tisch.sourceforge.net/`
[3] `https://www.libavg.de/wiki/index.php/Event_Handling`

most common virtual globe tasks only include the "what". Based on the results of this survey, we have developed a multi-touch virtual globe derived from an adapted virtual globe paradigm designed to widen the potential uses of the technology by helping its users to inquire about both the "what is where" and "why" of spatial distributions. We do not seek to provide users with full GIS (geographic information system) functionality, but rather we aim to facilitate the asking and answering of simple "why" questions about general topics that appeal to a wide virtual globe user base.

5.2.1 Introduction & Motivation

There exists myriad evidence of the dramatic rise in popularity of virtual globes. Google Earth, the most ubiquitous virtual globe, was downloaded over 100 million times in its first 15 months[1] of release. Fromer U.S. President George W. Bush has said that he uses Google Earth to look at his Texas ranch[2]. Moreover, the phenomenon has even inspired a Nature news feature [33]. The Nature article notes an important dichotomy between the features employed by the casual user of Google Earth and those used by the scientific audience. The author writes "to the casual user ... the appeal of Google Earth is the ease with which you can zoom from space right down to the street level" while the attraction of scientists and enthusiasts to the program lies in the fact that it is "an easy way into GIS software" (p. 776). While virtual globes' use as an entryway into the world of GIS cannot be understated, this dichotomy raises doubts about the ground-breaking nature of the technology on the large group of people who do not make the jump to advanced GIS packages. The results of a survey, discussed later in the paper, elicit further concerns about the superficiality of tasks performed by the average virtual globe user. As defined in the recently published National Research Council Report, Learning to Think Spatially, spatial thinking (in the geospatial domain) is a "dynamic process that allows us to describe, explain, and predict the structure and functions of objects and their relationships in real and imagined spatial worlds." (p. 33). A significant part of the spatial thinking process involves generation of hypotheses, pattern predictions, and tests of hypotheses. Essentially, when thinking spatially, individuals observe what patterns exist in the environment and seek to provide explanations for these patterns. In short, these individuals ask "what is where"and "why". GIS is

[1] http://www.google.com/press/pressrel/earth_election_guide.html
[2] http://www.guardian.co.uk/technology/2006/oct/27/news.usnews

increasingly heralded as the most probable support system for facilitating the spatial thinking process as it allows for spatialization [61] of non-spatial datasets. The spatial representation of data permits the individual to ask “why” questions i.e. why certain patterns or relationships exist in or between certain places questions that are difficult to formulate when the same data is experienced in a different format (e.g. a spreadsheet). With expertise in traditional GIS technology, these patterns and processes can be further explored using spatial statistics and other advanced operations, analyses certainly beyond the knowledge of the everyday Google Earth user. As virtual globe technologies become increasingly pervasive, much hope surrounds their capacity to potentially enhance spatial thinking ability among both K-12 students and non-expert users. However, as demonstrated in the results of a survey on the uses of virtual globes (see below), most individuals use these technologies simply for observational purposes, and little to no spatial thinking actually occurs. In other words, the majority of individuals seem to use these technologies to observe the “what” of spatial data (e.g. the location of their home or business and where it is in relation to other prominent geographic features), but moving beyond pure observation to questioning why certain patterns exist in the landscape proves out of reach to the casual user. As typical virtual globe technologies are not coupled with specific datasets or feature sets, and adding data to the existing software involves a certain level of expertise, the majority of individuals does not have access to the information or tools they need to ask the “why” questions. Therefore, the technologies do not, in their current form, typically support the spatial thinking process. Importantly, this research is not an effort to incorporate an easy-to-use GIS into a virtual globe software package. Other projects such as ArcGIS Explorer[1],Nasa World Wind and Google Earth itself have tackled this problem to at least a small extent. Our aim is entirely different. Rather than providing the user with advanced GIS functionality (e.g. spatial join, cluster analysis, buffers) to answer spatial thinking questions, our system facilitates the asking and answering of simple “why” questions, e.g. Why does this spatial feature display this value? What is the relationship between these two features? Our prototype enables this facilitation by demonstrating enhancements in two key areas: data type and interface. We introduce a new simple spatial thinking-oriented virtual globe data type called Explicitly Explanatory Spatial Data

[1] http://www.esri.com/arcgisexplorer/

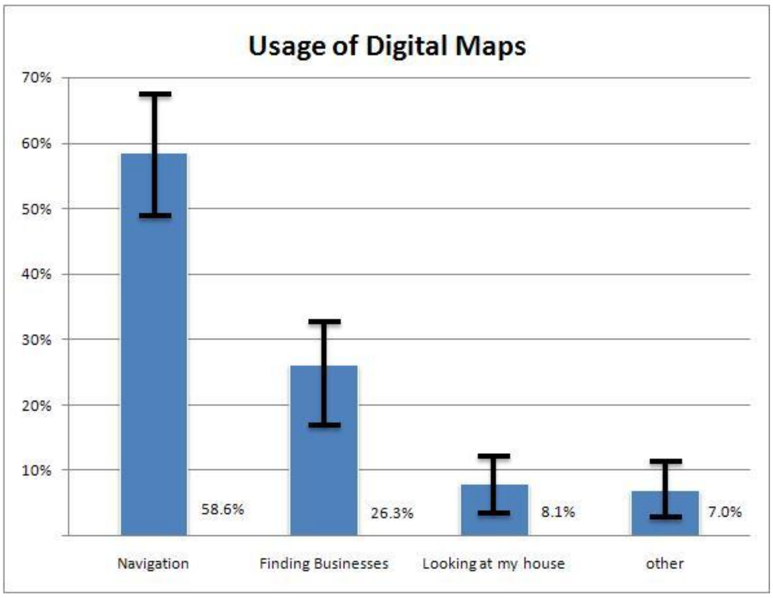

Figure 5.9: Usage of Digital Maps.

(EESD), which contains both a standard spatial layer and a new explicitly explanatory layer designed specifically to answer "why" questions. Two test case data sets are presented. The first is based on our previous WikEar[207] and Minotour [86] projects, which use Wikipedia to generate narratives between geotagged Wikipedia articles. The second uses a prototype of GeoSR [87], a new semantic relatedness-based system that is backed by a Wikipedia-based knowledge repository. The semantic relatedness literature originates in computational linguistics and seeks to define a single number to quantity the degree to which any two concepts are related [7].

5.2.2 Survey

A user survey was conducted to investigate the usage and user needs of virtual globes. The study included 120 participants: 60 female and 60 male. They were randomly selected in a pedestrian area in Mnster, Germany and had a mean age of 34.2 years (SD=8.7). The length of the survey was about 5 to 8 minutes, during which time each participant was asked ten questions about her or his knowledge and use of virtual globes, as well as digital maps. We also asked about digital maps to investigate the usage similarities and differences between the two geovisualization mediums. First, the participants were asked if they were aware of digital maps; 89.2% (± 5.5%) of the

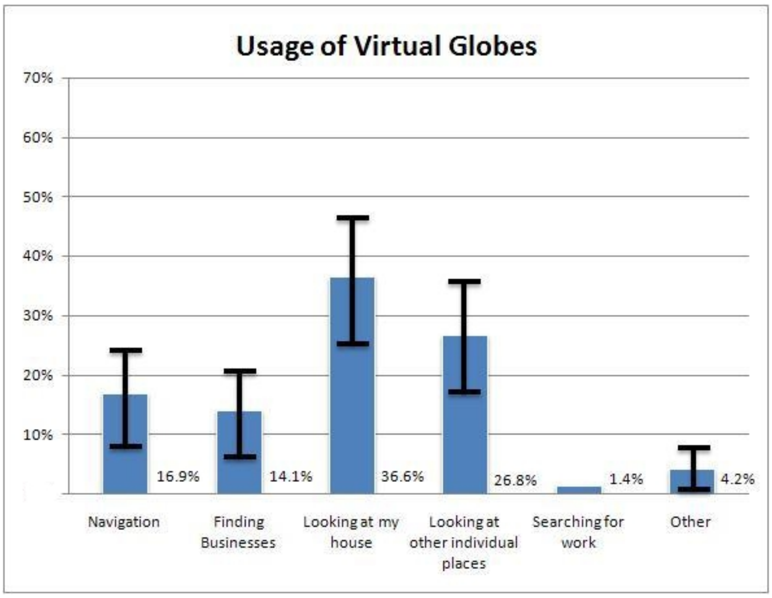

Figure 5.10: Usage of Virtual Globes.

participants answered in the affirmative, of whom 92.5% (82.5% ± 6.8% overall) use digital maps more than 5 times per month. When asked identical questions about virtual globes, 67.5% (± 8.3%) said that they were aware of virtual globes while 59.2% overall (± 8.8%) said they used them more than 5 times per month. We then asked users about the motivations behind their virtual globe and digital map use. Around half (53.4% ± 11.6%) said they used virtual globes for either looking at their own house or other individual places (e.g. a neighbor's house, their hotel from their last vacation, the city center). The second most common uses of virtual globes were navigation (16.9% ± 8.7%) and locating businesses (14.1% ± 8.1%). More esoteric responses, such as that of a roofer who said he used Google Earth to find roofs that needed repair, rounded out the respondents' uses. The distribution of digital map use was quite different than that of their virtual globe cousins, with over 50% of respondents saying that they used digital maps for navigation. More details on both results can be seen in figures 2 and 3. Finally, we asked respondents to compare and contrast the advantages of virtual globes and maps. They answered that the main advantages of digital maps were easy navigation and global map coverage. In contrast, the main advantages of virtual globes over digital maps were the ability to view high resolution satellite images and aerial photography overlaid with additional information such as geotagged Wikipedia articles and Panoramio[1] photos in a 3D environment.

5.2.3 Design Requirements

As noted in the introduction, the central design conclusion of the virtual globe survey is that the majority of tasks employed by virtual globe users are simplistic and do not require spatial thinking. As spatial thinking involves both noticing patterns in the landscape (whether it be a real or represented environment), and questioning the evolution of those patterns, simple observational activities do not constitute spatial thinking. Similarly, with navigation and business location (in this case), virtual globes are employed simply to answer the "what" question, as well as "where" certain features are in relation to one another. There is no "why" in the picture. It is also important to draw conclusions "albeit less firm ones" from trends that can be found in the unstructured and unsolicited responses from survey participants. First and foremost, users like the general idea of displaying the Earth in three dimensions, as they indicated they enjoyed viewing the Earth as it truly is. However, they noted that the interaction with a 3D environment was difficult and many expressed a desire for easier-to-control interfaces [29]. Finally, and most critically, over the half of the users indicated that they felt that virtual globes could be more useful to them if they only could figure out more tasks to perform with them (besides those they indicated). This can be interpreted as a desire to engage in more advanced tasks, likely those involving some spatial thinking.

5.2.4 Conceptual Design

Following the results of our survey, we developed a new virtual globe prototype designed to widen the potential uses of the technology by allowing users to spatially inquire about both "what" and "why".

There were two key challenges in developing the data type for our improved virtual globe prototype. The first was to design a general data structure that would enable users to both ask and answer spatial thinking questions. The second was to appeal to the thematic interests of the broad virtual globe user base. The former challenge is the topic of the first subsection and the latter is discussed in the second.

[1] `http://www.panoramio.com/`

Data There were two key challenges in developing the data type for our improved virtual globe prototype. The first was to design a general data structure that would enable users to both ask and answer spatial thinking questions. The second was to appeal to the thematic interests of the broad virtual globe user base. The former challenge is the topic of the first subsection and the latter is discussed in the second.

A Framework To Facilitate Answering "Why" Questions of Data GIS software for years has enabled users to engage in a large variety of advanced spatial thinking tasks. However, the design goal for this research is to facilitate simple versions of these tasks using intuitive paradigms in virtual globes. Our solution on the data side is the Explicitly Explanatory Spatial Data (EESD) type. Each EESD set is defined by two layers. The first layer is the standard spatial data layer that has been in use since the first GIS around 40 years ago. It can contain raster cells, points, polylines, polygons, or any other feature type that can be displayed on a virtual globe. This layer, in an abstract sense at least, also contains attribute data for the features. The second layer, the explicitly explanatory layer, holds the innovation. This layer contains explicit explanations for the attribute values and/or relationships present in the spatial data layer. It is hypothesized that explanation of these two properties of a spatial data layer, corresponding to the "objects" and "relationships" noted in the definition of spatial thinking found in the introduction, will best facilitate basic spatial thinking tasks. This layer must make it a trivial matter for the interface responding to a "why" query from the user to return an explanation.

We have implemented two examples of the EESD sets, the WikEar [207] data set, which is derived from our Minotour [86] work, and the data set generated by an early version of our GeoSR [87] project. Both have a spatial data layer that is generated from the large number of hand-geotagged articles in the English version of Wikipedia. The former EESD set is of the type that contains explanatory information about spatial relationships while the latter is focused on explaining single data values (although users will likely identify implicitly explained patterns as well). Before detailing the prototype EESD sets, however, it is important to discuss certain properties of the Wikipedia knowledge repository. First and foremost, it is necessary to acknowledge concerns about the risks of using Wikipedia data. Denning et al. [48] codified these risks into concerns over accuracy, uncertain expertise, volatility, coverage, and sources. However, Giles [71] reported that Wikipedia is comparable to the Encyclopedia Britannica in terms of number of serious errors and only slightly worse than Britannica when it comes to “factual errors, omissions, or misleading statements”. Regardless, given the requirements of this research: a natural language knowledge repository with both an extensive and intensive coverage of world knowledge, Wikipedia is by far the best choice. With over 2 million articles in the English version (as of submission) and 14 other language editions with over 100,000 articles (all methods described here work with all Wikipedia languages), Wikipedia is the largest Encyclopedia to ever exist. For the purposes of this research, Wikipedia articles can be split up into 3 groups:

1. articles without a geotag,
2. articles with a geotag and
3. articles about purely temporal phenomena (i.e. the article on the year “1983” or the date “October 1”).

We call articles in the first group “non-spatial articles” and articles in the second “spatial articles”. The third group exists because purely temporal articles have very defined relationships encoded in their links with other articles; linking to a temporal article is nothing more than providing an explicit temporal reference to the article, something that can be useful in some contexts but amounts to enormous noise in this work. Finally, the concept of a Wikipedia “snippet” is critical to both EESD sets. In a general sense, a Wikipedia snippet is simply a paragraph of a Wikipedia article. These paragraphs are unique in natural language text knowledge repositories in that they are almost entirely independent of one another. In other words, snippets almost never contain unexplained or incomplete textual references to other snippets. This is a direct result of the encyclopedic writing style that is the Wikipedia norm, as well as the collaborative nature of Wikipedia, in which the median number of authors per article (as of 2006) in the English version was over seven [32]. We have found experimentally that the only context necessary for fully understanding the vast majority of snippets is the title of the article to which the snippet belongs. One can further increase understanding of independent snippets by providing the hierarchy of headings in which the snippet resides (i.e. for the United States article, there are 3 snippets under “History→Native Americans and European Settlers” as of November

26, 2007). Minotour generates cohesive stories from a Wikipedia knowledge repository using a data mining methodology derived from narrative theory. The WikEar dataset contains human-narrated versions of Minotour's stories in an attempt to simulate future text-to-speech technology. The stories begin at one Wikipedia article a, end at a Wikipedia article b, and contain s snippets, each of which belong to a Wikipedia article on a narrative-theory defined optimal path from a to b through the Wikipedia Article Graph (WAG). In the WAG, each article is a vertex and each directional link between articles is an edge. The variables a, b, and s are all user-defined. The primary test case for Minotour and WikEar is the generation of educational tourism narratives. In this context, spatial Wikipedia articles are used for a and b, while non-spatial articles provide the snippets for the body of the narrative. Critically, applied in this manner, Minotour narratives, by definition, explain a relationship between the spatial entities that articles a and b describe. As such, operating with a layer of the spatial references of Wikipedia articles, the narratives form an explicitly explanatory layer for the relationships between the points in the spatial layer. Looking at the spatial layer, a user can ask, "Why are these two spatial entities related?" and the system can easily respond with an answer. With tens of thousands of spatial articles in the English Wikipedia, users are able to ask this very simple and general spatial thinking question about almost anywhere in the world. This simplicity and generality fits in with other typical virtual globe data layers (i.e. satellite photography), but also allows for the explicit answering of "why" questions. An early prototype of GeoSR is the backbone of the second EESD layer. GeoSR is based on our novel ExploSR semantic relatedness (SR) measure, the first adapted to the context of data exploration. The goal of SR measures is to identify a value that summarizes the number of relationships between two entities as well as the strength of these relationships. By analyzing the Wikipedia Article Graph (WAG), ExploSR derives such values between the entities represented by Wikipedia articles. The key variables looked at by ExploSR when examining any two articles a and b are the myriad paths from a to b (and vice versa) in the WAG and the scaled weight of the links in those paths. Link weights are determined by a mixture of article length, number of out-links (outdegree) between the linked articles, text position of those links, and several Wikipedia-specific variables. In addition to being the first semantic relatedness measure designed for use in a data exploration context, ExploSR is the first measure to utilize the WAG and the first measure that can be visualized in a reference system (in this case, a geographic one. Further discussion of the benefits of the WAG for this type of semantic relatedness application is merited. First, the WAG is ideal for SR measures designed with data explanations in mind because a natural language explanation is built into every outputted measure (see below). Secondly, the WAG is replete with both classical relationships, i.e. is-a (hypernymy and hyponymy) and has-a (meronymy and holonymy), and non-classical relationships [146]. We have found qualitatively that these non-classical relationships such as "spoke-at", "ate-a", "wrote-about", "tool-he-uses-to-look-at-ranch" to be far more important than their more standard cousins when evaluating SR measures on articles representing entities that belong to a commonly-used reference system, such as spatial and temporal articles. Abstractly, ExploSR takes a Wikipedia article as an input and returns a single semantic relatedness value from a to all Wikipedia articles of type T. Users can then query GeoSR for an explanation of any value, and GeoSR will return the snippets containing the links that form the path between a and b in the WAG, where $b \in T$. Typically, T will be a set of articles that all belong to some semantic reference system [119] (i.e. spatial or temporal). There are many geographic applications of GeoSR, two of which are used in our GeoSR-derived prototype EESD data set. The first occurs when a is a non-spatial article and T equals the set of spatial articles. This will result in all spatial articles having a semantic relatedness value to a. The second occurs when T again equals the set of spatial articles, but a is also a spatial article. While similar, these applications differ significantly in that one result in measures of theme-to-spatial entity relationships while the other outputs spatial entity-to-spatial entity relationships. This is illustrated in figure 5.11.

Both applications, however, provide excellent EESD sets. In both, the spatial data layer is a spatial visualization of the GeoSR measure in which the spatial entities depicted are those about which there are Wikipedia articles. In our example (see figure 5.14), this layer is represented as a graduated symbol map, with each spatial Wikipedia article depicted as a point in a geographic reference system. Other visualizations are also possible with improved georeferencing of Wikipedia articles (for instance, referencing articles about spatial entities of sufficiently large extent to polygons rather than points). The size of the symbol is defined by the value of the semantic relatedness measure, with bigger symbols indicating more and/or stronger relationships. The explanatory layer, is built directly into the ExploSR system as described above. Applied

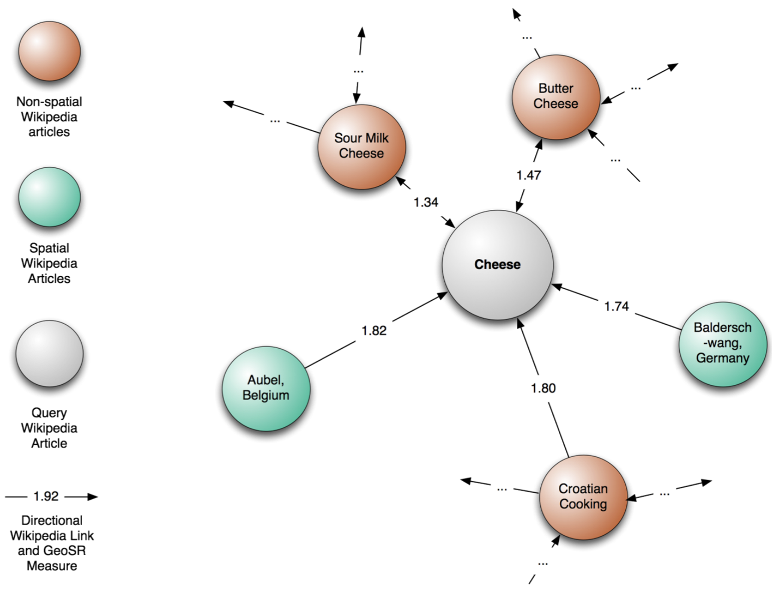

Figure 5.11: Generation of the EESD layer. Step one: Applying semantic relatedness values to the Wikipedia Article Graph based on a query entity (in this case "cheese"). Query entity must appear in Wikipedia. Step two: pick out the spatial articles.

in a geographic context, this amounts to every value visualized on the map having an explicit explanation found easily in the data set. This is illustrated in figure 5.12.

Similarly to the WikEar dataset, the GeoSR system generates data of broad, general appeal. Since a can be any article, the user is able to see how related all entities described by Wikipedia articles that are spatially referenced are to any entity in all of Wikipedia, from "multi-touch" to "George W. Bush" to "Rugby" to "Surfing". Importantly, both layers are easily applied to spatial subsets, or extents, of the globe. Using the measures on small extents will focus the graduated symbol visualization to allow maximum differentiation in relatedness in the region of study.

5.2.5 Basic Interaction Concepts

To benefit from the EESD layers, users need an intuitive way of interacting with them. To provide this intuitive interaction paradigm, we use the full advantages of our multi-touch surface. The basic spatial interaction tasks such as pan, rotate, zoom and tilt are implemented. For instance, the user can pan through the world with the flick of a finger or hand and use other multi-touch gestures to zoom, rotate, tilt or navigate.

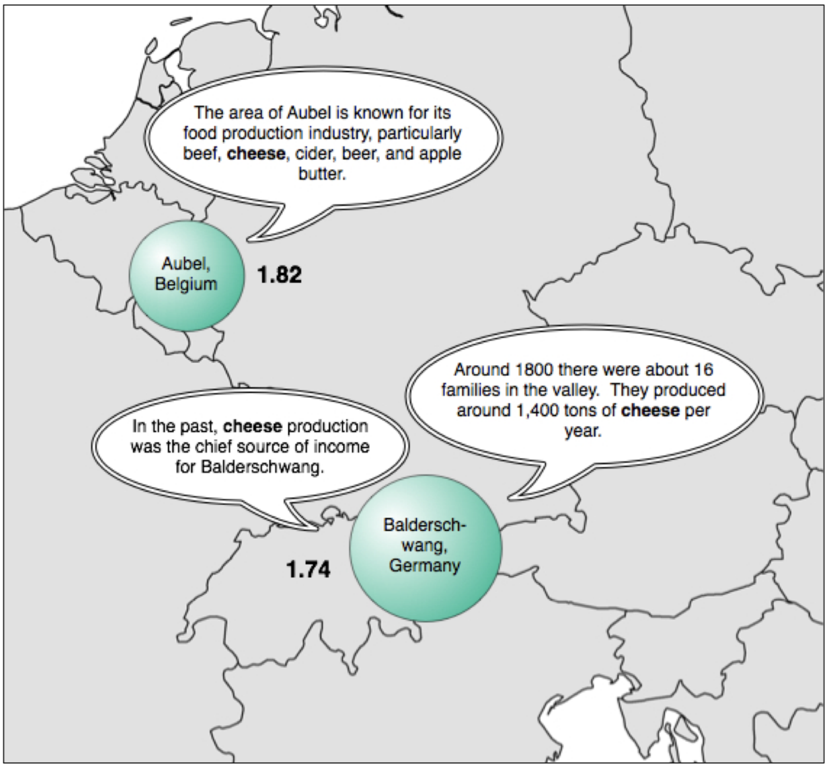

Figure 5.12: Explanation of EESD layer entities. Built into the measure is an explanatory component. Each value has a reason behind it: the links in the Wikipedia Article Graph that resulted in the value. These links can be presented along with their host paragraphs and the users will easily be able to see"Why" each value is as high or as low as it is. Reiterating, the why and the what is where are very intimately related.

"Click" and "double-click" are implemented with simple taps.

5.2.5.1 Interaction with the First (WikEar) EESD Layer

The interaction with the first EESD layer is straightforward. The user selects two spatial Wikipedia features by double-clicking (double-touching) Wikipedia icons (which indicate spatial Wikipedia articles) simultaneously with two fingers (see figure 5.13). The icon selected by the one hand is the start feature a and the other Wikipedia feature is the end feature b. The start feature, end feature and a line between them are highlighted and a story derived from Wikipedia (as described in the previous section) is read out to the user. Users can control the speed of playback by dragging their finger from the start point (in green) to the end point. By moving the finger from the end

Figure 5.13: Interaction with the first (Wikear) EESD Layer.

location to the start location story is derived by swapping the start and end point and the story is played back. By releasing her/his fingers from the multi-touch surface, a user can stop playback and can, for example, navigate to another place on the earth or request other information.

5.2.5.2 Interaction with the Second (ExploSR) EESD Layer

To interact with the second EESD layer, users must first define a region. This can be done by activating the “region definition mode” by touching a button and dragging the hand(s) or finger(s) over the multi-touch surface. After lifting her/his hand or hands from the multi-touch surface, a user sees a menu where she or he can select one or more different “themes” (which represent different articles a as input) for that region (in our prototype we have 25 pre-computed themes). By dragging a theme into the region (see figure 5.14, middle), users can explore semantic relatedness values for that “theme” in the region they selected (see figure 5.14, right). Clicking on a single symbol will provide the text-based explanation of the “why” of each value as described in the data section. Without dragging the “theme” into a predefined region, users can explore the relatedness of that “theme” at a global scale (see figure 5.15). This mode can be

Figure 5.14: Interaction with the second EESD Layer - Region Selection (left). Interaction with the second EESD Layer - Dropping a theme into a region (middle). Interaction with the second EESD Layer - Visualization of the result (right).

deactivated by disabling the EESD view.

5.2.6 Implementation

Both EESD layers operate from a significantly pre-processed version of the Wikipedia knowledge repository. The pre-processing takes as input one of the semi-regularly exported database backup dump files from Wikipedia. Currently only the English, German, and Spanish files are supported, but with the help of a translator it would be an easy matter to add support for any language version of Wikipedia. For the larger Wikipedias such as English and German, the size of these dump files is remarkable. The latest English dump file as of November 2007, for instance, was about 12.7GB of text. During the pre-processing stage, the dump file is parsed in a Java parsing engine to isolate article, snippet, link structure, spatial, and temporal information, which is then stored in a MySQL database in a variety of tables. A Java API to this database, which is named WikAPIdia is then used by the systems that generate both EESD layers. The API provides basic access to Wikipedia data as well as more advanced graph mining and spatiotemporal features, which are used by both Minotour and GeoSR. As noted above, our interface is rooted in a 1.8m x 2.2m FTIR-based multi-touch wall as described in the related work chapter. This wall consists of a 12mm thick acrylic plate, in which every four centimeters a hole for an infrared LED was drilled. The acrylic plate was mounted onto a wall and a wide-angle lens digital video camera (PointGrey Dragonfly2) equipped with a matching infrared band-pass filter was mounted orthogonally at a two-meter distance. As a projection screen, very inexpensive drafting paper was used. For

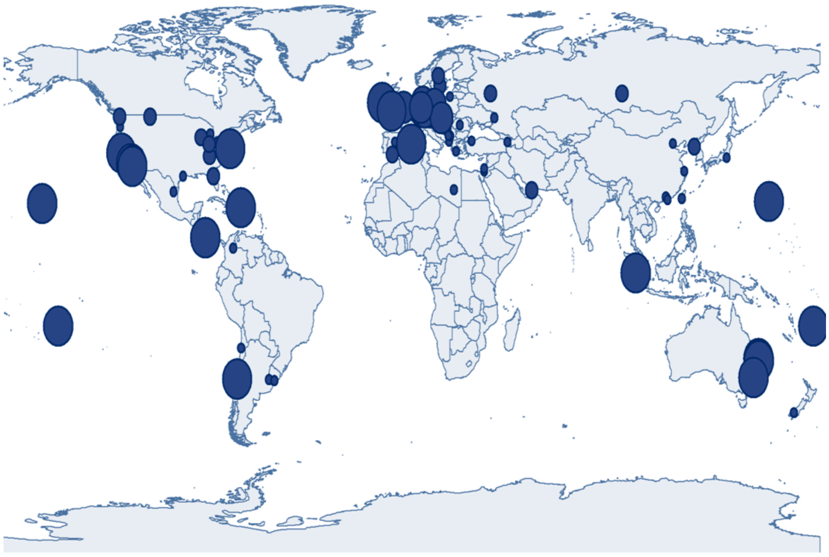

Figure 5.15: Visualization of the theme “surfing” (second EESD layer) on a global scale derived from the German Wikipedia.

the projector we used a Panasonic PT-AE1000E HD projector. To improve dragging operations, we placed a thin layer of silicon (Silka Clear 40) between the acrylic and the drafting paper. The Java-based Multi-Touch Library[1] developed at the Deutsche Telekom Laboratories and released under the GNU Public License was used for image processing. It contains a set of common algorithms designed to work with any multi-touch system such as routines to label connected components and track features. By using an application layer, it is easy to manipulate objects and transform (position, rotate, scale) them. The library also comes with a module for accessing cameras such as the PointGrey Dragonfly2. Our virtual globe is based on NASA's World Wind. The NASA World Wind visualization platform is open source and comes with a rich SDK for data set and interface customization, which we take advantage of with our EESD layers and multi-touch interaction.

5.2.7 User Study

Twelve randomly selected employees (9 male, 3 female) of the Institute for Geoinformatics in Münster, Germany (no one who was involved in the project was included)

[1] `http://code.google.com/p/multitouch/`

were asked to provide feedback on the interaction with the first and second EESD layers.

After explaining the possibilities of FTIR multi-touch surfaces, the users were asked three questions:

1. How would you choose two spatial features out of a group of features and establish a connection between them? (How to interact with the first layer)

2. How would you select an area as needed in our interaction with the second EESD?

3. How would you assign an attribute from a list to an area? The answers to these questions were as follows:

1. Eight of the 12 participants would select two features just by clicking (single-touching) the features' icons simultaneously with two fingers, just as we have implemented in our prototype. For icons a small distance apart, one participant would use two fingers of the same hand and for icons further apart one finger of both hands. The four others would double-click (one just single click)

2. Ten of 12 participants would select a region by circumscribing the area with one finger. One of these ten, a trained geographer, would use two fingers simultaneously. Two would use their whole hand to select the region as is established in our prototype. (We let users define a region by using their fingers or of using their complete hand).

3. Seven participants would touch the desired "theme" in the list and then touch inside the selected area. One of these seven users would perform these tasks all at once. The other five participants would drag the theme into the layers as is done in the prototype. When the seven participants who preferred the click interaction were told of the drag method, six of them agreed that the dragging methodology would be more fun and maybe more attractive for an interactive digital globe.

These initial results suggest that while we have to adapt some of our interaction with the virtual globe, most of our interaction paradigm is very intuitive. The informal study also convinced us that users still think in WIMP interaction styles. People have to use multi-touch surfaces more often to become accustomed to their possibilities. After improving the speed of the algorithms we want to formally evaluate the interaction with both EESD layers

5.2.8 Discussion & Future Work

A common theme the authors' previous collaborative work has been to bring to users of state-of-the-art consumer spatial technologies a fuller sense of knowledge about the world. Too often the gift of spatial context provided by these technologies is under-used by applications that only provide obvious functionality and ignore the users' inclination to explore and learn. For location-based services on mobile devices, this obvious functionality often involves pointing users to the nearest pizza parlor or pub. The corollary for virtual globes is, according to our survey, navigation, sightseeing, and, mainly, innocuous voyeurism. While these applications are certainly useful, much is lost, particularly with respect to the "objects and their relationships" that make up the inspiration for and the answer to spatial thinking questions. These thus-far missed opportunities do the greatest harm to geography education both intentional and incidental something that is widely recognized as severely lacking in many parts of the world. For virtual globes, geography education represents both a largely untapped financial market and a chance to enhance world knowledge. The possibility of virtual globes facilitating the spatial thinking process provides an exciting avenue for this technology to reach its full potential. We have provided a glimpse of what is possible when spatial thinking is enabled in a virtual globe. However, there is an extensive amount of future work yet to be done. First and foremost, more research must be completed into the current state of virtual globe use. A wider and more structured survey would be extremely useful and could be used to formally derive a requirements analysis for a spatial thinking-facilitating virtual globe prototype. Secondly, more robust theoretical framework for the Explicitly Explanatory Spatial Data (EESD) layers must be developed. Additionally, we must complete formal studies of the interaction with EESD data layers, particularly with regard to the degree to which it enhances spatial thinking. Separately, much work is being done to develop the EESD types used as prototypes in this work. For instance, implementation speed must be improved. Depending on the entity, the calculation of the second EESD layer can take up to 10 minutes using the Wikipedia data set. Additionally, cartographic research must be used to inform the visualization of these layers.

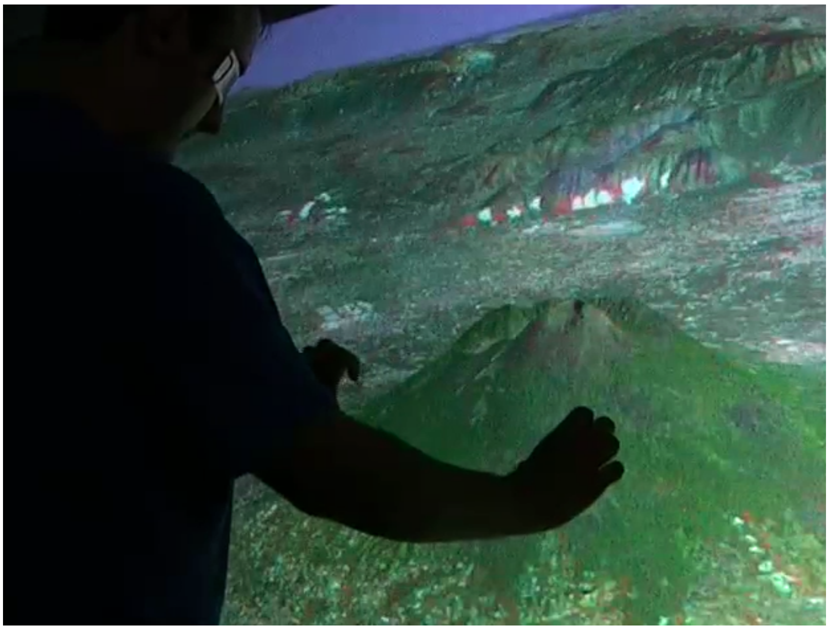

Figure 5.16: A user interacting with an *interscopic multi-touch surface* (iMUTS) in a landscape planning scenario in which depth cues are important, e. g., in a landslide risk management scenario. The geospatial data is displayed in anaglyph stereoscopic mode.

5.3 Multi-touch Interaction with Interscopic Data

In recent years, the visualization of and the interaction with 3D data has become more and more popular and widespread due to the requirements of numerous application areas. 2D desktop systems are often limited in cases in which natural and intuitive interfaces are desired. Sophisticated 3D user interfaces, as they are provided by VR systems consisting of stereoscopic projection and tracked input devices, are rarely adopted by ordinary users or even by experts often due to the overall complexity of the user interface. However, a major benefit of stereoscopy is binocular disparity that provides a better depth awareness. When a stereoscopic display is used, each eye of the user perceives a different perspective of the same scene. This can be achieved by either having the user wear special glasses or by using special 3D displays. In this section we discuss the challenges for iMUTS-based interaction paradigms, and present two new interaction techniques, which underline their benefits.

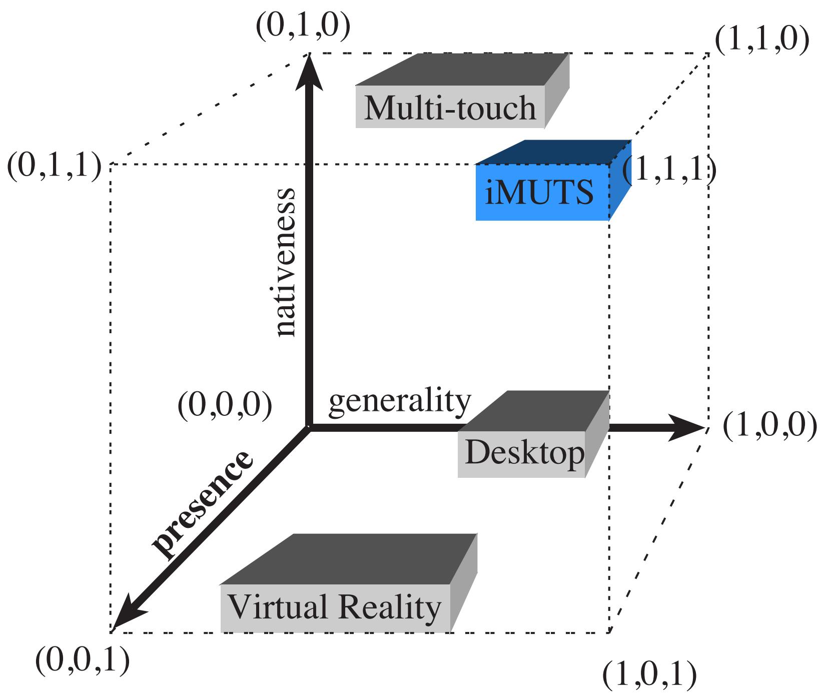

Figure 5.17: Taxonomy of different user interface paradigms and the integration of GUIs based on iMUTS. The three axes representing *generality*, *presence* and *nativeness*.

5.3.1 Interaction Challenges

While it appears quite obviously to use multi-touch for 2D interaction, one might argue that it is not useful to limit interaction with 3D data to a 2D touch surface. As mentioned in the introduction, the usage of complex VR systems still requires much user instrumentation, and it is still quite complex to interact in the 3D space [25]. And as a matter of fact, 2D interaction devices usually supporting only two DoFs [82; 196] are still in widespread use, although they are not the optimal devices for 3D interaction. These both facts motivated us to consider multi-touch interfaces with respect to the requirements of current 3D graphical applications.

5.3.1.1 Taxonomy of User Interfaces

In Figure 5.17 a taxonomy is illustrated which classifies iMUTS paradigms within the broad field of graphical user interfaces. This classification includes conventional desktop systems as well as VR and current multi-touch environments. The taxonomy model is based on a coordinate system involving three axes representing *generality*, *presence* and *nativeness*. Generality captures the variety of interaction tasks that can be performed with the corresponding user interface, presence measures the degree to which a user believes the VE is part of the physical surrounding or vice versa, and nativeness denotes how inartificial the user interface appears with respect to the required instrumentation. Current graphical user interfaces can be classified according to these characteristics. For example, as mentioned in Section 2.4.3.3, desktop-based environments are most appropriate for 2D interaction tasks, but are not optimal for immersive 3D interaction. Hence they are specialized rather than general in this case. Furthermore, virtual scenes are usually displayed monoscopically in desktop-based environments, with only a small field of view and no-head tracking is supported; therefore, the user's sense of presence is often lower in this case. Although interaction with traditional input devices is quite natural, the instrumentation with keyboard and/or mouse is less natural than using the hands directly as it is possible with multi-touch user interfaces. For these reasons, desktop systems are mapped to the area close to the origin $(0, 0, 0)$. VR systems increase presence and are suitable for 3D interaction, but lack support for 2D interaction (see Section 2.4.3.3). Furthermore, instrumentation by haptic devices such as data gloves and/or immersive display systems such as head-mounted-displays leads to non-native interaction. Hence we arrange VR systems similar to desktop-based environments, but closer to the front bottom edge. Current multi-touch user interfaces are suitable for ordinary desktop tasks, they provide the same sense of presence, but they also allow native interaction by the hands. Hence they are arranged above desktop systems showing increased nativeness. We believe that the iMUTS paradigm has the potential to increase presence (e. g., due to stereoscopic projection) in comparison to existing multi-touch user interfaces, and that they will support intuitive 2D as well as 3D interactions while not requiring additional instrumentation of the user. Hence, iMUTS user interfaces may meet most requirements for current GUIs. We admit that this taxonomy is neither perfect nor universal, but it points out some of the benefits that we believe iMUTS paradigms may provide.

To summarize: with iMUTS, humans may interact spontaneously while no heavy instrumentation is required. Further advantages of iMUTS are:

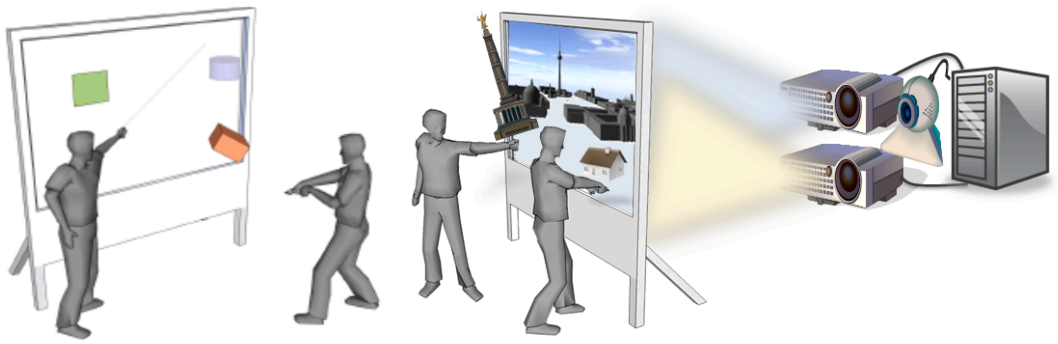

Figure 5.18: Illustration of two users interacting with stereo content, as well as monoscopic content (green rectangle: zero parallax, orange-colored box: negative parallax, purple-colored cylinder: positive parallax) (left). Illustration of two users interacting with a city planning application on an interscopic multi-touch surface (iMUTS) (right).

- Multiple Users: iMUTS easily supports the interaction between different users.
- Intuitivity: People have developed sophisticated skills for sensing and manipulating their physical environments [228]. iMUTS supports these skills.
- Spontaneity: User can switch between different tasks spontaneously.
- Costs: The presented iMUTS setup can be realized with low costs.

As mentioned in Section 2.4.3.3, interaction with stereoscopically displayed objects is still a challenging task [25], in particular when the interaction is restricted to a 2D touch surface. This is the main limitation of an iMUTS. In the following section, we explain this issue and discuss solutions, which have proven their usability in other domains.

5.3.1.2 Parallax Problems

In order to display graphical content stereographically, two half-images have to be generated, i. e., one for each eye. When using stereoscopic projection, a 3D impression occurs due to binocular disparity, which means that objects in space are projected to different positions on the screen. The corresponding horizontal displacement results in essentially three different stereoscopic display paradigms: negative, zero and positive parallax. Objects having zero parallax are displayed monoscopically and therefore are ideally suited for multi-touch interaction (see green-colored box in Figure 5.18). Both eyes perceive the same image, which causes a two-dimensional impression. As mentioned in the introduction, for such a situation multi-touch interfaces have considerable potential to enhance the interaction process, in particular when 2D manipulations are intended. Objects with positive parallax appear behind the touch screen and therefore cannot be accessed directly due to the screen limiting the reach of the user (see purple-colored cylinder in Figure 5.18). This is a problem for any kind of direct interaction in stereoscopic environments, and several approaches address this issue [25; 173], e. g., distant objects behind the screen can be selected by casting a virtual ray [144].

Objects displayed with negative parallax appear in front of the projection screen (see orange-colored box in Figure 5.18) When the user wants to interact with such objects by touching, s/he is limited to touching the area behind the objects since multi-touch screens capture only direct contacts. Therefore, the user virtually has to move fingers or her/himself through virtual objects, and the stereoscopic projection is disturbed. Consequently, immersion may get lost. This problem is a common issue known from two-dimensional representation of the mouse cursor within a stereoscopic image. While the mouse cursor can be displayed stereoscopically on top of stereoscopic objects [219], movements of real objects in the physical space, e. g., the user's hands, cannot be constrained such that they appear only on top of virtual objects. Therefore direct grabbing of objects in front of the touch screen is not possible, and moreover, the hands may interfere with the stereoscopic effect. When objects are displayed with negative parallax, any input devices tracked with 6 DoF support direct interaction. However, we admit that this requires an instrumentation of the user again. Hence, multi-touch devices in combination with image-based gesture recognition may be more appropriate, but have not been used in our setup until now.

5.3.2 Interaction Metaphors

As mentioned above, 3D visualization applications combine two-dimensional with three-dimensional content. While 3D data has the potential to benefit from stereoscopic display, visualization and interaction with 2D content should be restricted to two dimensions [5]. In this section, we discuss aspects that have to be taken into account when designing a multi-touch user interface for interscopic interaction. Therefore, we present interaction metaphors with *interscopic* data for a city planning scenario as well as for medical volume deformation. The described metaphors are not limited to the

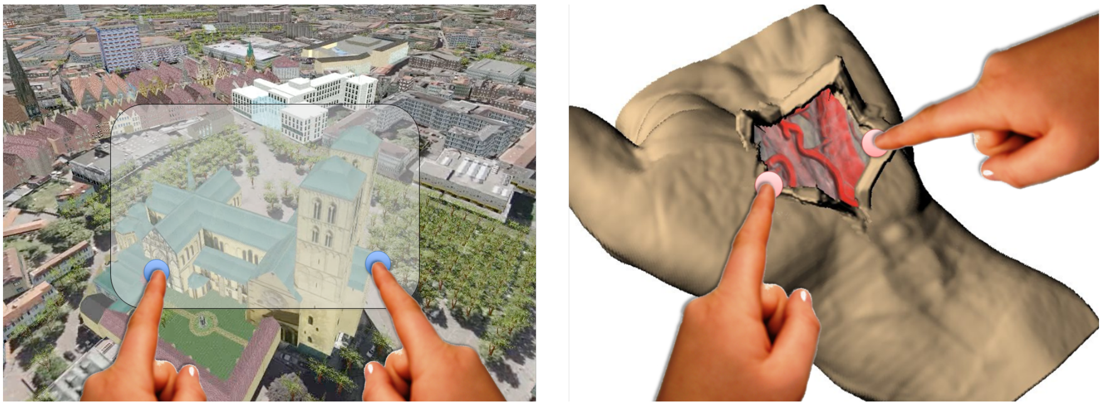

Figure 5.19: Two example applications which can be interfaced by novel user interfaces based on the iMUTS paradigms: (left) navigating through a virtual city model, and (right) multi-touch deformation of volume data in a medical scenario.

city or medical domain, but can be applied also to other scenarios. Both interaction concepts are motivated by the work of [245], which combines multi-touch sensing with a physics engine. In the first scenario, users can manipulate a virtual window analog to a plate on a ball and socket joint. This enables an intuitive way of traveling through a 3D city model. Although such 3D city models are usually represented by polygonal descriptions, the iMUTS interaction paradigms are not limited to polygonal data. To underline that also other formats, such as raster-based and volumetric data representations can be interfaced with iMUTS-based interaction paradigms, we present a second interaction metaphor with which users can directly manipulate a volume dataset by multi-touch deformation.

5.3.3 iMUTS - Technical Setup

In this section, we describe the system components of our iMUTS setup and discuss challenges for user interfaces based on such an interscopic multi-touch environment. There are several ways of combining multi-touch screens with a stereoscopic display in order to improve depth perception while providing intuitive interaction paradigms without further instrumentation of the user. Our multi-touch interscopic wall prototype is based on the FTIR principle introduced by Han [79]. We use same setup as described earlier. For stereoscopic projection we have tested two different setups landscape planning and volume deformation in a medical scenario.

Figure 5.20: Two examples of applications interfaced by the iMUTS interaction paradigms. (Left) a user interacts in a landscape planning scenario on an FTIR multi-touch wall with an anaglyph-based stereoscopic projection. (Right) a user performs volume deformation in a medical scenario using a large FTIR-based passive back projection iMUTS.

- a passive stereoscopic back projection with an FTIR based interactive wall (see Figure 5.20 (right)),
- and a simple anaglyph-based stereoscopic projection again based on an FTIR multi-touch wall (see Figure 5.20 (left)).

The passive stereoscopic projection screen is illustrated in Figure 5.18. Two DLP projectors with a resolution of 1248×1024 provide half images for the left and right eye of the user. The half images are linearly filtered such that users have to wear corresponding polarized glasses. In the case of the anaglyph mode, the half images can be displayed by one projector with a resolution of 1920×1080. Color masks applied to the rendering processes mask both half images such that the user can separate the half images with anaglyph glasses. For both setups the images are rendered by a computer with Intel dual-core processors, 4 GB of main memory and an nVidia GeForce 8800 GTX for rendering purposes.

5.3.4 Discussion & Future Work

In this section, we have introduced a new interscopic multi-touch paradigm that combines traditional 2D interaction performed in monoscopic mode with 3D interaction and stereoscopic projection. We discussed challenges and potentials for the use of multi-touch interfaces for the interaction with interscopic data. In addition, we have introduced two different systems of iMUTS:

- a passive back projection with an FTIR-based interactive wall (see Figure 5.20),
- and a simple anaglyph-based stereoscopic projection again based on an FTIR multi-touch wall.

We believe that iMUTS has great potential to fill the gap between WIMP and VR systems to form the basis of the next generation of 2D and 3D user interfaces. They provide intuitive, fast and spontaneous access to 3D information for multiple users at low costs without requiring user instrumentation. Moreover, we have outlined some challenges and limitations that might occur in such scenarios. Currently, multi-touch walls are horizontally or vertically mounted. VR-based display devices such as the responsive workbench allow users to turn the display from horizontal to vertical. In contrast to vertical multi-touch surfaces, horizontal ones provide the possibility to place physical objects on the surface. For some application domains, it might be beneficial to present the stereoscopic content on a non-planar surface like Microsoft's sphere [17]. With respect to the surface geometry, it might be possible that in some areas objects may not be placed due to instabilities caused by gravitation. The special problem of interaction with stereoscopic data displayed with negative parallax has not been addressed in detail within the scope of this paper and will be considered in future about iMUTS paradigms. However, in such a case, the user can use arbitrary input devices that can be tracked with six degrees of freedom. Another solution might be to allow a user to interactively change the parallax of objects by using a mobile device attached to the user's body as a "soft slider". If the touch surface is portable, the screen can be moved through the VE (analog to the 3D *Window on the World* metaphor) until desired objects are displayed with zero or negative parallax and interaction can be performed as described above. In order to provide stereoscopic images without the need for glasses, in future work we will set up a transparent multi-touch wall in front of an autostereoscopic display. Hence, natural and intuitive interfaces can be provided which do not require any instrumentation, but support 3D interaction. In addition we will carry out further user studies with this next prototype.

5.4 Multi-Touch & Tab Sized Devices

The development of FTIR technology has enabled the construction of large-scale, low-cost, multi-touch displays. These displays—capable of sensing fingers, hands, and whole arms—have great potential for exploring complex data in a natural manner and easily scale in size and the number of simultaneous users. In this context, access and security problems arise if a larger team operates the surface with different access rights. The team members might have different levels of authority or specific roles, which determines what functions they are allowed to access via the multi-touch surface. In this section we present first concepts and strategies to use a mobile phone to spontaneously authenticate and interact with sub-regions of a large-scale multi-touch wall.

5.4.1 Introduction & Motivation

Multi-touch surfaces are well suited for multi-user collaboration with large data sets, such as geographical or time-stamped data. In scenarios with large surfaces (i.e. more than 2 meters) and large groups of users (i.e., more than two) controlling access to content and functionality made available through the multi-touch surface is often an important requirement. However, although FTIR allows identifying a large number of contact points on the wall, it does not discriminate between different users. This makes it difficult to control who is issuing a command. This can lead to severe security problems if the multi-touch wall is used for triggering real-world events, as is the case in control room scenarios. For example, in an emergency response to a flooding event, where a team of experts needs to coordinate mobile forces on the ground (e.g., fire brigades) and monitor data on a geographical representation (e.g., flood level and degree of pollution of air and water), not all users should be able to manipulate all data presented on the multi-touch wall. Depending on the particular policy, only the commander of the fire brigade forces might be allowed to send a mobile unit to a new target (e.g., by pointing to the unit and the new destination). Authentication concepts known from desktop computing are not well suited for these settings, since they usually grant access to an application or the whole computer, rather than to a local area of the screen.

In this section we are addressing the problem that in some collaborative work situations the group of users of a multi-touch wall varies greatly in competence, hierarchical

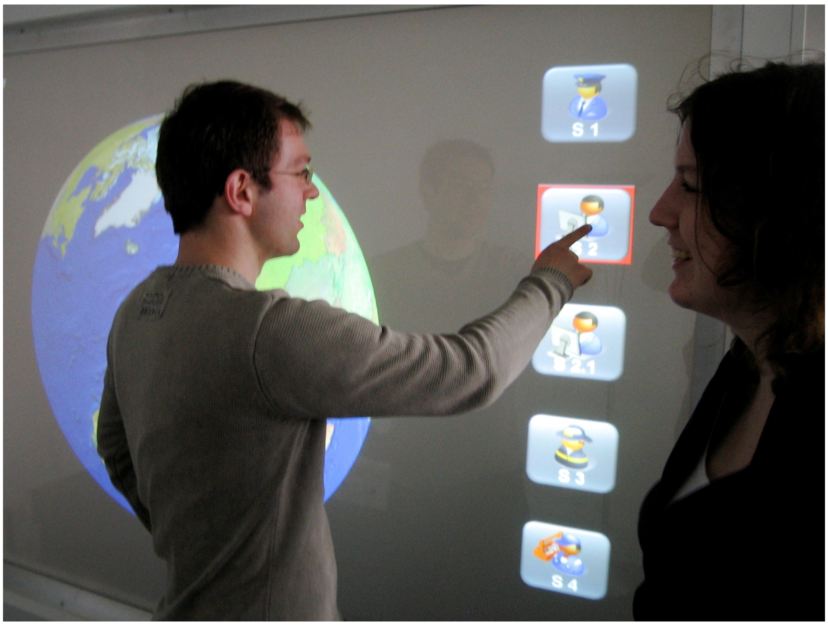

Figure 5.21: Multi-user interaction with a multi-touch wall in an emergency scenario without dedicated access control: The user is selecting an authentication level by pressing a button representing a certain role.

level, and decision-making authority, demanding a dedicated authentication and access mechanism for small regions of a multi-touch surface. We present a first solution for how to authenticate a user who wants to interact with a sub-region of a multi-touch wall. We present novel concepts that enrich the interaction with multi-touch surfaces by using a personal mobile device to spontaneously authenticate and interact with the multi-touch wall.

5.4.2 User Identification & Authentication

As already motivated in the introduction, collaborative work at a multi-touch surface often involves users with different roles, competencies, and scopes of expertise. In an emergency response scenario, for example, a media contact person may be allowed to visualize statistical data on the wall to get an up-to-date picture of the situation, while only the officer-in-charge may command emergency troops at the real emergency site. It would thus increase safety and security if the system could distinguish between users or if individual input events could be authenticated. This would also help in a later

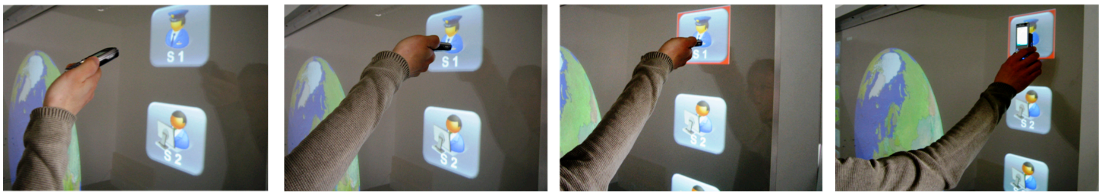

Figure 5.22: Interaction scheme to authenticate with a specific user role on an FTIR multi-touch surface: (i) The user touches the wall with the phone. (ii) The mobile phone flash light sends a light flash (or a camera flash) to indicate the region the user wants to interact with and at the same time initializes the authentication process. (iii+iv) The user can interact in his/her assigned role with the wall and do critical actions.

analysis of the events that took place, since critical operations could be attributed to individual users.

Even in such a scenario we would like to retain the direct-touch interaction scheme of FTIR multi-touch surfaces as much as possible. We assume that most interactions are allowed for every user and that only a small subset of interactions are critical, e.g., because they trigger external real-world events such as sending troops to a specific position. It therefore seems to be acceptable if these critical operations require a slightly higher interaction effort than the other operations.

The minimum requirement to support the above scenario would be to identify the user who generates the critical input event. The system could then check whether the identified user is authorized to trigger the associated action. A better solution would be to also cryptographically authenticate the user attempting the input action instead of mere identification. Of course, it would be best to continuously authenticate each individual contact point, e.g., each contact point during a dragging operation. However, this is not possible given bare finger input and current FTIR technology. It is also not necessary for enabling scenarios like the one outlined above. A solution in which a user "logs in" to a small region in order to gain exclusive access to the region until the user releases that region again does not seem to be adequate, because we assume that, in general, quick access to all parts of the multi-touch surface is required.

We therefore propose to identify — and if possible also authenticate —users in the case of critical operations by using a mobile device as a mediator. We assume that the device contains a flash light and Bluetooth connectivity, and is able to detect touch events with an integrated microphone or accelerometer. If a microphone or an accelerometer is not available, a button could be pressed when the device is touching the region. We further assume that the FTIR system has a second camera that detects light flashes in the visible range. The basic identification scheme (without cryptographic authentication) works as follows:

1. The user touches a region of the wall with the phone.

2. The phone detects the touch event with its built-in accelerometer or microphone and generates a light flash. Simultaneously it sends the user' ID via Bluetooth. (Optionally, microphones can be installed at the multi-touch surface as proposed in [167] to determine the position of touch event on the surface.)

3. The surface detects the light flash at a certain position and receives the user ID via Bluetooth. The light flash can be distinguished from finger touch events, because it produces a bright light strobe in the visible range, whereas finger touch events are detectable mainly in the infrared range.

4. The surface either detects the light flash first or receives the user ID via Bluetooth first. Both events have to be received within a short time window Δt. If either one is missing or if they are more than Δt apart, the protocol is aborted. If more than one flash event and one ID event are detected during a time window extending from Δt before the first event and Δt after the second event, this is considered as a collision.

5. If a collision was detected the server asks one of the devices that have sent an ID to repeat the procedure. Here also random backoff procedures could be used to resolve the collision, in which the device waits a random amount of time before a retransmission is attempted (c.f. Ethernet media access).

6. If a unique association of position and user ID is found the server looks up the authorization data for the object at the respective position and checks whether the user is allowed to perform the action. If so, a positive response is sent via Bluetooth and the action is executed. In addition, visible feedback on the region is given to indicate success or failure.

Figure 5.23: General interaction scheme to identify a user with a certain area on an FTIR multi-touch surface: (i) The user touches the wall with the phone. (ii) The mobile phone flash light sends a light flash (or a camera flash) to indicate the region the user wants to interact with and at the same time initializes the authentication process. (iii+iv) The user is identified can interact in his/her assigned role with the wall and do critical actions. The more detail scheme is described in the body of that paper.

Similar to our approach is the work of Mayrhofer et al. [138]. They present a method for establishing and securing spontaneous interactions on the basis of spatial references which are obtained by accurate sensing of relative device positions. In their work they implemented an interlocked protocol using radio frequency messages and ultrasonic pulses for verifying that two devices share a secret.

The above algorithm uniquely identifies input events on individual regions, even with multiple simultaneous users generating finger input events and multiple users generating phone touch events. If a user touches some other object this will generate only a Bluetooth ID event, but no flash event will be detected by the surface, so the algorithm will abort or a collision with another user will happen. The algorithm is guaranteed to uniquely associate user identities to regions if both events are generated and sensed within Δt.

A shortcoming of this algorithm is that it is not cryptographically secure. An attacker could forge a user ID and thus execute unauthorized critical operations on behalf of another user. We identified the following requirements for an algorithm that authenticates input on a sub-region of the multi-touch wall to support the above scenario:

- The main goal is to ensure that critical operations are only executed by authorized users. The authentication scheme thus has to prove the identity as well as the input position of the user who attempts the operation.

- The system should log all critical interactions for later analysis and documentation. Ideally, the system should also ensure non-repudiation of critical interactions. It should be possible to reconstruct who was responsible for which interaction.

- The system should allow for easy and spontaneous authentication without requiring too much effort and without interfering with other simultaneous users who perform non-critical operations.

With Bluetooth we have a high bandwidth connection but we cannot determine the position on the multi-touch surface where the user actually touched the surface. With the flash light we have a very low bandwidth data channel and way to detect the input position. We assume that the multi-touch surface server and all mobile devices that are allowed to interact with the surface have a pair of cryptographic keys — a public key, a private key, and a corresponding certificate.

We propose the following preliminary authentication scheme. In order to prevent forging, the user ID is signed with the private key of the mobile device before sending it to the server. To prevent replay attacks a timestamp and a sequence number are included in the authentication request. The authentication protocol proceeds as follows:

1. The user touches region of the wall with the phone.

2. The phone detects the touch event with its built-in accelerometer or microphone and generates a light flash. Simultaneously it sends the message m via Bluetooth:

   $$m = enc(R', pubKey_{server})$$

   with

   $$R' = (R, sign(hash(R), privKey_{device}))$$

   $$R = (opcode, userID, time, seq.nr., rand.delay)$$

   $$opcode = inputrequest$$

   We assume that only the device knows $privKey_{device}$ and thus only it is able to generate a valid "input request" message.

3. The surface detects the light flash at a certain position and receives m via Bluetooth. If the content of m cannot be verified it is discarded. Verification includes the signature, the timestamp, and the sequence number for that device.

4. As above, if more than one flash event and one ID event are detected during a time window extending from Δt before the first event and Δt after the second event, this is considered as a collision.

5. As above, if a collision was detected the process is repeated.

6. As above, authorization is performed and feedback is given accordingly.

We assume that a valid signature of the message sent via Bluetooth can only be generated by the device containing the private key. Therefore the server can be sure that a successfully verified ID stems from an authentic input request. If an attacker produces or replays an input request, verification will fail at the server. However, an attacker can produce flash events. If we assume that the authentic device produces a flash event as well, the attacker can only produce a collision.

A problem occurs, if a device generates an input request, but the corresponding light flash is not detected by the surface. This could happen if a touch event is triggered while not facing the surface. In this case the light flash would never reach the surface and an attacker could produce a light flash on some random display region.

To solve this problem, a second light flash could be produced after a random delay whose duration is sent in m (see step 2 above). The attacker would then have to guess the right delay and produce the second flash at exactly the right moment. If the server detects a flash before the indicated delay, the procedure is aborted. The security of this approach depends on the accuracy with which the camera can detect the light flashes. In the current setup, the camera runs at 30 Hz, which severly limits the bandwidth of the visual channel. An obvious way to get a higher bandwidth is to increase the frame rate of the camera. We are also working on other solutions. One idea is to introduce a light back channel. A challenge could be sent by projecting a pattern on the surface next to the detected light spot. The camera of a mobile device is normally located next to the light flash and could detect the challenge and send it back to the server (signed and encrypted). This approach has the advantages that the back channel via the mobile device camera has a higher bandwidth and we can be sure that the user is actually interacting with the right sub-regions of a large-scale multi-touch wall.

For the implementation we use a Nokia 5500 with a built-in flash light and the Nokia N95 using its built-in camera flash. A camera image (recorded by a DragonFly camera

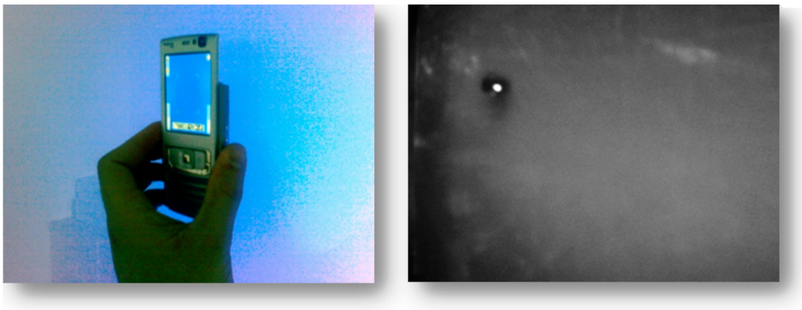

Figure 5.24: User is touching the multi-touch wall with a mobile device. Raw camera image of the phone flash using a Dragonfly Camera with an infrared filter.

with an infrared filter) of the raw camera image and the N95 touching the multi-touch surface can be seen in Figure 5.24.

5.4.3 Conclusions & Future Work

We addressed the problem of spontaneous authentication of individual input actions in the context of large-scale multi-touch FTIR surfaces. We described an access mechanism for small sub-regions of the surface that is capable of authenticating multiple simultaneous users. Users have to touch the wall with their personal mobile device for spontaneous authentication and interaction. We still have to do user tests on the usability and general acceptability of the proposed scheme. We intend to do a formal security analysis of the method and to evaluate it with real users in an emergency operation center. As future work, we plan to add additional functionality to our prototype. As an example, while the user touches the surface, the front camera can take a photo of the user and we can verify if the right user acts with the mobile device. Other functionalities beyond the authentication problem can be easily added. For example, in our emergency response scenario, a secured voice call connection could be easily established by touching the icon of a first responder troop on the surface. We also experiment with other output modalities like the display light or, if available, the IrDA port.

6

Conclusions, Contributions & Future Work

In this chapter, the general contributions of this thesis are summarized. While minor contributions of individual research projects are discussed when the projects are described, this section focuses on major contributions that span all or significant portions of the thesis. This is followed by a brief broad-based discussion of future work. Finally, this chapter and the whole thesis are concluded with some final remarks.

6.1 Major Contributions

Digital spatial information will play an increasingly larger role in many of our day-to-day and longer-term decisions. Therefore, good user interfaces are needed to interact with this information. This thesis explores the design and evaluation of new tangible and natural interfaces for spatial information within the framework of ubiquitous computing.

The main core contributions are in three areas.

1. New findings about how people interact with analog spatial information in combination with digital spatial information are derived. This is done via empirical studies as presented in sections 4 and 5. Both studies provides unique and new information that can be reused by other researchers to improve the design of interface for spatial information. This new findings could easily generalize to digital-analog information interfaces in any domain.

2. Novel interfaces are developed to solve some of the issues raised in our digital-analogue information interaction studies. This contribution can be generalized and transferred to other domains. Indeed, the combination of analogue and digital information is becoming increasingly important to interface design in general, especially for devices that support people in their daily life and tasks, where paper-based artifacts still play an important role. As one example the iBookmark concept will apply for many years to come. This is an entirely new narrative schema that opens the door for an entirely new class of stories and can be easily ported to new interfaces that just come up for example Apples iPad.

3. Tools, prototypes and methods for new user interfaces for spatial information are developed. The consideration of all three of Weiser's device classes is unique. By exploring this wide scope of a device class range, interesting insights are provided, even for research well outside the spatial domain. In addition, as more and more spatial technologies are developed, the robust user interfaces found in this thesis can serve as a guide to interface design moving into the future.

This research takes place in the context of very rapid technological change in the area of its focus. Therefore the contributions have to be placed in this context. This and the shifting environment is best exemplified by examining the meteoric advancement of the Google Maps Mobile application.

- In late 2006, Google introduced a Java application called Google Maps for Mobile. Some of the web-based site's features were provided in the application.

- One year later, Google Maps Mobile 2.0 was released. It introduced localization services based on Cell-ID and WiFi.

- At the of end of 2008, Google Maps Mobile was ported to (nearly) all mobile platforms. Features like a friend finder were introduced.

- Just a few weeks ago, the free Google Maps Navigation was released in conjunction with Google Android OS 2.0 Eclair on the Motorola Droid. This version added voice commands, traffic reports, and street view support.

In times of frantic progress, quality and complete understanding is often sacrificed for quick action and knee-jerk reactions. The rapid advancement in the spatial data domain highlights the need for robust and quality research in the area. It is hoped that the new findings and design practices developed here can serve as a guide, no matter how far and fast spatial technologies advance. In addition the developed prototypes show how serve as good examples to interface design moving into the future. Interesting more and more devices come up that supports the hardware specifications that are needed for the several prototypes. Apple's iPad (a GPS enabled eBook reader) or Samsung's camera projector unit called Samsung i8520 Beam.

To sum this up the following table provides an overview on the newly developed prototypes and interaction techniques in this thesis. We also highlight to which contribution area (as described above) the prototypes and interaction techniques contribute most.

Project Name or Interaction Technique	**Contribution**	**1)**	**2)**	**3)**
Mobile Map Interaction	These new interfaces that allow the user to combine the advantages of paper maps with the advantages of mobile devices, a novel interaction principle are called MMI. This implementation uses a mobile phone as a magic lens that the user can combine with a paper map to get additional dynamic and personalized information. With Wikear, SightQuest and the mobile response scenario we successfully showed how analog maps and dynamic digital screens can be combined in various domains. We show how the combination of a high resolution analog paper map in combination		X	x
Mobile Projection	We have presented different application classes of interfaces utilizing a mobile camera-projector unit. The interfaces all focus on the augmentation of real word objects in the environment. We showed how the different spatial setups of camera and projector units effect the possible applications and the physical interaction space.		X	X

PhotoMap	With Photomap we have presented our initial explorations into the technical feasibility and associated usability implications of allowing GPS-enabled mobile phones to support the capture, georeferencing and subsequent display of traditional 'YOU ARE HERE' map signage. Detailed minor contributions are described within the section.	x	X	x
GeoBook & iBookmark	We have presented a large scale quantitive survey of the German geocaching community. This lead us to the design of the GeoBook. In addition we presented the first combination of an eBook reader with a GPS Device to create stories that change in response to the location of the eBook itself, as well as other context variables.	X		X
Multi-touch Virtual Globe	The possibility of virtual globes facilitating the spatial thinking process provides an exciting avenue for this technology to reach its full potential. We have provided a prototype of what is possible when spatial thinking is enabled in a virtual globe and described a detailed technical description.	x	X	X
Multi-touch with stereoscopic data	We have introduced a new interscopic multi-touch paradigm that combines traditional 2D interaction performed in monoscopic mode with 3D interaction and stereoscopic projection. We discussed challenges and potentials for the use of multi-touch interfaces for the interaction with interscopic data. In addition, we have introduced two different systems of iMUTS		X	x
Multi-touch and pad sized devices	We addressed the problem of spontaneous authentication of individual input actions in the context of large-scale multi-touch FTIR surfaces. We described an access mechanism for small sub-regions of the surface that is capable of authenticating multiple simultaneous users.		X	

Table 6.1: Prototypes or interaction techniques and their contributions. In the last three columns their share for the main contributions is highlighted. A "X" shows a strong contribution to the main contributions and a "x" a minor one.

6.2 Future Work

As Schmidt [203] says "with each prototype finished, each system evaluated, and each paper published, a number of new issues that pose interesting challenges usually appear". These new issues are explored in the future work sections of the individual chapters. However, several overarching themes emerge from a complete examination of the future work suggested in this thesis. These themes are as follows:

Virtual vs. analog media: The transfer of results of this thesis to other domains where virtual and analog information are often used together. As an example it would be interesting to design a interface where all media is interconnected – analog and digital one. If one reads a text (printed on a piece of paper), the according video jumps to the right frame displayed on this mobile devices next to the user. This combination of the advantages of media types of both worlds can be transferred to various domains, starting from education, entertainment or supporting every day tasks.

Approaching the analog/digital divide from both sides: We claim to put more focus on closing the gap between analogue and digital media from the analogue side rather than the digital side. Following the ideas of Gentner [70] - for example labeling a light switch with *1* and *0* - we think that this a very important area for future work. This is also inline with the dual reality paradigm proposed by Lifton [126] or similar ideas by Brandherm [26].

More prototypes and field testing: The development of additional prototypes in different areas addressed in this thesis, e.g. the area of mobile camera projector units. In addition, prototypes performance in the "real world / in the wild" must be examined. Many of the shown examples have market potentials and can be turned into a product. The MMI or PhotoMap are just two examples. While the iPad – seen mainly as an eBook reader – comes with a build in GPS, the iBookmark application could become very powerful on that device. In addition the MMI principle can be transferred to the retail domain as a mobile product lens as shown in figure 6.1.

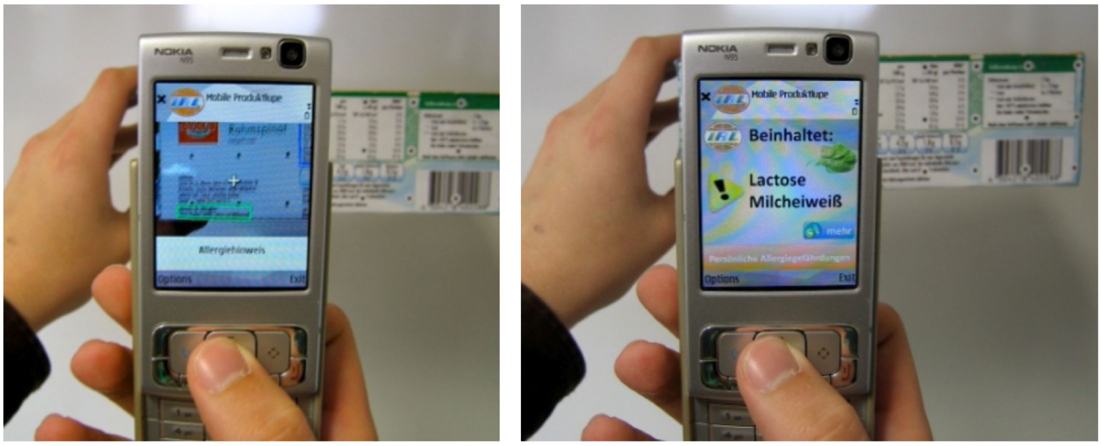

Figure 6.1: The mobile product lens. The MMI principle is transferred to the retail domain – A user can explore various features of the product, e.g. the ingredients and connected it with allergy warnings.

Understanding the user: Longitudinal studies how the new technologies presented affect daily lives are needed to understand the effects of the technology on the users. The short time frame of a PhD is often not enough to investigate that in detail.

6.3 Concluding Remarks

What Schmidt[203] not says is that in contrast to pure technical innovation the contingency of design on other (human) factors, to an extent not often acknowledged. I claim that this is very important for doing "good" research (whatever we define as "good research") and this is a major strength of this thesis.

Kevin Arthur blog entry[1] provides a good summary of two articles of Norman and Ihde that discuss this point. In an article[2], design/HCI guru Don Norman claims that "major innovation comes from technologists who have little understanding of all this research stuff: they invent because they are inventors."

Norman begins his article as follows:

> I've come to a disconcerting conclusion: design research is great when it comes to improving existing product categories but essentially useless

[1] http://www.touchusability.com/2009/12/1

[2] http://jnd.org/dn.mss/technology_first_needs_last.html

> when it comes to new, innovative breakthroughs. I reached this conclusion through examination of a range of product innovations, most especially looking at those major conceptual breakthroughs that have had huge impact upon society as well as the more common, mundane small, continual improvements. Call one conceptual breakthrough, the other incremental. Although we would prefer to believe that conceptual breakthroughs occur because of a detailed consideration of human needs, especially fundamental but unspoken hidden needs so beloved by the design research community, the fact is that it simply doesn't happen. Grand conceptual inventions happen because technology has finally made them possible rather than because of design research into people's needs. Major innovation comes from technologists who have little understanding of all this research stuff: they invent because they are inventors.

Norman's article is in line with an essay by philosopher of technology Don Ihde [96].

> I am attempting to show that the design situation is considerably more complex and less transparent than it is usually taken to be. Both the designer-materiality relation, and the artifact-user relations are complex and multistable. While it is clear that a new technology, when put to use, produces changes in practices – all of the examples show that – these practices are not of any simple "deterministic" pattern. The results are indeterminate but definite, but also multiple and diverse. Moreover, both intended results and unintended results are unpredictable in any simple way, and yet results are produced.

As we said in the introduction of this subsection it seems that both support the claim that, contingency of design on other (human)factors, to an extent not often acknowledged. To make this more concrete, take the example of the telephone (cited in both essays and also highlighted by Arthur). Norman's point is that the telephone wasn't invented in response to research into people's needs; it was invented because the technology was ready. Ihde's point is that when the telephone was invented, the inventor had a different use in mind than what later happened:

> Here the designer intent was for an amplifying device capable of transmitting a voice over distance, and intended as a prosthetic technology for the hard-of-hearing (Bell's mother) the party line on which all the neighbors chatted was not foreseen, let alone the subsequent telephone wiring of early 20th century America.

Following Ihde and Norman, we encourage readers to consider all factors that affect technology design and use, not just those presented here. Therefore, we do not claim that the factors analyzed in this thesis are the only factors that are important when designing user interfaces for spatial information. More multi-disciplinary research is needed to understand how technology impacts our daily lives. More researchers have to back up the technology innovations with users needs.

References

[1] Abowd, G., Atkeson, C., Hong, J., Long, S., Kooper, R., and Pinkerton, M. Cyberguide: A mobile context-aware tour guide. *Wireless Networks 3*, 5 (1997), 421–433.

[2] Alvarado, C., and Davis, R. SketchREAD: A multi-domain Sketch Recognition Engine. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (2004), ACM New York, NY, USA, pp. 23–32.

[3] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and Macintyre, B. Recent advances in augmented reality. *Computer Graphics and Applications, IEEE 21*, 6 (2001), 34–47.

[4] Azuma, R., et al. A survey of augmented reality. *Presence-Teleoperators and Virtual Environments 6*, 4 (1997), 355–385.

[5] Balakrishnan, R. A Grant of 3D. Why aren't we using 3D user interfaces, and will we ever? (3DUI Keynote), 2006.

[6] Baldauf, M., Fröhlich, P., and Reichl, G. Gestural interfaces for micro projector-based mobile phone applications. In *Adjunct Proceedings Ubicomp 2009* (2009).

[7] Banerjee, S., and Pedersen, T. Extended gloss overlaps as a measure of semantic relatedness. In *International Joint Conference on Artificial Intelligence* (2003), vol. 18, pp. 805–810.

[8] Barkhuus, L., Chalmers, M., Tennent, P., Hall, M., Bell, M., Sherwood, S., and Brown, B. Picking pockets on the lawn: The development of tactics and strategies in a mobile game. In *Proceedings of UbiComp 2005* (2005), Springer, pp. 358–374.

[9] Baudisch, P., Good, N., Bellotti, V., and Schraedley, P. Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (2002), ACM New York, NY, USA, pp. 259–266.

[10] Baudisch, P., and Rosenholtz, R. Halo: a technique for visualizing off-screen objects. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2003), ACM, pp. 481–488.

[11] Baus, J., Cheverst, K., and Kray, C. A Survey of Map-based Mobile Guides. *Map-based mobile services-Theories, Methods, and Implementations* (2003), 197–216.

[12] Baus, J., Krüger, A., and Wahlster, W. A resource-adaptive mobile navigation system. In *Proceedings of the 7th international conference on Intelligent user interfaces* (2002), ACM New York, NY, USA, pp. 15–22.

[13] Bell, B., Feiner, S., and Höllerer, T. View management for virtual and augmented reality. In *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology* (New York, NY, USA, 2001), ACM, pp. 101–110.

[14] Bellotti, V., and Bly, S. Walking away from the desktop computer: distributed collaboration and mobility in a product design team. In *CSCW '96: Proceedings of the 1996 ACM conference on Computer supported cooperative work* (New York, NY, USA, 1996), ACM, pp. 209–218.

[15] Benford, S., Crabtree, A., Flintham, M., Drozd, A., Anastasi, R., Paxton, M., Tandavanitj, N., Adams, M., and Row-Farr, J. Can you see me now? *ACM Transactions on Computer-Human Interaction (TOCHI) 13*, 1 (2006), 100–133.

[16] Benko, H., and Feiner, S. Balloon selection: A multi-finger technique for accurate low-fatigue 3d selection. *3D User Interfaces 7* (2007), 79–86.

[17] Benko, H., Wilson, A. D., and Balakrishnan, R. Sphere: multi-touch interactions on a spherical display. In *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology* (New York, NY, USA, 2008), ACM, pp. 77–86.

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

[19] Billinghurst, M., Kato, H., and Poupyrev, I. Magicbook: transitioning between reality and virtuality. In *CHI '01: CHI '01 extended abstracts on Human factors in computing systems* (2001), ACM, pp. 25–26.

[20] Bimber, O., Grundhöfer, A., Zollmann, S., and Kolster, D. Digital illumination for augmented studios. *Journal of Virtual Reality and Broadcasting 3*, 8 (2006).

[21] Bimber, O., and Raskar, R. Modern approaches to augmented reality. In *International Conference on Computer Graphics and Interactive Techniques* (2006), ACM New York, NY, USA.

[22] Blasko, G., Coriand, F., and Feiner, S. Exploring interaction with a simulated wrist-worn projection display. In *ISWC 05 IEEE International Symposium on Wearable Computers* (2005), pp. 2–9.

[23] Board, C. Maps as models. *Models in geography* (1967), 671–725.

[24] Bordwell, D. *Narration in the fiction film.* Routledge, 1985.

[25] Bowman, D., Kruijff, E., LaViola, J., and Poupyrev, I. *3D User Interfaces: Theory and Practice.* Addison-Wesley, 2004.

[26] Brandherm, B., Ullrich, S., and Prendinger, H. Simulation of sensor-based tracking in Second Life. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems: demo papers* (2008), International Foundation for Autonomous Agents and Multiagent Systems, pp. 1689–1690.

[27] Brocklehurst, E. The NPL Electronic Paper Project. *International Journal of Man-Machine Studies 34*, 1 (1990), 69–95.

[28] Brown, B., and Chalmers, M. Tourism and mobile technology. In *ECSCW'03: Proceedings of the eighth conference on European Conference on Computer Supported Cooperative Work* (Norwell, MA, USA, 2003), Kluwer Academic Publishers, pp. 335–354.

[29] Buchholz, H., Bohnet, J., and Dollner, J. Smart and physically-based navigation in 3D geovirtual environments. In *In Proceedings of Ninth International Conference on Information Visualisation* (2005), pp. 629–635.

[30] Buchmann, V., Violich, S., Billinghurst, M., and Cockburn, A. Fingartips: gesture based direct manipulation in augmented reality. In *GRAPHITE '04: Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia* (New York, NY, USA, 2004), ACM, pp. 212–221.

[31] Budanitsky, A., and Hirst, G. Evaluating wordnet-based measures of lexical semantic relatedness. *Computational Linguistics 32*, 1 (2006), 13–47.

[32] Buriol, L., Castillo, C., Donato, D., Leonardi, S., and Millozzi, S. Temporal analysis of the wikigraph. In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence* (2006), IEEE Computer Society, pp. 45–51.

[33] Butler, D. The web-wide world. *Nature (London) 439*, 7078 (2006), 776–778.

[34] Buxton, W., and Myers, B. A study in two-handed input. *SIGCHI Bull. 17*, 4 (1986), 321–326.

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

[36] Chavez, D., Courtright, R., and Schneider, I. Over the river and through the woods. *Parks & Recreation (Ashburn) 39*, 4 (2004), 68–72.

[37] Chavez, D., Schneider, I., and Powell, T. The Social-Psychology of a Technology Driven Outdoor Trend: Geocaching in the USA. *Proceedings of HICSS* (2004).

[38] Cheverst, K., Davies, N., Mitchell, K., Friday, A., and Efstratiou, C. Developing a context-aware electronic tourist guide: some issues and experiences. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2000), ACM, pp. 17–24.

[39] Cheverst, K., Schöning, J., Krüger, A., and Rohs, M. Photomap: Snap, grab and walk away with a "you are here" map. *In Proc. of MobileHCI '08 : Workshop on Mobile Interaction with the Real World* (2008).

[40] Chun, W.-S., Napoli, J., Cossairt, O. S., Dorval, R. K., Hall, D. M., II, T. J. P., Schooler, J. F., Banker, Y., and Favalora, G. E. Spatial 3-D Infrastructure: Display-Independent Software Framework, High-Speed Rendering Electronics, and Several New Displays. In *In Stereoscopic Displays and Virtual Reality Systems XII* (2005), vol. 5664, pp. 302–312.

[41] Conner, D. B., Snibbe, S. C., Herndon, K. P., Robbins, D. C., Zeleznik, R. C., and van Dam, A. Three-Dimensional Widgets. In *Symposium on Interactive 3D Graphics* (1992).

[42] Cook, C., Heath, F., and Thompson, R. A Meta-Analysis of Response Rates in Web-or Internet-Based Surveys. *Educational and Psychological Measurement 60*, 6 (2000), 821.

[43] Costanza, E., and Robinson, J. A region adjacency tree approach to the detection and design of fiducials. *Vision, Video and Graphics (VVG* (2003), 63–70.

[44] Costanza, E., Shelley, S., and Robinson, J. Introducing audio d-touch: A tangible user interface for music composition and performance. *Proc. of the 6th Intl Conf. on Digital Audio Effects (DAFX)* (2003), 63–70.

[45] Davies, N., Cheverst, K., Dix, A., and Hesse, A. Understanding the role of image recognition in mobile tour guides. In *MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services* (New York, NY, USA, 2005), ACM, pp. 191–198.

[46] Davies, N., Cheverst, K., Mitchell, K., and Friday, A. 'caches in the air': Disseminating tourist information in the guide system. In *WMCSA '99: Proceedings of the Second IEEE Workshop on Mobile Computer Systems and Applications* (Washington, DC, USA, 1999), IEEE Computer Society, p. 11.

[47] Davis, R. Sketch understanding in design: Overview of work at the MIT AI lab. In *Sketch Understanding, Papers from the 2002 AAAI Spring Symposium* (2002), pp. 24–31.

[48] Denning, P., Horning, J., Parnas, D., and Weinstein, L. Wikipedia risks. *Commun. ACM 48*, 12 (2005), 152–162.

[49] Dietz, P., and Leigh, D. DiamondTouch: A Multi-user Touch Technology. *Proceedings of the 14th annual ACM symposium on User interface software and technology* (2001), 219–226.

[50] Dix, A., Finlay, J., Abowd, G., and Beale, R. *Human-computer interaction*. Prentice hall, 2004.

[51] Downs, R. Using resistive touch screens for human/machine interface. Tech. rep., Texas Instruments Incorporated, Texas Instruments, Post Office Box 655303, Dallas, Texas 75265, 2005.

[52] Drascic, D., and Milgram, P. Perceptual issues in augmented reality. In *SPIE Volume 2653: Stereoscopic Displays and Virtual Reality Systems III* (1996), pp. 123–134.

[53] Dyer, M. *The Essential Guide To Geocaching: Tracking Treasure With Your GPS*. Fulcrum Publishing, 2004.

[54] Echtler, F., Pototschnig, T., and Klinker, G. An led-based multitouch sensor for lcd screens. In *TEI '10: Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction* (New York, NY, USA, 2010), ACM, pp. 227–230.

[55] Echtler, F., Sielhorst, T., Huber, M., and Klinker, G. A short guide to modulated light. In *TEI '09: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction* (New York, NY, USA, 2009), ACM, pp. 393–396.

[56] Egenhofer, M. J. Query Processing in Spatial-Query-by-Sketch. *Journal of Visual Languages and Computing 8*, 4 (1997), 403–424.

[57] Elmqvist, N., Axblom, D., Claesson, J., Hagberg, J., Segerdahl, D., So, Y., Svensson, A., Thoren, M., and Wiklander, M. 3DVN: a mixed reality platform for mobile navigation assistance. Tech. rep., Technical Report, 2006.

[58] Elrod, S., Bruce, R., Gold, R., Goldberg, D., Halasz, F., Janssen, W., Lee, D., McCall, K., Pedersen, E., Pier, K., et al. Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (1992), ACM New York, NY, USA, pp. 599–607.

[59] Elson, J., Howell, J., and Douceur, J. R. Mapcruncher: integrating the world's geographic information. *SIGOPS Oper. Syst. Rev. 41*, 2 (2007), 50–59.

[60] Epps, J., Lichman, S., and Wu, M. A study of hand shape use in tabletop gesture interaction. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2006), ACM, pp. 748–753.

[61] Fabrikant, S. Evaluating the usability of the scale metaphor for querying semantic spaces. *Lecture notes in computer science* (2001), 156–172.

[62] Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Personal and Ubiquitous Computing 1*, 4 (1997), 208–217.

[63] Fitzmaurice, G., and Buxton, W. The Chameleon: Spatially aware palmtop computers. In *Conference on Human Factors in Computing Systems* (1994), ACM New York, NY, USA, pp. 451–452.

[64] Fitzmaurice, G. W. Situated information spaces and spatially aware palmtop computers. *Commun. ACM 36*, 7 (1993), 39–49.

[65] Fitzmaurice, G. W., Ishii, H., and Buxton, W. A. S. Bricks: laying the foundations for graspable user interfaces. In *CHI '95: Proceedings of the SIGCHI conference on Human factors in computing systems* (1995), ACM Press/Addison-Wesley Publishing Co., pp. 442–449.

[66] Florence, J., Hornsby, K., and Egenhofer, M. The GIS wallboard: Interactions with Spatial Information on large-scale Displays. *International Symposium on Spatial Data Handling 7* (1996), 449–463.

[67] Freitag, U. Map functions. *The Selected Main Theoretical Issues Facing Cartography, Cologne: International Cartographic Association* (1993), 9–19.

[68] Furmanski, C., Azuma, R., and Daily, M. Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visual and obscured information. In *ISMAR '02: Proceedings of the 1st International Symposium on Mixed and Augmented Reality* (2002), IEEE Computer Society, p. 320.

[69] Gartner, G., Cartwright, W., and Peterson, M. *Location Based Services and TeleCartography*. Springer, 2007.

[70] Gentner, D., and Nielsen, J. The anti-mac interface. *Commun. ACM 39*, 8 (1996), 70–82.

[71] Giles, J. Special report: Internet encyclopedias go head to head. *Nature 438*, 15 (2005), 900–901.

[72] Golledge R. G. Geographical Perspectives on Spatial Cognition. *Advances in Psychology 96* (1993), 16–46.

[73] Greaves, A., Hang, A., and 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* (2008), ACM, pp. 527–530.

[74] Greenberg, S., and Buxton, B. Usability evaluation considered harmful (some of the time). In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 111–120.

[75] Grossman, T., Wigdor, D., and Balakrishnan, R. Multi-finger gestural interaction with 3d volumetric displays. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (2004), ACM New York, NY, USA, pp. 61–70.

[76] Gustafson, S., Baudisch, P., Gutwin, C., and Irani, P. Wedge: clutter-free visualization of off-screen locations. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 787–796.

[77] Hachet, M., Pouderoux, J., Guitton, P., and Gonzato, J. TangiMap: A tangible interface for visualization of large documents on handheld computers. In *Proceedings of Graphics Interface 2005* (2005), Canadian Human-Computer Communications Society, p. 15.

[78] Haller, M., Brandl, P., Richter, C., Seifried, T., Leitner, J., and Gokcezade, A. Interactive displays and next-generation interfaces. In *Hagenberg Research*, B. et al., Ed. Springer, 2009, pp. 433–472.

[79] Han, J. Y. Low-cost Multi-touch Sensing through Frustrated Total Internal Reflection. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology* (New York, NY, USA, 2005), ACM, pp. 115–118.

[80] Han, Y., and Wagner, R. An efficient and fast parallel-connected component algorithm. *Journal of the ACM (JACM) 37*, 3 (1990), 626–642.

[81] Hang, A., Rukzio, E., and Greaves, A. Projector phone: a study of using mobile phones with integrated projector for interaction with maps. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services* (2008), ACM, pp. 207–216.

[82] Hanson, A. J., and Wernert, E. Constrained 3D Navigation with 2D Controllers. In *Proceedings of Visualization '97* (1997), IEEE Computer Society Press, pp. 175–182.

[83] Harley, J., and Woodward, D. *The History of Cartography: Cartography in prehistoric, ancient, and medieval Europe and the Mediterranean.* University Of Chicago Press, 1987.

[84] Härmä, A., Jakka, J., Tikander, M., Karjalainen, M., Lokki, T., Hiipakka, J., and Lorho, G. Augmented reality audio for mobile and wearable appliances. *J. Audio Eng. Soc 52*, 6 (2004), 618–639.

[85] Haugen, E. The Semantics of Icelandic Orientation. *Cognitive Anthropology: Readings* (1969), 330–353.

[86] Hecht, B., Dara-Abrams, D., Starosielski, N., Goldsberry, K., Dillemuth, J., Roberts, J., and Clarke, K. Minotour: A location-aware mobile tour application that weaves a spatial tale from wikipedia. In *Proceedings of the 2007 Meeting of The AAG* (2007).

[87] Hecht, B., and Raubal, M. GeoSR: Geographically explore semantic relations in world knowledge. In *Eleventh AGILE International Conference on Geographic Information Science (AGILE'08)* (2008), Springer, pp. 95–114.

[88] Hecht, B., Rohs, M., Schöning, J., and Krüger, A. WikEye–Using Magic Lenses to Explore Georeferenced Wikipedia Content. *Proceedings of the 3rd International Workshop on Pervasive Mobile Interaction Devices (PERMID)* (2007).

[89] Hecht, B., Starosielski, N., and Dara-Abrams, D. Generating Educational Tourism Narratives from Wikipedia. In *Association for the Advancement of Artificial Intelligence Fall Symposium on Intelligent Narrative Technologies (AAAI-INT'07)* (2007), pp. 37–44.

[90] Hedley, N., Billinghurst, M., Postner, L., May, R., and Kato, H. Explorations in the use of augmented reality for geographic visualization. *Presence: Teleoperators & Virtual Environments 11*, 2 (2002), 119–133.

[91] Henrysson, A., Billinghurst, M., and Ollila, M. Face to face collaborative AR on mobile phones. In *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality, 2005. Proceedings* (2005), pp. 80–89.

[92] Hodges, S., Izadi, S., Butler, A., Rrustemi, A., and Buxton, B. Thinsight: versatile multi-touch sensing for thin form-factor displays. In *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology* (New York, NY, USA, 2007), ACM, pp. 259–268.

[93] Höllerer, T., Feiner, S., Hallaway, D., Bell, B., Lanzagorta, M., Brown, D., Julier, S., Baillot, Y., and Rosenblum, L. User interface management techniques for collaborative mobile augmented reality. *Computers & Graphics 25*, 5 (2001), 799–810.

[94] Höllerer, T., Feiner, S., and Pavlik, J. Situated documentaries: Embedding multimedia presentations in the real world. In *Proc. ISWC* (1999), vol. 99, pp. 79–86.

[95] Hubona, G. S., Wheeler, P. N., Shirah, G. W., and Brandt, M. The relative contributions of stereo, lighting, and background scenes in promoting 3d depth visualization. *ACM Trans. Comput.-Hum. Interact. 6*, 3 (1999), 214–242.

[96] Ihde, D. The designer fallacy and technological imagination. *Defining technological literacy: Towards an epistemological framework* (2006), 121–131.

[97] Iliffe, J. *Datums and Map Projections for Remote Sensing, GIS and Surveying.* Whittles Publishing, 2000.

[98] International Organization for Standardization. Ergonomic requirements for office work with visual display terminals (VDTs) – Requirements for non-keyboard input devices. *ISO/IEC 9241–9* (2000).

[99] Isbister, K., and Doyle, P. Web guide agents: Narrative context with character. *ADVANCES IN CONSCIOUSNESS RESEARCH 46* (2002), 229–244.

[100] Ishii, H., and Ullmer, B. Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems* (1997), ACM, pp. 234–241.

[101] Jones, O., Fleuriot, C., and Williams, M. 'A New Sense of Place?' Mobile'Wearable' Information and Communications Technology Devices and the Geographies of Urban Childhood. *Children s Geographies 1*, 2 (2003), 165–180.

[102] Julier, S., Baillot, Y., Brown, D., and Lanzagorta, M. Information filtering for mobile augmented reality. *IEEE Comput. Graph. Appl. 22*, 5 (2002), 12–15.

[103] Jung, Y., Keil, J., Behr, J., Webel, S., Zöllner, M., Engelke, T., Wuest, H., and Becker, M. Adapting x3d for multi-touch environments. In *Web3D '08: Proceedings of the 13th international symposium on 3D web technology* (New York, NY, USA, 2008), ACM, pp. 27–30.

[104] Kähäri, M., and Murphy, D. MARA-sensor based augmented reality system for mobile imaging. *ISMAR2006, Oct* (2006).

[105] Kaltenbrunner, M., and Bencina, R. reacTIVision: a computer-vision framework for table-based tangible interaction. *Proceedings of the 1st international conference on Tangible and embedded interaction* (2007), 69–74.

[106] Kaltenbrunner, M., Bovermann, T., Bencina, R., and Costanza, E. TUIO: A protocol for table-top tangible user interfaces. In *Proc. of the The 6th Intl. Workshop on Gesture in Human-Computer Interaction and Simulation* (2005).

[107] Kantarjiev, C., Demers, A., Frederick, R., Krivacic, R., and Weiser, M. Experiences with X in a Wireless Environment. In *Mobile & Location-Independent Computing Symposium on Mobile & Location-Independent Computing Symposium* (1993), USENIX Association, p. 11.

[108] Kasday, L. R. US patent #4,484,179: Touch position sensitive surface.

[109] Kato, H., Tachibana, K., Tanabe, M., Nakajima, T., and Fukuda, Y. Magiccup: a tangible interface for virtual objects manipulation in table-top augmented reality. In *Augmented Reality Toolkit Workshop, 2003. IEEE International* (Oct. 2003), pp. 75–76.

[110] Kelley, M. *Local Treasures: Geocaching Across America*. Center for American Places; University Presses Marketing, 2005.

[111] Kish, G., and Buache, P. Early Thematic Mapping: The Work of Philippe Buache. *Imago Mundi 28* (1976), 129–136.

[112] Klein, G., and Murray, D. Parallel tracking and mapping for small AR workspaces. In *6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007. ISMAR 2007* (2007), pp. 1–10.

[113] Klippel, A., Freska, C., and Winter, S. You-are-here maps in emergencies-: The danger of getting lost. *Journal of spatial science 51*, 1 (2006), 117–131.

[114] Kray, C., and Baus, J. A survey of mobile guides. In *Workshop on HCI in mobile guides, 5th Int. Symposium on HCI with Mobile Devices and Services. Udine, Italy* (2003).

[115] Kray, C., Baus, J., Zimmer, H., Speiser, H., and Krüger, A. Two path preposition: Along and past. *Spatial Information Theory* (2001), 263–277.

[116] Krieger, M., Stark, E. M., and Klemmer, S. R. Coordinating tasks on the commons: designing for personal goals, expertise and serendipity. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems* (New York, NY, USA, 2009), ACM, pp. 1485–1494.

[117] Krüger, A., Baus, J., Heckmann, D., Kruppa, M., and Wasinger, R. *Adaptive Mobile Guides*, vol. 4321. Springer, Berlin, 2007, pp. 521–549.

[118] Krüger, A., Butz, A., Müller, C., Stahl, C., Wasinger, R., Steinberg, K.-E., and Dirschl, A. The connected user interface: realizing a personal situated navigation service. In *IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces* (New York, NY, USA, 2004), ACM, pp. 161–168.

[119] Kuhn, W. Semantic reference systems. *International Journal of Geographical Information Science 17*, 5 (2003), 405–409.

[120] Küpper, A. *Location-based services*. Wiley, 2005.

[121] Lee, S., Buxton, W., and Smith, K. A multi-touch three dimensional touch-sensitive tablet. *ACM SIGCHI Bulletin 16*, 4 (1985), 21–25.

[122] Lehikoinen, J., and Suomela, R. WalkMap: developing an augmented reality map application for wearable computers. *Virtual Reality 6*, 1 (2002), 33–44.

[123] Levine, M. You-Are-Here Maps: Psychological Considerations. *Environment and Behavior 14*, 2 (1982), 221.

[124] Lewis, G. Eearthcaching – An Adventure Game with an Earthy Twist. *GSA Today 15*, 1 (2005), 38–38.

[125] Li, C., and Willis, K. Modeling context aware interaction for wayfinding using mobile devices. In *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (New York, NY, USA, 2006), ACM, pp. 97–100.

[126] Lifton, J., and Paradiso, J. Dual Reality: Merging the Real and Virtual. In *Proceedings of the First International ICST Conference on Facets of Virtual Environments (FaVE)* (2007).

[127] Löchtefeld, M., Schöning, J., Rohs, M., and Krüger, A. LittleProjectedPlanet: An Augmented Reality Game for Camera Projector Phones. In *Mobile Interaction with the Real World (MIRW 2009), Mobile HCI Workshop* (2009).

[128] Low, K., Welch, G., Lastra, A., and Fuchs, H. Life-sized projector-based dioramas. In *Proceedings of the ACM symposium on Virtual reality software and technology* (2001), ACM New York, NY, USA, pp. 93–101.

[129] Lowe, D. Object recognition from local scale-invariant features. *International Conference on Computer Vision 2* (1999), 1150–1157.

[130] Lynch, K. *The image of the city.* MIT Press, 1960.

[131] MacEachren, A. *How maps work.* Guilford Press New York, 1995.

[132] Maceachren, A., and Brewer, I. Developing a Conceptual Framework for Visually-enabled Geocollaboration. *International Journal of Geographical Information Science 18*, 1 (2004), 1–34.

[133] Makri, A., Arsenijevic, D., Weidenhausen, J., Eschler, P., Stricker, D., Machui, O., Fernandes, C., Maria, S., Voss, G., and Ioannidis, N. ULTRA: An Augmented Reality system for handheld platforms, targeting industrial maintenance applications. In *11th International Conference on Virtual Systems and Multimedia, Ghent, Belgium* (2005).

[134] Malik, S., and Laszlo, J. Visual touchpad: a two-handed gestural input device. *Proceedings of the 6th international conference on Multimodal interfaces* (2004), 289–296.

[135] Mallos, J. B. US patent #4,346,376: Touch position sensitive surface.

[136] Matherson, L., Wright, V., Inman, C., and Wilson, E. Get Up, Get Out with Geocaching: Engaging Technology for the Social Studies Classroom. *Social Studies Research and Practice 3*, 3 (2008).

[137] Matsushita, N., and Rekimoto, J. HoloWall: designing a finger, hand, body, and object sensitive wall. *Proceedings of the 10th annual ACM symposium on User interface software and technology* (1997), 209–210.

[138] Mayrhofer, R., Gellersen, H., and Hazas, M. An Authentication Protocol using Ultrasonic Ranging. Tech. Rep. COMP-002-2006, Lancaster University, October 2006.

[139] McGonigal, J. ?This Is Not a Game?: Immersive Aesthetics and Collective Play. In *Melbourne DAC 2003 Streamingworlds Conference Proceedings* (2003), Citeseer, pp. 116–25.

[140] McGonigal, J. Why I Love Bees: A Case Study in Collective Intelligence Gaming. *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning* (2007), 199–227.

[141] Mehra, S., Werkhoven, P., and Worring, M. Navigating on handheld displays: Dynamic versus static peephole navigation. *ACM Transactions on Computer-Human Interaction (TOCHI) 13*, 4 (2006), 457.

[142] Milgram, P., and Kishino, F. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems 77*, 12 (1994), 1321–1329.

[143] Miller, N. Wikipedia and the Disappearing" Author". *ETC.: A Review of General Semantics 62*, 1 (2005), 37–41.

[144] Mine, M. Virtual Environments Interaction Technqiues. Tech. Rep. TR95-018, UNC Chapel Hill Computer Science, 1995.

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

[146] Morris, J., and Hirst, G. Non-classical lexical semantic relations. In *Workshop on Computational Lexical Semantics, Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics* (2004), pp. 46–51.

[147] Moscovich, T. Multi-touch interaction. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2006), ACM, pp. 1775–1778.

[148] Moscovich, T., and Hughes, J. F. Indirect mappings of multi-touch input using one and two hands. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 1275–1284.

[149] Motamedi, N. Hd touch: multi-touch and object sensing on a high definition lcd tv. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 3069–3074.

[150] Mueller, R. E. US patent #3,846,826: Direct television drawing and image manipulation system.

[151] Mulloni, A., Wagner, D., Barakonyi, I., and Schmalstieg, D. Indoor positioning and navigation with camera phones. *IEEE Pervasive Computing 8*, 2 (2009), 22–31.

[152] Murray, J. *Hamlet on the Holodeck.* MIT press, 1998.

[153] Myers, B. A. A Taxonomy of Window Manager User Interfaces. *IEEE Comput. Graph. Appl. 8*, 5 (1988), 65–84.

[154] Newman, J., Schall, G., Barakonyi, I., Schürzinger, A., and Schmalstieg, D. Wide-Area Tracking Tools for Augmented Reality. *Advances in Pervasive Computing 207* (2006), 2006.

[155] Newman, J., Schall, G., and Schmalstieg, D. Modelling and handling seams in wide-area sensor networks. In *Proc. of ISWC* (2006).

[156] Nielsen, J. Participation Inequality: Encouraging more Users to Contribute. *Jakob Nielsen's Alertbox 9* (2006).

[157] Nielsen, J., and Bellcore, M. Iterative user-interface design. *Computer 26*, 11 (1993), 32–41.

[158] Norrie, M., and Signer, B. Overlaying paper maps with digital information services for tourists. *Information and Communication Technologies in Tourism 2005* (2005), 23–33.

[159] O'Hara, K. Understanding geocaching practices and motivations. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 1177–1186.

[160] Olwal, A. LightSense: enabling spatially aware handheld interaction devices. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR 2006)* (2006), pp. 119–122.

[161] O'Neill, M. Theory and research in design of'you-are-here'maps. *Visual information for everyday use-design and research perspectives, Taylor & Francis, London* (1999), 225–238.

[162] Oviatt, S. Mulitmodal Interactive Maps: Designing for Human Performance. *Human-Computer Interaction 12*, 1 & 2 (1997), 93–129.

[163] Paelke, V., and Brenner, C. Development of a mixed reality device for interactive on-site geo-visualization. In *Proceedings of 18th Simulation and Visualization Conference* (2007).

[164] Paelke, V., and Reimann, C. Vision-Based Interaction–A First Glance at Playing MR Games in the Real-World Around Us. *Proc. of PerGame Workshop in conjunction with PERVASIVE '05* (2005).

[165] Palen, L., and Liu, S. Citizen Communications in Crisis: Anticipating a Future of ICT–supported Public Participation. *In Proc. of CHI* (2007), 727–736.

[166] Papay, G. Funktionen der kartographischen Darstellungsformen. *Petermanns Geographische Mitteilungen 117*, 3 (1973), 234–239.

[167] Paradiso, J., and Leo, C. Tracking and Characterizing Knocks Atop Large Interactive Displays. *Sensor Review 25*, 2 (2005), 134–143.

[168] Parkinson, B., and Spilker, J. *Global Positioning System: Theory and Applications.* Aiaa, 1996.

[169] PELTONEN, P., KURVINEN, E., SALOVAARA, A., JACUCCI, G., ILMONEN, T., EVANS, J., OULASVIRTA, A., AND SAARIKKO, P. It's mine, don't touch!: interactions at a large multi-touch display in a city centre. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 1285–1294.

[170] PETERS, J. W. *Complete Idiot's Guide to Geocaching.* Alpel Publishing, 2004.

[171] PICKLES, J. *A history of spaces: cartographic reason, mapping, and the geo-coded world.* Routledge, 2004.

[172] PIEKARSKI, W., AND THOMAS, B. Tinmith-Metro: new outdoor techniques for creating city models withan augmented reality wearable computer. In *Wearable Computers, 2001. Proceedings. Fifth International Symposium on* (2001), pp. 31–38.

[173] PIERCE, J., FORSBERG, A., CONWAY, M., HONG, S., ZELEZNIK, R., AND MINE, M. Image Plane Interaction Techniques in 3D Immersive Environments. In *ACM Symposium on Interactive 3D Graphics* (1997), pp. 39–44.

[174] PIPER, B., RATTI, C., AND ISHII, H. Illuminating clay: a 3-d tangible interface for landscape analysis. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems* (2002), ACM, pp. 355–362.

[175] PREECE, J., ROGERS, Y., AND SHARP, H. Interaction design: beyond human-computer interaction.

[176] RAPER, J., GARTNER, G., KARIMI, H., AND RIZOS, C. A critical evaluation of location based services and their potential. *Journal of Location Based Services 1*, 1 (2007), 5–45.

[177] RASKAR, R., BEARDSLEY, P., VAN BAAR, J., WANG, Y., DIETZ, P., LEE, J., LEIGH, D., AND WILLWACHER, T. RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors. *ACM Transactions on Graphics (TOG) 23*, 3 (2004), 406–415.

[178] RASKAR, R., VAN BAAR, J., BEARDSLEY, P., WILLWACHER, T., RAO, S., AND FORLINES, C. iLamps: geometrically aware and self-configuring projectors. In *ACM SIGGRAPH 2005 Courses* (2005), ACM, p. 5.

[179] RASKAR, R., WELCH, G., LOW, K., AND BANDYOPADHYAY, D. Shader Lamps: Animating real objects with image-based illumination. In *Rendering Techniques 2001: Proceedings of the Eurographics Workshop in London, United Kingdom, June 25-27, 2001* (2001), Springer Verlag Wien, p. 89.

[180] RATH, O., SCHÖNING, J., ROHS, M., AND KRÜGER, A. Sight quest: A mobile game for paper maps. In *Intertain 2008: Adjunct Proceedings of the 2nd International Conference on INtelligent TEchnologies for interactive enterTAINment* (2008).

[181] REILLY, D., RODGERS, M., ARGUE, R., NUNES, M., AND INKPEN, K. Marked-up maps: combining paper maps and electronic information resources. *Personal and Ubiquitous Computing 10*, 4 (2006), 215–226.

[182] Reilly, D., Welsman-Dinelle, M., Bate, C., and Inkpen, K. Just point and click?: using handhelds to interact with paper maps. In *MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices & services* (New York, NY, USA, 2005), ACM, pp. 239–242.

[183] Reimann, C., and Paelke, V. Computer vision based interaction techniques for mobile augmented reality. In *Proc. 5th Paderborn Workshop Augmented and Virtual Reality in der Produktentstehung* (2006), HNI, pp. 355–362.

[184] Reitmayr, G., and Drummond, T. Going out: Robust model-based tracking for outdoor augmented reality. In *Proceedings of 5th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2006)* (2006), pp. 109–118.

[185] Reitmayr, G., Eade, E., and Drummond, T. Localisation and interaction for augmented maps. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality* (2005), IEEE Computer Society Washington, DC, USA, pp. 120–129.

[186] Reitmayr, G., and Schmalstieg, D. Collaborative augmented reality for outdoor navigation and information browsing. In *Proc. Symposium Location Based Services and TeleCartography* (2004), pp. 31–41.

[187] Rekimoto, J. NaviCam- A magnifying glass approach to augmented reality. *Presence: Teleoperators and Virtual Environments 6*, 4 (1997), 399–412.

[188] Rekimoto, J. Smartskin: an infrastructure for freehand manipulation on interactive surfaces. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2002), ACM, pp. 113–120.

[189] Riegelsberger, J., and Nakhimovsky, Y. Seeing the bigger picture: a multi-method field trial of google maps for mobile. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 2221–2228.

[190] Robinson, A. Elements of cartography. *Soil Science 90*, 2 (1960), 147.

[191] Rohs, M. Marker-Based Embodied Interaction for Handheld Augmented Reality Games. *Journal of Virtual Reality and Broadcasting 4*, 5 (2007).

[192] Rohs, M., Schleicher, R., Schöning, J., Essl, G., Naumann, A., and Krüger, A. Impact of item density on the utility of visual context in magic lens interactions. *Personal and Ubiquitous Computing 13*, 8 (2009), 633–646.

[193] Rohs, M., Schöning, J., Krüger, A., and Hecht, B. Towards real-time markerless tracking of magic lenses on paper maps. *Adjunct Proceedings of the 5th Intl. Conference on Pervasive Computing (Pervasive), Late Breaking Results* (2007), 69–72.

[194] Rohs, M., Schöning, J., Raubal, M., Essl, G., and Krüger, A. Map navigation with mobile devices: virtual versus physical movement with and without visual context. In *Proceedings of the 9th international conference on Multimodal interfaces* (2007), ACM New York, NY, USA, pp. 146–153.

[195] Rolland, J. P., and Fuchs, H. Optical versus video see-through head-mounted displays. In *in Medical Visualization. Presence: Teleoperators and Virtual Environments* (2000), pp. 287–309.

[196] Salzman, T., Stachniak, S., and Stürzlinger, W. Unconstrained vs. Constrained 3D Scene Manipulation. In *8th IFIP International Conference on Engineering for Human-Computer Interaction* (2001), vol. 2254.

[197] Sanneblad, J., and Holmquist, L. Ubiquitous graphics: combining hand-held and wall-size displays to interact with large images. In *Proceedings of the working conference on Advanced visual interfaces* (2006), ACM, p. 377.

[198] Sareika, M., and Schmalstieg, D. Urban Sketcher: Mixed Reality on Site for Urban Planning and Architecture. In *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality* (2007), IEEE Computer Society, pp. 1–4.

[199] Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reitinger, B., and Schmalstieg, D. Handheld Augmented Reality for underground infrastructure visualization. *Personal and Ubiquitous Computing* (2008), 1–11.

[200] Schall, G., and Schmalstieg, D. Interactive Urban Models Generated from Context-Preserving Transcoding of Real-Wold Data. *Proceedings of the 5th International Conference on GIScience (GISCIENCE 2008)* (2008).

[201] Schlatter, B., and Hurd, A. Geocaching-21st-century hide-and-seek. *Journal of Physical Education, Recreation & Dance 76*, 7 (2005), 28–32.

[202] Schmalstieg, D., and Reitmayr, G. Augmented Reality as a Medium for Cartography Multimedia Cartography, 2006.

[203] Schmidt, A. Ubiquitous Computing–Computing in Context. *Lancaster University, UK* (2002).

[204] Schöning, J., Bartindale, T., Olivier, P., Jackson, D., Krüger, A., and Kitson, J. iBookmark: locative texts and place-based authoring. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (2009), ACM New York, NY, USA, pp. 3775–3780.

[205] Schöning, J., Brandl, P., Daiber, F., Echtler, F., Hilliges, O., Hook, J., Löchtefeld, M., Motamedi, N., Muller, L., Olivier, P., Roth, T., and von Zadow, U. Multi-touch surfaces: A technical guide. Tech. rep., Technical University of Munich, 2008.

[206] Schöning, J., Hecht, B., Raubal, M., Krüger, A., Marsh, M., and Rohs, M. Improving interaction with virtual globes through spatial thinking: helping users ask "why?". In *IUI '08: Proceedings of the 13th international conference on Intelligent user interfaces* (New York, NY, USA, 2008), ACM, pp. 129–138.

[207] Schöning, J., Hecht, B., Rohs, M., and Starosielski, N. WikEar- Automatically Generated Location-Based Audio Stories between Public City Maps. *Adjunct Proc. of Ubicomp 07* (2007), 128–131.

[208] Schöning, J., Hecht, B., and Starosielski, N. Evaluating automatically generated location-based stories for tourists. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 2937–2942.

[209] Schöning, J., Krüger, A., Cheverst, K., Rohs, M., Löchtefeld, M., and Taher, F. PhotoMap: using spontaneously taken images of public maps for pedestrian navigation tasks on mobile devices. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (2009), ACM, pp. 14–24.

[210] Schöning, J., Krüger, A., and Müller, H. Interaction of mobile camera devices with physical maps. In *Adjunct Proceeding of the Fourth International Conference on Pervasive Computing* (2006), pp. 121–124.

[211] Schöning, J., Krüger, A., and Müller, H. J. Interaction of mobile devices with maps. In *Advances in Pervasive Computing 2006. Adjunct Proceedings of the 4th International Conference on Pervasive Computing* (2006), vol. 27, Oesterreichische Computer Gesellschaft.

[212] Schöning, J., Krüger, A., and Olivier, P. Multi-touch is dead, long live multi-touch. *CHI 2009: Workshop on Multi-touch and Surface Computing* (2009).

[213] Schöning, J., Rohs, M., Kratz, S., Löchtefeld, M., and Krüger, A. Map torchlight: a mobile augmented reality camera projector unit. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (2009), ACM New York, NY, USA, pp. 3841–3846.

[214] Schöning, J., Rohs, M., and Krüger, A. Mobile interaction with the real world. In *MIRW 2008:Workshop on Mobile Interaction with the Real World* (2008).

[215] Schöning, J., Rohs, M., Krüger, A., and Stasch, C. Improving the Communication of Spatial Information in Crisis Response by Combining Paper Maps and Mobile Devices. *Lecture Notes In Computer Science* (2009), 57–65.

[216] Shakespeare, W. *King Henry V.* Classic Books Company, 2001.

[217] Sharma, R., Poddar, I., Ozyildiz, E., Kettebekov, S., Kim, H., and Huang, T. Toward Interpretation of Natural Speech/Gesture: Spatial Planning on a Virtual Map. *Proceedings of ARL Advanced Displays Annual Symposium* (1999), 35–9.

[218] Sherman, E. *Geocaching: Hike and Seek with Your GPS.* APress, 2004.

[219] Steinicke, F., Ropinski, T., Bruder, G., and Hinrichs, K. Interscopic User Interface Concepts for Fish Tank Virtual Reality Systems. In *Proceedings of the Virtual Reality* (2007), pp. 27–34.

[220] Steinicke, F., Schöning, J., Krüger, A., and Hinrichs, K. Multi-touching cross-dimensional data:towards direct interaction in stereoscopic display environments coupled with mobile devices. In *AVI 2008: Workshop on designing multi-touch interaction techniques for coupled private and public displays* (May 2008).

[221] Sun Microsystems, Inc. . Java platform, micro edition, `http://java.sun.com/javame/index.jsp`, 2009. [Online; accessed 6-February-2009].

[222] Surdick, R. T., Davis, E. T., King, R. A., and Hodges, L. F. The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. *Presence 6*, 5 (1997), 513–531.

[223] Szalavri, Z., and Gervautz, M. The personal interaction panel - a two-handed interface for augmented reality. In *COMPUTER GRAPHICS FORUM* (1997), pp. 335–346.

[224] Szulborski, D. *This is not a game: A guide to alternate reality gaming.* New-Fiction Publishing, 2005.

[225] Telaar, D. Geocaching - Eine kontextuelle Untersuchung der deutschsprachigen Geocaching-Community. Master's thesis, University of Münster, Germany, 2007.

[226] Tversky, B. Cognitive Maps, Cognitive Collages, and Spatial Mental Models. *In Proc. of COSIT 2003* (1993), 14–34.

[227] Ullmer, B., and Ishii, H. The metadesk: models and prototypes for tangible user interfaces. In *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology* (1997), ACM, pp. 223–232.

[228] Ullmer, B., and Ishii, H. Emerging frameworks for tangible user interfaces. *IBM Systems Journal 39*, 3 (2000), 915–931.

[229] Urquhart, K., Cartwright, W., Miller, S., Mitchell, K., Quirion, C., and Benda, P. Ensuring useful cartographic representations in location-based services. In *Proceedings 21st International Cartographic Conference* (2003), pp. 10–16.

[230] Valli, A., and Linari, L. Natural interaction sensitivetable. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 2315–2318.

[231] Veas, E., and Kruijff, E. VespAR: design and evaluation of a handheld AR device. In *7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 2008. ISMAR 2008* (2008), pp. 43–52.

[232] Wagner, D., Pintaric, T., Ledermann, F., and Schmalstieg, D. Towards Massively Multi-User Augmented Reality on Handheld Devices. *Third International Conference on Pervasive Computing* (2005), 208–219.

[233] Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., and Schmalstieg, D. Pose tracking from natural features on mobile phones. In *7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 2008. ISMAR 2008* (2008), pp. 125–134.

[234] Wagner, D., and Schmalstieg, D. First Steps towards Handheld Augmented Reality. *Proceedings of the 7th IEEE International Symposium on Wearable Computers* (2003), 127–135.

[235] Wagner, D., and Schmalstieg, D. Artoolkitplus for pose tracking on mobile devices. In *Computer Vision Winter Workshop* (2007), pp. 6–8.

[236] Want, R., Schilit, B., Adams, N., Gold, R., Petersen, K., Goldberg, D., Ellis, J., and Weiser, M. The PARCTAB ubiquitous computing experiment. *KLUWER INTERNATIONAL SERIES IN ENGINEERING AND COMPUTER SCIENCE* (1996), 45–97.

[237] Wasinger, R., and Krüger, A. Multi-modal interaction with mobile navigation systems. *Special Journal Issue Conversational User Interfaces, it - Information Technology 46*, 6 (2004), 322–331.

[238] Wasinger, R., Stahl, C., and Krüger, A. M3i in a pedestrian navigation exploration system. In *Proceedings of the 5th International Symposium on Human-Computer Interaction with Mobile Devices and Services (Mobile HCI). Udine* (Udine, Italy, 2003), L. Chittaro, Ed., Springer, pp. 481–485.

[239] Weiser, M. The computer for the 21 stcentury. *ACM SIGMOBILE Mobile Computing and Communications Review 3*, 3 (1999), 3–11.

[240] Werner, S., and Habel, C. Spatial reference systems. *Spatial Cognition and Computation 1*, 4 (1999), 3–7.

[241] White, R. M. US patent #4,484,179: Tactile sensor employing a light conducting element and a resiliently deformable sheet.

[242] Wikipedia. Geocoin — Wikipedia, The Free Encyclopedia, 2008. [Online; accessed 10-April-2009].

[243] Wikipedia. Indium tin oxide — wikipedia, the free encyclopedia, 2008. [Online; accessed 28-July-2008].

[244] Wikipedia. Geocaching — Wikipedia, The Free Encyclopedia, 2009. [Online; accessed 16-April-2009].

[245] Wilson, A. D., Izadi, S., Hilliges, O., Garcia-Mendoza, A., and Kirk, D. Bringing physics to the surface. In *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology* (New York, NY, USA, 2008), ACM, pp. 67–76.

[246] Woodward, C., Lahti, J., Rönkkö, J., Honkamaa, P., Hakkarainen, M., Jäppinen, J., Rainio, K., Siltanen, S., and Hyväkkä, J. Case Digitalo-A range of virtual and augmented reality solutions in construction application. In *24th W78 Conference, Maribor. In: 24th W*, vol. 78, Citeseer, pp. 529–540.

[247] Worboys, M., and Duckham, M. *GIS: A Computing Perspective, Second Edition*, 2 ed. CRC, 5 2004.

[248] Worboys, M., and Duckham, M. *GIS: A Computing Perspective, Second Edition*, 2 ed. CRC, 5 2004.

[249] Wu, M., and Balakrishnan, R. Multi-finger and Whole Hand Gestural Interaction Techniques for Multi-user Tabletop Displays. *Proceedings of the 16th annual ACM symposium on User interface software and technology* (2003), 193–202.

[250] Yee, K. Peephole displays: pen interaction on spatially aware handheld computers. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2003), ACM New York, NY, USA, pp. 1–8.

[251] Yuan, Y., and Detlor, B. Intelligent Mobile Crisis Response Systems. *Commun. ACM 48*, 2 (2005), 95–98.

[252] Zipf, A. User-Adaptive Maps for Location-Based Services (LBS) for Tourism. In *Information and Communication Technologies in Tourism 2002: Proceedings of the International Conference in Innsbruck, Austria, 2002* (2002), Springer.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen oder indirekt bernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

Saarbrücken, 22.02.2010

Apendix

- Vita of Johannes Schöning
- Full list of articles of Johannes Schöning

curriculum vitæ

Johannes Schöning

Date of Birth	October 4th 1982
Place of birth	in Georgsmarienhütte, Germany
Nationality	German
Address	Mozartstr. 22 66111 Saarbrücken

Education

02.09.1989–01.06.1993	Elementary school Bad Laer, Germany
02.08.1993–31.07.1995	Two-grade middle school Bad Laer, Germany
04.08.1995–12.06.2002	Secondary school Bad Iburg. Received Abitur (A-level) qualifying for university admission (grade 1.7, 1 best - 5 worst)
01.10.2002–01.08.2003	Civilian service in the "DJK - Sportschule" in Münster
Languages	English, Latin

Studies

01.10.2003-01.07.2007	Diploma in Geoinformatics at the Institute for Geoinformatics, University of Münster, Germany
01.04.2006	Received Pre-diploma ($\sim$ Bachelor) in Geoinformatics (grade 1.6, 1 best - 5 worst)
01.04.2005-01.04.2007	Teaching assistent as for "Digital Cartography" and research assistant in the group of Prof. Antonio Krüger
01.10.2005-01.07.2007	Student research assistant in the group of Prof. Dr. Antonio Krüger
01.12.2006–31.05.2007	Diploma theses *Interaction of Mobile Devices with Maps*; Advisors: Prof. Dr. Antonio Krüger (Uni Münster) and Prof. Dr. Martin Raubal (University of California at Santa Barbara)
01.02.2007–31.05.2007	Semester abroad at the *University of California at Santa Barbara* to complete diploma thesis with the support of DAAD (German Academic Exchange Service)
01.07.2007	Diploma in Geoinformatics (grade 1.2, 1 best - 5 worst)
01.10.2007-30.4.2009	Researcher and PhD Student at the Institute for Geoinformatics, University of Münster, Germany
since 01.05.2009	Researcher at the DFKI, Saarbrücken (German Research Center for Artificial Intelligence).

Voluntary academic services

Organization of Workshops: Multi-touch Workshop 2006, Multi-touch Bootcamp 2007,Advanced Navigation 2007, Advanced Navigation 2008, TIPUGG 2008, LocWeb 2009, Next generation Interfaces for Education 2010, Ubiprojection 2010 and LocWeb 2010.

Program Committee: MIRW 2009, AGILE 2009, Fun & Games 2008, LocWeb 2008, AGILE 2009.

Reviewing: UIST, CHI, Ubicomp, Pervasive, IUI, MobileHCI, INTERACT, and others.

Chairs: Demo Chair ITS 2009, General Chair ITS 2010.

Selected Publications

52 peer reviewed articles in Conference Proceedings, Journals and Books (a full list of articles is attached.)

May 2006 Johannes Schöning, Antonio Krüger & Hans Jörg Müller: *Interaction of Mobile Camera Devices with Physical Maps.* Pervasive 2006: Adjunct Proceedings of the 4th International Conference on Pervasive Computing, (2006).

May 2007 Michael Rohs, Johannes Schöning, Antonio Krüger & Brent Hecht: *Towards - Real-Time Markerless Tracking of Magic Lenses on Paper Maps.* Pervasive 2007: Adjunct Proceedings of the 5th International Conference on Pervasive Computing, (2007).

September 2007 Michael Rohs, Johannes Schöning, Martin Raubal, Georg Essl & Antonio Krüger: Map Navigation with Mobile Devices: *Virtual versus Physical Movement with and without Visual Context.* ICMI 2007: Proceedings of the 9th International Conference on Multimodal Interfaces, (2007).

January 2008 Johannes Schöning, Brent Hecht, Martin Raubal, Antonio Krüger, Meri Marsh & Michael Rohs: *Improving Interaction with Virtual Globes through Spatial Thinking: Helping users Ask "Why?".* IUI 2008: Proceedings of the International Conference on Intelligent User Interfaces, (2008).

September 2009 Johannes Schöning, Keith Cheverst, Markus Löchtefeld, Antonio Krüger, Michael Rohs & Faisal Taher: *Photomap: Using Spontaneously taken Images of Public Maps for Pedestrian Navigation Tasks on Mobile Devices.* Mobile HCI 2009: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Service, (2009).

Awards / Grands

September 2006 Best Poster Presentation, GIScience2006, Münster, Germany

Oktober 2006	2nd Place, MotoFRWD2006, Frankfurt, Germany
Februar 2007	DAAD Scholarship for visiting UCSB, USA
September 2008	Best Extend Abstract, GiScience2006, Salt Lake City, USA.
September 2008	Travel Scholarship. GIScience 2008: Oak Ridge National Lab.
Januar 2009	Nokia Contest, Calling All Innovators: Region Winner Europe.
May 2009	BMBF Contest Winner “Alltagstauglich”
September 2009	Best Paper Award, Mobile HCI, Bonn, Germany

Voluntary services

1998–2002	Schooling as a mediator at secondary school Bad Iburg, Germay
1999–2002	Youth-group leader in fold Saint ”Mariä Geburt” Bad Laer, Germany
2000–2007	Youth-group leader for the Deutscher Alpenverein (German Moutaineering Foundation)
2002	Volunteer at the European Championships in Athletics, Munich, Germany
2003	Schooling as high rope course trainer
2000 and 2005	Volunteer at the World Youth Day in Rome and Cologne
2006, 2007	Student volunteers at various international conferences

Hobbies

Mountaineering, climbing, perform theater plays, rugby, travelling, reading

J. Schöning

Saarbrücken, December 14, 2009

Publication record: Johannes Schöning

The list contains all publications (minimum 2 reviews)

1. Johannes **Schöning**, Antonio Krüger & Hans Jörg Müller: Interaction of Mobile Camera Devices with Physical Maps. Pervasive 2006: Adjunct Proceedings of the 4th International Conference on Pervasive Computing, (2006)

2. Johannes **Schöning**, Jan Torben Heuer, Hans Jörg Müller & Antonio Krüger: The Marauders Lens. Proc. of the 4th International Conference on GIScience, Extended Abstracts, (2006)

3. Hans Jörg Müller, Johannes **Schöning** & Antonio Krüger: Mobile Mobile Map Interaction - Evaluation in an Indoor Scenario. Workshop on Mobile and Embedded Interactive Systems, Informatik 2006 Gesellschaft für Informatik e.V., (2006)

4. Johannes **Schöning**, Ilja Panov & Carsten Keßler: No Vertical Limit -Conceptual Design of a LBS for climbers. Workshop Mobile Spatial Interaction (MSI) on CHI 2007, (2007)

5. Krzysztof Janowicz & Johannes **Schöning**: Mobile Map Interaction for Local News. Workshop Mobile Spatial Interaction (MSI) on CHI 2007, (2007)

6. Brent Hecht, Michael Rohs, Johannes **Schöning** & Antonio Krüger: Wikeye - Using Magic Lenses to Explore Georeferenced Wikipedia Content. Pervasive 2007: Workshop on Pervasive Mobile Interaction Devices (PERMID), (2007)

7. Michael Rohs, Johannes **Schöning**, Antonio Krüger & Brent Hecht: Towards Real-Time Markerless Tracking of Magic Lenses on Paper Maps. Pervasive 2007: Adjunct Proceedings of the 5th International Conference on Pervasive Computing, (2007)

8. Michael Rohs, Johannes **Schöning**, Martin Raubal, Georg Essl & Antonio Krüger: Map Navigation with Mobile Devices: Virtual versus Physical Movement with and without Visual Context. ICMI 2007: Proceedings of the 9th International Conference on Multimodal Interfaces, (2007)

9. Johannes **Schöning**: Interaktion von mobilen Geräten mit öffentlichen statischen Karten. Kartographische Nachrichten (in german) Issue 10/07, (2007)

10. Johannes **Schöning**, Brent Hecht, Michael Rohs & Nicole Starosielski: WikEar: Automatically Generated Location-Based Audio Stories between Public City Maps. Ubicomp 2007: Adjunct Proceedings of the 9th International Conference on Ubiquitous Computing, (2007)

11. Johannes **Schöning**, Michael Rohs & Antonio Krüger: Paper Maps as an Entry Point for Tourists to Explore Wikipedia content. ICMI 2007: Adjunct Proceedings of the 9th International Conference on Multimodal Interfaces, (2007)

12. Oliver Rath, Johannes **Schöning**, Michael Rohs & Antonio Krüger: Sight Quest: A Mobile Game for Paper Maps. Intertain 2008: Adjunct Proceedings of the 2nd International Conference on INtelligent Technologies for interactive enterTAINment , (2008)

13. Johannes **Schöning**, Brent Hecht, Martin Raubal, Antonio Krüger, Meri Marsh & Michael Rohs: Improving Interaction with Virtual Globes through Spatial Thinking: Helping users Ask „Why?“. IUI 2008: Proceedings of the International Conference on Intelligent User Interfaces, (2008)

14. Johannes **Schöning**, Brent Hecht & Nicole Starosielski: Evaluating Automatically Generated Location-Based Stories for Tourists. CHI 2008: Adjunct Proceedings of the 26th International Conference on Human Factors in Computing Systems, (2008)

15. Jana Gliet, Antonio Krüger, Johannes **Schöning** & Otto Klemm: Image Geo-Mashups: The Example of an Augmented Reality Weather Cam. AVI 2008: Proceedings of the 9th International Conference on Advanced Visual Interfaces, (2008)

16. Johannes **Schöning**, Michael Rohs & Antonio Krüger: Using Mobile Phones to Spontaneously Authenticate and Interact with Multi-Touch Surfaces. AVI 2008: Workshop on designing multi-touch interaction techniques for coupled private and public displays PPD , (2008)

17. Frank Steinicke, Johannes **Schöning**, Antonio Krüger, Klaus Hinrichs: Multi-Touching Cross-Dimensional Data: Towards Direct Interaction in Stereoscopic Display Environments coupled with Mobile Devices. AVI 2008: Workshop on designing multi-touch interaction techniques for coupled private and public displays PPD , (2008)

18. Johannes Schöning, Michael Rohs, Antonio Krüger & Christoph Stasch: Improving the Communication of Spatial Information in Crisis Response by Combining Paper Maps and Mobile Devices. Mobile Response 2008: Proceedings of International Symposium on Mobile Information Technology for Emergency Response, (2008)

19. Johannes **Schöning**, Michael Rohs & Antonio Krüger: Mobile Interaction with the „Real World“. Mobile HCI 2008: Workshop on Mobile Interaction with the Real World MIRW, (2008)

20. Keith Cheverst, Johannes **Schöning**, Antonio Krüger & Michael Rohs: Photomap: Snap, Grab and Walk away with a ‘YOU ARE HERE’ Map. Mobile HCI 2008: Workshop on Mobile Interaction with the Real World MIRW, (2008)

21. Johannes **Schöning** & Antonio Krüger: Multi-Modal Navigation through Spatial Information. GIScience 2008: Proc. of the 4th International Conference on GIScience, Extended Abstracts, (2008)

22. Brent Hecht & Johannes **Schöning**: Mapping the Zeitgeist. GIScience 2008: Proc. of the 4th International Conference on GIScience, Extended Abstracts, (2008)

23. Johannes **Schöning**, Michael Rohs & Antonio Krüger: Spatial Authentication on Large Interactive Multi-Touch Surfaces. IEEE Tabetop 2008: Adjunct Proceedings of IEEE Tabletops and Interactie Surfaces, (2008)

24. Johannes **Schöning**, Florian Daiber & Antonio Krüger: Advanced Navigation Techniques for Spatial Information Using Whole Body Motion. HCI 2008: Workshop on Whole Body Interation: The Future of the Human Body. Whole Body Interaction II, (2008).

25. Johannes **Schöning**, Peter Brandl, Florian Daiber, Florian Echtler, Otmar Hilliges, Jonathan Hook, Markus Löchtefeld, Nima Motamedi, Tim Roth, David Smith, Ulrich von

Zadow: "Build your Own" Multi-touch Surface: Bootcamp on Construction & Implementation of Multi-touch Surfaces. IEEE Tabetop 2008: Proceedings of IEEE Tabletops and Interactie Surfaces, (2008)

26. Peter Konopatzky, Michael Scholz, Sven Reißig, Johannes **Schöning**, Philipp Verhoeven, Markus Löchtefeld & Antonio Krüger: xChase: A Location-Based Multi-User Pervasive Game Using a Lightweight Tracking Framework. FnG 2008: Proceedings of International Conference on Fun and Games, (2008)

27. Johannes **Schöning**, Peter Brandl, Florian Daiber, Florian Echtler, Otmar Hilliges, Jonathan Hook, Markus Löchtefeld, Nima Motamedi, Laurence Muller, Patrick Olivier, Tim Roth, Ulrich von Zadow: Multi-Touch Surfaces: A Technical Guide. Technical Report TUM-I0833: Technical Reports of the Technical University of Munich, (2008)

28. Johannes **Schöning**, Antonio Krüger & Patrick Olivier: Multi-Touch is Dead, Long live multi-touch. CHI 2009: Workshop on Multi-touch and Surface Computing, (2009)

29. Thomas Bartoschek & Johannes **Schöning**: Trends und Potentiale von virtuellen Globen in Schule und Wissenschaft. GIS Buisness, (2008)

30. Erik Wilde, Susanne Boll, Keith Cheverst, Peter Fröhlich, Ross Purves & Johannes **Schöning**: Location and the Web (LocWeb 2009). CHI 2009: Adjunct Proceedings of the 27th International Conference on Human Factors in Computing Systems, (2009)

31. Frank Steinicke, Johannes **Schöning**, Antonio Krüger, Klaus Hinrichs: Interscopic Multi-Touch Surfaces: Using bimanual Interaction for intuitive Manipulation of Spatial Data. IEEE 3DUI 2009: Proceedings of the IEEE Symposium on 3D User Interfaces, (2009)

32. Johannes **Schöning**, Markus Löchtefeld, Keith Cheverst & Antonio Krüger: Photomap Nokia Contest, Calling All Innovators: Region Winner Europe presented at MWC Barcelona, (2009)

33. Florian Daiber, Christoph Stasch, Alexander C. Walkowski, Johannes **Schöning** & Antonio Krüger: Multi-Touch- und Multi-User-Interaktion zur Verbesserung des kollaborativen Arbeitens in Katastrophenstäben. Geoinformatik 2009: In Proc. of Geoinformatik 2009 (german only), (2009)

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

35. Johannes **Schöning**, Florian Daiber, Antonio Krüger & Michael Rohs: Using Hands and Feet to Navigate and Manipulate Spatial Data. CHI 2009: Adjunct Proceedings of the 27th International Conference on Human Factors in Computing Systems, (2009)

36. Johannes **Schöning**, Tom Bartindale, Patrick Olivier, Dan Jackson, Antonio Krüger & Jim Kitson: iBookmark: Locative Texts and Place-based Authoring. CHI 2009: Adjunct Proceedings of the 27th International Conference on Human Factors in Computing Systems, (2009)

37. Florian Daiber, Johannes **Schöning**, Antonio Krüger: Whole Body Interaction with Geospatial Data. Smart Graphics 2009: Proceedings of the 9th International Symposium on Smart Graphics, (2009)

38. Michael Rohs, Robert Schleicher, Johannes **Schöning**, Georg Essl, Anja Naumann & Antonio Krüger: Impact of Item Density on the Utility of Visual Context in Magic Lens Interactions. PUC 4/2009: Personal and Ubiquitous Computing, (2009)

39. Johannes **Schöning**, Frank Steinicke, Antonio Krüger, Klaus Hinrichs: Bimanual Interaction with Interscopic Multi-Touch Surfaces. INTERACT 2009: 12th IFIP TC13 Conference in Human-Computer Interaction , (2009)

40. Michael Rohs, Johannes **Schöning**, Robert Schleicher, Georg Essl, Anja Naumann & Antonio Krüger: Impact of Item Density on Magic Lens Interactions. Mobile HCI 2009: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Service, (2009)

41. Johannes **Schöning**, Keith Cheverst, Markus Löchtefeld, Antonio Krüger & Michael Rohs: Photomap: Using Spontaneously taken Images of Public Maps for Pedestrian Navigation Tasks on Mobile Devices. Mobile HCI 2009: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Service, (2009)

42. Michael Rohs, Johannes **Schöning**, Robert Schleicher, Georg Essl, Anja Naumann & Antonio Krüger: Impact of Item Density on Magic Lens Interactions. Mobile HCI 2009: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Service, (2009)

43. Benjamin Proß, Johannes **Schöning** & Antonio Krüger: iPiccer: Automatically Retrieving and Inferring Tagged Location Information from Web Repositories. Mobile HCI 2009: Adjunct Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Service, (2009)

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

45. Johannes **Schöning**: Do we need further multi-touch affordances? Touch Affordances: Workshop on Touch Affordances in conjunction with INTERACT, (2009)

46. Stefan Below, Thorsten Diekhof, Jörg Possin, Mareike Kritzler, Johannes **Schöning** & Antonio Krüger: GeoInformationRouter - Localization in WLAN Environments Via Infrastructure. Ubicomp 2009: Adjunct Proceedings of the 11th International Conference on Ubiquitous Computing, (2009)

47. Johannes **Schöning**: Advanced Natural and Tangible Interfaces for Spatial Information. Ubicomp 2009: Adjunct Proceedings of the 11th International Conference on Ubiquitous Computing Doctoral Consortium, (2009)

48. Gerrit Kahl, Karin Leichtenstern, Johannes **Schöning**, Lübomira Spassova & Antonio Krüger: A Contextual Learning Game for Toddlers Installed on an Interactive Display Attached to a Shopping Cart. PerED 2009: Workshop in adjunction with the 11th International Conference on Ubiquitous Computing, (2009)

49. Gerhard Schall, Johannes **Schöning**, Volker Paelke & Georg Gartner: Augmented Maps & Environments: Approaches, Applications & Interactions. In "Advances in Web-based GIS, Mapping Services and Applications", ISPRS book by CRC Press/ Taylor and Francis, 2009, (2009)

50. Robert Schleicher, Michael Rohs & Johannes **Schöning**: Eye movements when using a mobile phone to search on wall maps. ECEM 2009: Proceedings of the 15th European Conference on Eye Movements, (2009)

51. Lübomira Spassova, Johannes **Schöning**, Gerrit Kahl & Antonio Krüger: Innovative Retail Laboratory. AmI 2009: 3rd European Conference on Ambient Intelligence, (2009)

52. Johannes **Schöning**, Jonathan Hook, Nima Motamedi, Patrick Olivier, Florian Echtler, Peter Brandl, Laurence Muller, Florian Daiber, Otmar Hilliges, Markus Löchtefeld, Tim Roth, Dominik Schmidt & Ulrich von Zadow: Building Interactive Multi-touch Surfaces. JGT 2009: Journal of Graphics Tools, (2009)

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

54. Edward Tse, Johannes Schöning, Yvonne Rogers, Chia Shen, Gerald Morrison: Next Generation of HCI and Education: Workshop on UI Technologies and Educational Pedagogy. CHI 2010: Adjunct Proceedings of the 27th International Conference on Human Factors in Computing Systems, (2010)

55. Sven Gehring, Markus Löchtefeld, Johannes **Schöning**, Dominic Gorecky, Peter Stephan, Antonio Krüger & Michael Rohs: Mobile Product Customization. CHI 2010: Adjunct Proceedings of the 27th International Conference on Human Factors in Computing Systems, (2010)

56. Max Peiffer, Dagmar Kern, Johannes **Schöning**, Tanja Döring, Antonio Krüger & Albecht Schmidt: A Multi-Touch Enabled Steering Wheel – Exploring the Design Space. CHI 2010: Adjunct Proceedings of the 27th International Conference on Human Factors in Computing Systems, (2010)