Pasi Välkkynen

Physical Selection in Ubiquitous Computing

Physical Selection in Ubiquitous Computing

Pasi Välkkynen

Thesis for the degree of Doctor of Philosophy to be presented with due permission for public examination and criticism in Pinni Building, Auditorium B1097 at University of Tampere, on the 30th of November 2007 at 12 o'clock noon.

Supervisor: Professor Roope Raisamo, PhD
                Department of Computer Sciences
                University of Tampere
                Finland

Opponent: Associate Professor Morten Fjeld, Dr. Tech
                Computer Science and Engineering
                Chalmers University of Technology
                Sweden

Reviewers: Professor Jukka Riekki, Dr. Tech
                Department of Electrical and Information Engineering
                University of Oulu
                Finland

                Research Leader Juha Lehikoinen. PhD
                Nokia Research Center
                Nokia
                Finland

Abstract

In ubiquitous computing, the computing devices are embedded into the physical environment so that the users can interact with the devices at the same time as they interact with the physical environment. The various devices are connected to each other, and have various sizes and input and output capabilities depending on their purpose. These features of ubiquitous computing create a need for interaction methods that are radically different from the desktop computer interactions.

Physical selection is an interaction task for ubiquitous computing and it is used to tell the user's mobile terminal which physical object the user wants to interact with. It is based on tags that identify physical objects or store a physical hyperlink to digital information related to the object the tag is attached to. The user selects the physical hyperlink by touching, pointing or scanning the tag with the mobile terminal that is equipped with an appropriate reader. Physical selection has been implemented with various technologies, such as radio-frequency tags and readers, infrared transceivers, and optically readable tags and mobile phone cameras.

In this dissertation, physical selection is analysed as a user interaction task, and from the implementation viewpoint. Different selection methods – touching, pointing and scanning – are presented. Touching and pointing have been studied by implementing a prototype and conducting user experiments with it. The contributions of this dissertation include an analysis of physical selection in the ubiquitous computing context, suggestions for visualising the physical hyperlinks in both the physical environment and in the mobile terminal, and user requirements for physical selection as a part of an ambient intelligence architecture.

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Roope Raisamo from University of Tampere. His encouragement, support and advice were essential for completing this work.

I was honoured to have Professor Jukka Riekki from University of Oulu and Research Leader Juha Lehikoinen from Nokia as the pre-examiners of my dissertation. Their constructive comments were helpful, especially in clarifying the contribution of my work.

I am particularly grateful for Ilkka Korhonen for directing my PhD studies in VTT, especially in the early phases of this work when I was not yet sure where I should be heading. Eija Kaasinen supported my work with helpful comments and advice. I want to address my thanks to both for continuous support, and for advice during the writing of individual publications, many of which we wrote together.

I want to address my special thanks to my other colleagues at VTT whose help and support influenced my work greatly. Timo Tuomisto has been my closest associate during the past years and most of the publications in this dissertation have been published with him. I would not be at this point without his experience and advice. I also want to thank Marketta Niemelä for her enthusiasm and knowledge, and for helping me in this work in so many ways. Heikki Ailisto and Lauri Pohjanheimo are my co-authors in several publications and the source of many insightful discussions.

I want to thank my other colleagues from VTT who have worked with me in the Mimosa, Minami and Physical Browsing in Ambient Intelligence (PB-AmI) projects: Katja Rentto, Veikko Ikonen, Valtteri Pirttilä, Heikki Seppä, Johan Plomp and Luc Cluitmans. They have provided me with uncountable ideas for this work. I am also thankful for the support of the User Interfaces research group: our group leader Timo Urhemaa, my office roommate Jussi Mattila, and Juhani Heinilä. I am grateful for Jukka Perälä for supporting my leave of absence for finalising this dissertation, and making sure it received funding from VTT. Many of the figures in this dissertation were created initially for the PB-AmI project by Tiina Kymäläinen and have been used to illustrate the ideas in several publications.

Contents

Appendix II of this publication is not included in the PDF version.
Please order the printed version to get the complete publication
(http://www.vtt.fi/publications/index.jsp)

List of Publications

This dissertation is based on the following publications, which are reproduced here by permission:

I   Pasi Välkkynen, Ilkka Korhonen, Johan Plomp, Timo Tuomisto, Luc Cluitmans, Heikki Ailisto and Heikki Seppä, A user interaction paradigm for physical browsing and near-object control based on tags, in *Proc. Physical Interaction Workshop on Real World User Interfaces*, 2003, 31–34.

II   Heikki Ailisto, Lauri Pohjanheimo, Pasi Välkkynen, Esko Strömmer, Timo Tuomisto and Ilkka Korhonen, Bridging the physical and virtual worlds by local connectivity-based physical selection, in *Personal and Ubiquitous Computing* 10(6), Springer-Verlag London, 2006, 333–344.

III   Pasi Välkkynen, Lauri Pohjanheimo and Heikki Ailisto, Physical Browsing, in Thanos Vasilakos and Witold Pedrycz (eds.), *Ambient Intelligence, Wireless Networking, and Ubiquitous Computing*, Artech House, 2006, 61–81.

IV   Timo Tuomisto, Pasi Välkkynen and Arto Ylisaukko-Oja, RFID tag reader system emulator to support touching, pointing and scanning, in A. Ferscha, R. Mayrhofer, T. Strang, C. Linnhoff-Popien, A. Dey, A. Butz and A. Schmidt (eds.), *Advances in Pervasive Computing, Adjunct Proceedings of Pervasive 2005*, Austrian Computer Society, 2005, 85–88.

V   Pasi Välkkynen, Marketta Niemelä and Timo Tuomisto. Evaluating touching and pointing with a mobile terminal for physical browsing. *Proc. 4th Nordic Conference on Human-computer Interaction*. ACM Press, 2006, 28–37.

VI   Pasi Välkkynen, Timo Tuomisto and Ilkka Korhonen, Suggestions for visualising physical hyperlinks, in *Proc. Pervasive Mobile Interaction Devices*, 2006, 245–254.

VII   Pasi Välkkynen. Hovering: visualising RFID hyperlinks in a mobile phone. *Proc. Mobile Interaction with the Real World*, 2006, 27–29.

VIII   Eija Kaasinen, Timo Tuomisto and Pasi Välkkynen, Ambient functionality – use cases, in *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies*, ACM Press, 2005, 51–56.

IX   Eija Kaasinen, Marketta Niemelä, Timo Tuomisto, Pasi Välkkynen and Vladimir Ermolov, Identifying User Requirements for a Mobile Terminal Centric Ubiquitous Computing Architecture, in *Proc. Int. Workshop on System Support for Future Mobile Computing Applications*, IEEE, 2006, 9–16.

The papers are referred to in the text by the above Roman numerals. I was the main contributor to publications I, III and V–VII. My role in paper II was the analysis of physical browsing from the user's point of view; in paper IV in the user interaction design of the described system; and in papers VIII and IX the user requirements for physical selection as a part of the described architecture, in addition to participating in the whole writing process.

List of symbols

AR   Augmented Reality

Auto-ID   Automatic Identification

GPS   Global Positioning System

GUI   Graphical User Interface

HCI   Human-Computer Interaction

HF   High Frequency

ID   Identification or Identifier

IR   Infrared

IrDA   Infrared Data Association

NFC   Near-Field Communication

OCR   Optical Character Recognition

PDA   Personal Digital Assistant

RF   Radio Frequency

RFID   Radio Frequency Identification

TUI   Tangible User Interface

Ubicom   Ubiquitous Computing and Communication

UHF   Ultra High Frequency

UI   User Interface

URI   Universal Resource Identifier

URL   Universal Resource Locator

WLAN   Wireless Local Area Network

WPAN   Wireless Personal Area Network

1. Introduction

Physical browsing is an interaction paradigm that allows associating digital information with physical objects. It can be seen analogous to browsing the World Wide Web: the physical environment contains links to digital information and by selecting these physical hyperlinks, various services can be activated.

In physical browsing, the interaction happens via a *mobile terminal* – such as a mobile phone or a Personal Digital Assistant (PDA). The links are implemented as *tags* that can be read with the terminal, for example Radio Frequency Identifier (RFID) tags that are read with a mobile phone augmented with an RFID reader. The basis of physical browsing is *physical selection* – the interaction task with which the user tells the mobile terminal which link the user is interested in and wants to activate. After the selection an *action* occurs, for example, if the tag contains a Universal Resource Identifier (URI, "web address"), the mobile phone may display the associated web page in the browser. Optimally, the displayed information is somehow related to the physical object itself, creating an association between the object and its digital counterpart. This user interaction paradigm is best illustrated with a simple scenario [Välkkynen *et al.,* 2006]:

> Joe has just arrived on a bus stop on his way home. He touches the bus stop sign with his mobile phone and the phone loads and displays him a web page that tells him the expected waiting times for the next buses so he can best decide which one to use and how long he must wait for it. While he is waiting for the next bus, he notices a poster advertising a new interesting movie. Joe points his mobile phone at a link in the poster and his mobile phone displays the web page of the movie. He decides to go see it in the premiere and "clicks" another link in the poster, leading him to the ticket reservation service of a local movie theatre.

This scenario illustrates some aspects of physical browsing and physical selection compared to some other interaction paradigms:

1. Joe can select physical hyperlinks using physical gestures like touching and pointing at them instead of navigating through the menus of his device.

2. He has a mobile interaction device, which acts as a medium for the digital services from the links instead of the links activating on an environmental display.

3. Joe uses the mobile terminal for touching and pointing at tags, instead of touching and pointing with his bare hands.

4. The interaction is explicit, that is, Joe touches and points at things instead of activating the services implicitly just by being in the vicinity of the bus stop.

5. Finally, the services he activates (bus schedule, movie web page, ticket reservation) are all related to the physical objects and the location around him.

The list above introduces the basic concept of physical selection: a mobile terminal and tag based interaction technique, which is intended for interacting with the physical world and its entities.

The research questions discussed in this dissertation are the following:

- How to provide the user with a usable way to select an object in a physical environment in order to access the digital counterpart of the physical object?

- What kinds of selection methods are needed in physical selection?

- How should they be implemented from the point of view of the user?

- How should the physical hyperlinks be visualised so that they support user-friendly physical selection?

The research method includes purely analytical method of studying the available technologies and applying knowledge from human-computer interaction (HCI) literature to deduce some basic principles for user-friendly physical selection and how physical selection can be applied in a broader context. In addition to that, we have built prototypes and studied empirically the feasibility of physical selection and some details of the different selection methods. As this dissertation is a part of an ongoing work, it is not possible to provide definite answers to all themes in this dissertation. The further research, which will hopefully shed more light on the questions raised in this dissertation, is outlined in the discussion in

Chapter 8. However, in this dissertation it is already possible to contribute a body of knowledge regarding the selection methods, how they fit in the broader context, how they work in some detail and how they can be implemented.

The main content of this dissertation consists of nine published papers. In papers I, III, V, VI and VII, I have been the primary or only author. In paper II, I discuss physical browsing in general and my contribution in the paper consists primarily of discussing physical selection and the user interaction in physical browsing. In paper IV is described a prototype system in which my main contribution is within the design of the user interaction. In papers VIII and IX, a methodology and user requirements for a ubiquitous computing architecture that includes physical selection as an interaction paradigm, is described. My main contribution in those papers is in the sections specifying requirements for physical selection. The main contributions of this dissertation are:

1. a physical browsing and physical selection framework including different selection methods, actions and technologies, suitable for a variety of tasks in ubiquitous computing,

2. analysis on how to implement the different selection methods,

3. comparison of different techniques for touching and pointing,

4. suggestions for visualising the links in both the physical environment and in the mobile terminal using the hovering concept, and

5. user requirements for physical selection within a ubiquitous computing architecture.

The structure of this dissertation is as follows. In Chapter 2 the role of physical selection in the context of ubiquitous computing is discussed and in Chapter 3 some relevant ubiquitous computing technologies are introduced. In Chapter 4 the evolution of interfaces in which physical and digital worlds are connected is discussed and in Chapter 5 is presented a selection of existing systems in which physical selection is applied. In Chapter 6 selection is discussed as an interaction task in various environments and in Chapter 7 the publications and their themes are introduced. In Chapter 8 my work is related to that of others'. The dissertation is concluded in Chapter 9.

2. Physical Selection in Ubiquitous Computing Context

In this Chapter, the concept of ubiquitous computing is discussed briefly, and how physical browsing relates to it. Next, two other similar broad concepts, computer-augmented environments and physically based user interfaces, are presented. Finally, the concepts and vocabulary of physical selection and the different selection methods are introduced.

2.1 Ubiquitous Computing

Mark Weiser described his idea of ubiquitous computing as a "computing environment in which each person is continually interacting with hundreds of nearby wirelessly interconnected computers" [Weiser, 1991], and later as "making many computers available throughout the physical environment, while making them effectively invisible to the user." [Weiser, 1993] The goal of Weiser was to achieve "the most effective kind of technology, that which is essentially invisible to the user." This invisibility is not to be taken literally: the user can still see the technology if he wants to, but the ubiquitous computers disappear like text in our current surroundings; not invisible but out of our minds until needed.

Weiser [1991] wrote that "The most profound technologies are those that disappear." The constant background presence of writing – the first information technology, which is now ubiquitous in industrialised countries – "does not require active attention, but the information is conveyed ready for use at a glance." Computer-based information technology, in contrast, had not yet become part of the environment (and at the time of writing this, sixteen years later, it still has not reached that level in a large scale). Weiser thought that the idea of a personal computer itself is misplaced and that isolated portable and hand-held computers are only a transitional step toward achieving the real potential of information technology. Such machines cannot truly make computing an integral, invisible part of the way people live their lives. Therefore Weiser was trying to conceive a new way of thinking about computers in the world, one that takes into account the natural human environment and allows the computers themselves

to vanish into the background. Ubiquitous computing does not just mean computers that can be carried around. A mobile phone or a PDA alone, even with network access is not a ubiquitous computer, because the focus of the interaction is still the one device.

According to Weiser's visions [1991; 1993], ubiquitous computers will come in different sizes, each suited to a particular task. Weiser classified several research prototypes he had been involved in into three categories, tabs, pads and boards. Tabs are the smallest components of ubiquitous computing, expanding on the usefulness of existing inch-scale computers. The Active Badge Location System [Want *et al.*, 1992] is an example of a tab-based system. The next step up in size is the pad, similar to mobile and laptop computers, intended not to be personal but available to anyone who happens to pick one up [Weiser, 1993]. Yard-size displays, boards, such as the Liveboard [Elrod *et al.*, 1992] or DynaWall [Streitz *et al.*, 1999], are the largest class of ubiquitous computers[1].

The essential features of ubiquitous computing are thus [Weiser, 1991; Weiser, 1993; Abowd and Mynatt, 2000]:

- computing embedded into the physical environment so that the user can interact with the computers at the same time as interacting with the physical environment,

- computers disappearing from the active focus of the user, until they are needed,

- computers connected to each other, and

- computers having various sizes, and input and output capabilities depending on their purpose.

---

[1] Researchers have traditionally tried to categorise their and others' work into these three categories, but as Weiser himself stated [1993], these prototypes were just examples for practical reasons, and the real power of the concept comes not from any one of these devices; instead it emerges from the interaction of all of various devices. Therefore, in this dissertation, I will consciously not try to fit the physical browsing devices narrowly between tabs and pads.

The list above illustrates some features of physical selection as an interaction technique for ubiquitous computing. The tags enable embedding computing into the physical environment, and optimally the tags also disappear from the active focus until they are needed. If we see tags just as simple identifiers or distributed memories, the only real computer in physical selection is the mobile terminal, typically connected to other computers via a network and using a tag reader to complement its other input capabilities.

Coulouris *et al.* [2005] describe ubiquitous computing as a special case of a *distributed system* because of the requirement of computers communicating with each other. A distributed system is one in which "components located at networked computers communicate and co-ordinate their actions only by passing messages". Their definition of *mobile computing* is also related to ubiquitous computing: "the performance of computing tasks while the user is on the move, or visiting places other than her usual environment". This mobile (and at the same time ubiquitous) computing is made possible by the miniaturisation and portability of computational devices.

Abowd and Mynatt [2000] state that distributing computing into the physical space creates the desire to break away from the traditional desktop-based human-computer interaction, which will change the relationship between humans and computers. Continual interaction changes the computers from a localised tool to a constant companion.

2.2 Computer-Augmented Environments

Computer-Augmented Environment is a concept very similar to ubiquitous computing. If the name is taken literally, it means an environment that is augmented with computers or devices with computational capabilities – exactly as one aspect of ubiquitous computing. However, computer-augmented environments grew from the combination of ubiquitous computing and augmented reality (AR), but the terms have become more or less intermingled. Common to both approaches is the emphasis on the physical world and the tools that enhance our everyday activities [Wellner *et al.*, 1993]. The important issue is that computer-augmented environments research has produced several ideas and prototypes that include similar ideas as physical browsing.

Augmented interaction is a term with which Rekimoto and Nagao [1995] described interaction in computer-augmented environments. It is a style of HCI that aims to reduce computer manipulations by using environmental information as implicit input. The approach of Rekimoto and Nagao was based on mobile computers that can read for example tags and display information on a transparent screen, overlaying the information with real-world objects.

Common to ubiquitous computing and augmented interaction is that they both aim to create a computer augmented real world environment instead of building a virtual environment in the computer. According to Rekimoto and Nagao [1995], the main difference between ubiquitous computing and augmented interaction is the number of computers; augmented interaction is designed as an interaction paradigm for a single portable or wearable computer that uses context information as implicit commands. Ubiquitous computing bridges the gap between physical and virtual worlds by spreading a large number of computers around the environment [Weiser, 1991]. The two approaches are complementary and can support each other. Rekimoto and Nagao see mobile terminals as personal assistants that can help the user in a ubiquitous computing environment.

Because of its emphasis on mobile terminals and tag-based interaction, augmented interaction is very similar to physical browsing. However, augmented interaction differs from physical browsing in the explicit-implicit axis. Physical browsing is explicit, that is, the user has to touch, point or scan an object to select it. In augmented interaction, the mobile terminal implicitly reads information from the environment and displays it to the user, which is similar to the notifying selection method.

2.3 Physically Based User Interfaces

Another concept close to ubiquitous computing and computer-augmented environments is physically based user interfaces. In physically-based user interfaces the interaction is based on computationally augmented physical artefacts. Tangible user interfaces (TUI) are closely related to physically-based interfaces, being based on tangible and graspable objects as interaction devices. The main idea of tangible interfaces is that physical objects can at the same time act as input devices to control digital information, and as output devices to present that digital information.

With regards to physical selection, we are more interested about the concept of computationally augmenting physical artefacts than making the artefacts act as input and output devices at the same time. However, much research has been conducted in the field of tangible user interfaces, that is relevant to physical selection as well, and therefore I will discuss tangible user interfaces in a separate Chapter.

Want *et al.* [1999] describe the goal in physically-based user interfaces as "*to seamlessly blend the affordances and strengths of physically manipulatable objects with virtual environments or artifacts, thereby leveraging the particular strengths of each*". In this integration, physical input artefacts are linked to electronic graphical objects. Want *et al.* state that physically-based user interfaces support our everyday tools and objects and the affordances of these objects, and they augment them computationally. This allows casual interaction with these augmented everyday tools, using "natural manipulations and associations".

2.4 Physical Browsing and Ubiquitous Computing

An important issue in ubiquitous computing is how to interact with the devices and in particular how to transfer information between the devices in an intuitive and direct way [Streitz and Russell, 1998]. The directness is an important motivation for physical selection, which is intended to be more direct than menu based interactions.

Want *et al*. [1998] list three key problems in ubiquitous computing:

1. Locating people and objects at low cost and reliably.
2. Coordination of active objects into a single user interface. There are two further challenges within this problem:
    2.1. What are the expectations of the user about the linkage of the physical objects; what are the semantics the user associates with the objects?
    2.2. How can those semantics be delivered?
3. Coordination of real and virtual objects.

Items 2 and 3 of the list relate directly to physical selection. Physical selection relies heavily on the expectations of the user about the linkage of physical and digital objects. The visualisation of physical hyperlinks can help with directing the user's expectations. Additionally, coordination between the physical and digital worlds is a challenge in physical selection.

According to Abowd and Mynatt [2000], ubiquitous computing suggests new paradigms of interaction. They see three interaction themes in ubiquitous computing:

1. natural interfaces,
2. context-aware applications, and
3. automated capture and access.

Ubiquitous computing gets the applications away from the desktop [Abowd and Mynatt, 2000]. This will make the interaction between humans and computers unlike the desktop paradigm with keyboard, mouse and display interactions, and more similar to the way humans interact with the physical world. There will be a rich variety of communication capabilities between humans and computation. Graspable, or tangible, interfaces is one class of natural interfaces. Other forms of natural interfaces include speech and pen input, gaze-based interfaces and gesture recognition. Handling other forms of input as easily as keyboard and mouse input is important for enabling rapid development of applications with effective natural interfaces. Because of the recognition-based input, natural interfaces introduce a new set of problems: they permit new and various mistakes. However, physical selection is not a recognition-based input technique. Instead, typical reading of a tag is very unambiguous compared to recognition-based techniques such as gesture recognition or speech input.

Ailisto *et al.* [2003a] see two different approaches to ubiquitous computing and communication (Ubicom): distributed and mobile terminal centred. In the distributed approach, there are inexpensive, smart widgets embedded into appliances and even into people. In the mobile terminal centred approach, at the centre is the user who carries a mobile terminal (mobile phone, PDA or even a wrist computer). The terminal is connected to the local environment and also has global connectivity. It is already common to carry a hand-held device with nearly always-on networking connectivity and significant computational capabilities [Patel and Abowd, 2003]. Physical browsing is essentially part of the

latter, mobile terminal centred, approach, even though the tags are embedded in the environment.

Ailisto *et al.* [2003a] divide paradigms related to the user interfaces and ubiquitous computing into three types:

1. Natural and disappearing user interfaces;
2. User interfaces relying on displays and input devices in the local infrastructure; and
3. User interfaces using mobile personal devices.

They see paradigms one and two linked with the distributed ubiquitous computing model and the third one with the mobile terminal centred model, into which also physical browsing fits. It should be noted that these paradigms are overlapping, and they can complement each other. Ailisto *et al.* [2003a] say that using the user interface of a mobile terminal in Ubicom application systems is a natural part of the mobile terminal-centred Ubicom model. The mobile terminal is not only used for external or long-range communication, but also for local ubiquitous services. As an example, they state that a user can choose the address of a device by proximity or active pointing – physical selection. Physical browsing and physical selection are thus part of interaction in a ubiquitous computing environment, and they also depend on the ubiquitous computing technologies, which I will discuss in the next Chapter.

2.5 Physical Selection

The main topic within physical browsing in this dissertation is physical selection, that is, how the user accomplishes the selection task with a mobile terminal and tags. For physical selection, we have introduced four selection methods: touching, pointing, scanning and notifying[2].

---

[2] These selection methods are also known as TouchMe, PointMe, ScanMe and NotifyMe, respectively.

2.5.1 Concepts and Vocabulary

A *tag* is the technical device that a *reader* embedded in a mobile terminal can read. There are several implementation technologies for tags and their corresponding readers, for example RFID tags and RFID readers, which are introduced in a later Chapter.

The tag implements a *link*, or physical hyperlink. Link is something the user sees and interacts with, by physical selection. Thus, the *user selects links*, and the *reader reads tags*, in the terms of this dissertation. Optimally, the user should not have to know implementation details of the tag, but the mobile terminal hardware would take care of the reading details. In real world with varying technologies with incompatible readers and varying reading ranges this goal is not normally met.

Physical selection, which is the main topic of this dissertation, is an interaction task. By physical selection, the users tell the terminal with which physical object they want to interact.

After selection, typically some *action* occurs. This action can be displaying information related to the physical object to which the tag is attached, or it may be any digital service the terminal is capable of offering. Selection and action are independent of each other [Paper II] and the selection method should not affect the action. Also Bowman and Hodges [1997] state that it is important to consider grabbing (selection) and manipulation (action) as separate issues.

Physical browsing means in this dissertation the whole process of browsing in ubiquitous computing environment. When the user has a need to accomplish a task, or she notices an interesting link, she selects the link with the mobile terminal and after the selection an action occurs.

2.5.2 Touching

Touching is a centimetres-range selection method in which the user brings the terminal very close to the tag to select the tag. The implementation of touching requires that either the mobile terminal or the tag can sense the proximity of the

other. If the communication range of the implementation technology is very short, as for example in near-field RFID (see Chapter 3 for further details), touching is implemented 'for free'. If the communication range of the implementation technology is longer, there has to be a technique for determining which tag is closest to the reader. That can be accomplished in various ways, for example by power analysis or by periodically limiting the communication range to a few centimetres.
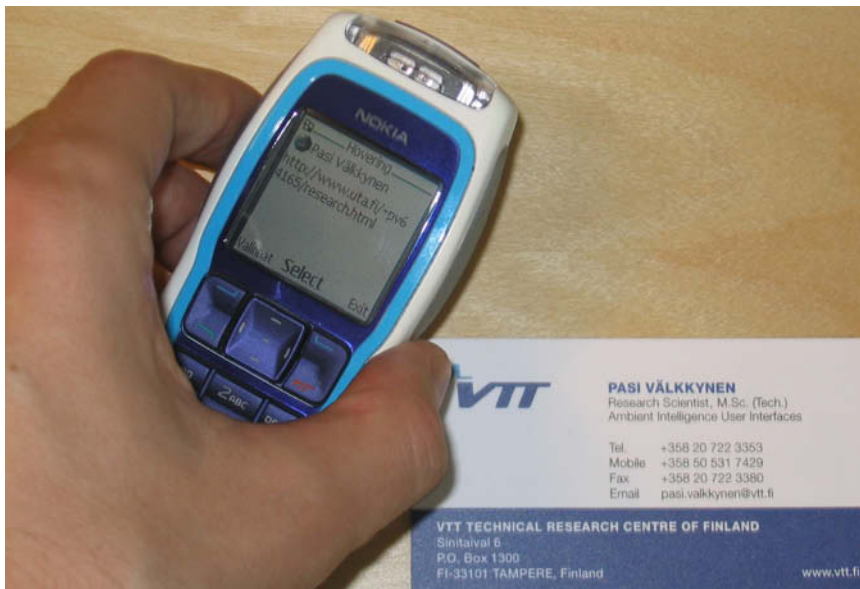
Figure 1. Touching a link embedded in a business card with a mobile phone.

Touching is the most direct and unambiguous of the selection methods. It is suitable for selecting links near the user, and in environments in which the link density is high. When touching a link, the user has to know the exact location of the tag and even the placement of the reader inside the terminal.

2.5.3 Pointing

Pointing is a long-range selection method that is suitable for selecting one tag from a distance. The distance depends on the communication range of the implementation technology, but ideally it should be several metres [Välkkynen *et al.*, 2006] or as much as is needed for the user to select a visible tag. To point at a link, the user has to know the exact location of the link.

Figure 2. User pointing at a link in a movie poster with a PDA.

Implementation of pointing is heavily dependent on the technology. Some technologies, such as infra-red, are easy to make directional whereas some, such as RFID, typically have more or less uniform spherical reading field and special methods are required to determine the relative alignment of the reader and the tag.

2.5.4 Scanning

Scanning is a long-range selection method that combines the physical location of the user and the graphical user interface of the mobile terminal. When the user scans for links, all links within the reading range are displayed on a list and the user will be able to make the final selection of the desired link by selecting it from the list. Scanning for nearby Bluetooth devices with a mobile phone is a familiar example of scanning. The implementation of scanning requires a long, preferably several metres, reading range and preferably uniform spherical reading field so that all tags within range will be read.

Figure 3. Scanning for services within the immediate environment.

Scanning is suitable for situations in which the user does not know the exact locations of the links, or what is available in the environment. Touching and pointing are more direct selection methods when the link location and content are known, and scanning is intended to cover the other situations. One of the main challenges with scanning is the naming of the links for display so that the user can predict what action will follow from selecting the link.

2.5.5 Notifying

Notifying means tag or terminal initiated selection. A tag can be active and broadcast its presence to all readers that are listening, or a reader can periodically search for tags in the vicinity and alert the user if something is found. This selection method seems somewhat similar to scanning, but the important difference is that scanning is initiated by an explicit request from the user, and in notifying a technological device initiates the selection.

Notify is listed here for the sake of completeness. We have not implemented notify in any of our prototypes because it is not a user-initiated selection method. This kind of active selection method would have to be studied thoroughly because the user does not initiate the selection. Some questions that would need to be answered include when the terminal may offer a link to the user, how the links are filtered to provide appropriate links preferably related to the context of the user and how to keep the user in control of what links are offered.

2.6 Summary

Physical selection is based on a mobile terminal and tags embedded in the environment and physical objects. The user can select a physical object for interaction by using the mobile terminal to read a link to digital information from the tag. In the following Chapters various aspects of physical selection are discussed. In Chapter 3, selected local connectivity technologies and how they either enable physical selection, or benefit from adding physical selection to their procedures, are described. Then in Chapter 4, the relation of tangible user interfaces and physical selection from the point of view of combining the digital and physical worlds are discussed. In Chapter 5, some previous systems that have used physical selection in their interactions are described. The selection interaction task and how it has been used in various environments is presented in Chapter 6. In Chapter 7, the themes of the publications in this dissertation are introduced, and in Chapter 8 various aspects of a user-friendly physical selection system are discussed. Chapter 9 concludes this dissertation.

3. Selected Technologies Related to Physical Selection

In this Chapter, a selection of ubiquitous computing technologies that are related to physical browsing, are discussed. These technologies allow identification of physical objects and storing data or links to data to enable physical selection. Wireless local connectivity technologies enable further actions after physical selection, for example connecting the mobile terminal to an external device via Bluetooth. On the other hand, some wireless technologies benefit from physical selection, especially when the normal connection method is to search for all devices within range (scanning) and then connect to them via the graphical user interface of the mobile terminal.

In ubiquitous computing and communication, users can access different kinds of services via local communication. Local communication usually means "wireless short-range communication links restricted around a single user, inside a room, inside a hall or inside a corresponding space out-of-doors." [Ailisto et al., 2003a] These systems consist of portable user devices that are moving with the humans, and the local infrastructure that consists of fitted items in some environment [Ailisto et al., 2003a]. Therefore the focus in this Chapter is in the local-connectivity technologies, and on technologies that can be used to identify physical objects.

3.1 Automatic Identification Systems

Automatic identification (Auto-ID) procedures emerged from logistics, industry and manufacturing in which it is useful to be able to track objects and containers [Finkenzeller, 2003]. In the Auto-ID procedures, a tag is attached to the tracked object and a reader device is used to read the tag and thus track the object. With the emergence of ubiquitous computing research, tracking physical objects outside the original domains of logistics was found useful as the location of a physical object near the user can often be important context information. Commonly used Auto-ID technologies include Radio Frequency Identification (RFID) and different camera-based technologies.

The common barcode is a binary code comprising a field of bars and gaps arranged in a parallel configuration, in which the bars and gaps represent the data elements [Finkenzeller, 2003]. An example of a barcode-based physical browsing system is WebStickers [Ljungstrand and Holmquist, 1999; Ljungstrand *et al.*, 2000]. Barcodes have later been expanded into two-dimensional matrix codes that typically code their data as black white and squares in a matrix formation. Matrix codes can be read with a mobile phone camera, as demonstrated for example by Rohs and Zweifel [2005]. Optical Character Recognition (OCR) is a technology in which text readable by humans can be read by a machine [Finkenzeller, 2003]. OCR is not specifically related to identification technologies, but it enables physical browsing if the text is for example a web address, email address or some other information the mobile terminal reading the text can use. Close to OCR is the commercial image recognition system developed by Mobot[3]. In Mobot, the user takes a picture of a specific advertisement with a camera phone, sends it to a server for decoding and gets digital material as a response if the advertisement is recognised correctly.

3.2 Radio-Frequency Identification

RFID refers to identification systems in which power and data are transferred without contact. RFID systems consist of two physically separate components [Finkenzeller, 2003]:

1. Transponder, the data-carrying device on the object. The transponder consists of a coupling element and a microchip. The transponder is commonly called "tag" and it is also referred to as tag in this dissertation.

2. Interrogator, a device that transfers power to the tag and reads data from it. The interrogator is commonly called "reader" but it may (depending on the technology) also write data to the tag. In this dissertation, it is referred to as reader.

---

[3] See www.mobot.com

There are many types of RFID, but at the highest level, RFID devices can be divided into two classes: active and passive [Want, 2006]. Active tags use a power source whereas passive tags do not require batteries or maintenance; instead the tag reader powers the tag with its electrical or magnetic field during the data transfer operation. Passive tags have typically very low cost and use very low power. Other differentiation criteria are physical coupling method (electric, magnetic or electromagnetic) and range (from a few millimetres to 15 metres) [Finkenzeller, 2003]. Passive tags are often low-end tags and active tags may have microcontrollers and multiple interfaces to the environment (though passive tags can now have sensors too) [Stanford, 2003].

3.2.1 Principles of RFID Systems

The reading ranges of different RFID systems vary from a few millimetres to up to fifteen metres. The physical coupling method (electric, magnetic or electromagnetic) largely defines the rough scale of the range [Finkenzeller, 2003; Want, 2006]. In close coupling (electric and capacitive) the reading range is in the region of 0.1 cm to 1 cm, in remote coupled systems the range can be up to one metre and long range coupling can reach a range of several metres. From the ubiquitous computing point of view, the two significant technologies are remote coupled and long range. Close coupling allows transferring greater amounts of power to the tag, so even a microprocessor (which consumes more power than a simple microchip) is possible. In close coupling systems the tag is inserted into the reader or placed onto a designated surface on the reader. This physical interaction method somewhat limits the usefulness of close coupling systems in ubiquitous computing applications, which is why close coupling and capacitive coupling are not discussed further in this dissertation.

In remote coupled systems the reading range is up to 1 m. Coupling is typically inductive (magnetic), which means that the tag comprises an electronic data-carrying device, usually a single microchip, and a large area coil that functions as an antenna [Finkenzeller, 2003; Want, 2006]. Inductively coupled tags are almost always passive (all the energy for the microchip has to be provided by the reader). The antenna coil of the reader generates a high-frequency electromagnetic field, which generates voltage in the antenna of the tag by inductance. This coupling is similar to coupling used in the coils of transformers. Data can be

transferred between the tag and the reader by load modulation. The tag can switch a load resistor on and off at its antenna, which modulates the amplitude of the voltage at the antenna of the reader. The reader can then measure the voltage changes in its antenna and get the data signal. Near-Field RFID is based on Faraday's principle of magnetic induction (current creates magnetic field, which in turn creates current). Tags that use near-field coupling send the data back to the reader using load modulation.

Far-field RFID tags capture electromagnetic waves from a dipole antenna attached to the reader [Want, 2006]. A smaller dipole antenna in the tag receives this energy to power the tag. The difference with near-field RFID is that the tag is beyond the reader's near field and information can not be passed back to the reader using load modulation [Finkenzeller, 2003]. Instead, far field RFID uses back scattering and reflects the signal back to the reader. In long range systems the reading range is significantly over 1 m, which makes them suitable for physical selection by pointing and scanning. Long range tags and readers use electromagnetic waves in the Ultra-High Frequency (UHF) and microwave range, and backscattering for reflecting the data back to the reader. Reading range is typically three metres for passive tags and may be over fifteen metres for active tags. The power of the reader is still used for data transmission; battery on the tag is used for powering the microchip in cases in which the reader can not transfer enough power for the chip.

In order to achieve long ranges of up to 15 m or to be able to operate tags with greater power consumption at an acceptable range, backscatter tags often have a backup battery to supply power to the chip of the tag [Finkenzeller, 2003]. Data transmission between the tag and the reader relies exclusively on the power of the electromagnetic field the reader emits. The tag antenna reflects a portion of the electromagnetic power the reader emits back to the reader. By altering the load of its antenna using a chip-controlled load resistor, the tag can transmit data with the reflected power.

RFID tags can be manufactured as paper thin "labels" [Finkenzeller, 2003]. These labels can be printed over and thus support visualisations and even optically readable tags printed on them.

3.2.2 RFID beyond Identification

Conceptualizing RFID tags as simply ID tags underestimates their capabilities: battery-powered tags can have local computing power, persistent storage and communication capabilities [Stanford, 2003]. Even passive tags can contain far more information than a simple ID. They can also convey information that extends beyond data stored in an internal memory and include data that onboard sensors created dynamically.

The energising signal from the reader can carry commands to the tag, including data to be written into the tag [Want, 2004]. The written data can be read from the tag alongside with the ID, making the tag a "memory tag". The tag data can also be sensor readings, but the sensors are somewhat limited, because without an external power source, they can only be powered when the tag is powered, and the available power is very small and available only for a limited time.

Passive RFID tags can be integrated with sensors, which can be used for detecting pointing, in addition to more traditional uses of sensors for measurements (see Figure 4). For example, Opasjumruskit *et al.* [2006] designed a temperature sensor-equipped tag and a companion hand-held reader to experiment with their sensor. Philipose et al. [2005] have built an accelerometer sensor that can be read with an RFID reader. Although these example systems used sensors unsuitable for detecting pointing, the important issue to note is the possibility to interface a sensor with a passive RFID tag, making pointing beam detection possible with a different sensor.
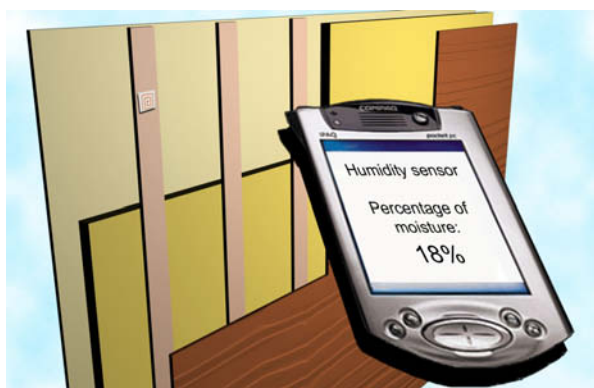
Figure 4. A sensor-equipped RFID tag. RFID tags do not require line-of-sight, allowing for example this kind of hidden humidity sensors inside structures.

Physical selection can utilise both memory tags and sensor-equipped tags. Onboard memory can contain for example a URL and a title for the URL, helping to visualise the tag information in the mobile terminal before the terminal has to query this information from network servers. Sensors can be used to detect pointing, which requires that the reader or the tag knows whether the reader is aligned towards the tag or not. We have experimented with photosensors to implement pointing this way.

3.3 Near-Field Communication

Near-Field Communication (NFC) is a technology for short-range (centimetres) wireless communication between devices [Ecma, 2004]. The devices are typically mobile terminals and consumer electronics devices, but also smart cards and low-frequency RF tags can be read with NFC readers. NFC is intended to "provide a mechanism by which wireless mobile devices can communicate with peer devices in the immediate locality, rather than rely on the discovery mechanisms of popular short-range radio standards" [Want, 2006].

NFC can be used to streamline the discovery process by passing wireless Media Access Control addresses and channel-encryption keys between radios through a near-field coupling side channel, which, when limited to 20 cm, lets users enforce their own physical security for encryption key exchange [Want, 2006]. The benefits are thus easier connection and/or association and more secure communications establishment.

Interaction in Near Field Communication is based on making the objects virtually touch each other, and as such it implements the touch selection method. Philips[4] has defined four "applications" for NFC interaction:

1. **Touch and Go**: device has ticket/access code stored in it and the user brings it close to the reader. Simple data capture, for example picking up a URL from environment, also fits in this category.

---

[4] See http://www.philips.com, http://www.nfc-forum.org.

2. **Touch and Confirm**: the user confirms the interaction by entering a password or otherwise accepting the transaction. This is used in for example mobile payment.

3. **Touch and Connect**: for linking two devices.

4. **Touch and Explore**: explore the capabilities of a device to find out what it offers.

These "applications" actually describe both selection method and action; they could be extended to "Point and Go" or "Scan and Connect" as well. This demonstrates the independence of selection and action: we choose one selection method to select a link and then some action occurs and the action is generally independent of the selection method used. These four actions of NFC demonstrate well the variety of actions available after physical selection. Touch and Explore is especially interesting from the tag visualisation point of view, and the "hovering" concept presented in the visualisation papers is very similar to this.

3.4 Infrared

Infrared communications is based on transceivers that modulate noncoherent infrared light. The communicating transceivers must be within the line of sight of each other either directly or via reflection from a light-coloured surface such as the ceiling of the room [Stallings, 2000].

One well known infrared application is the remote control of entertainment electronics such as television. Typically, the remote control translates the inputs of the user into signals it sends in one direction, from the remote to the device to be controlled. The target device uses an infrared photodiode to detect the infrared light and then converts the infrared signal back into digital patterns. The communication in this case is uni-directional from the remote control device to the target device.

IrDA[5] (Infrared Data Association) is an example of bi-directional infrared communication technology, which is intended as a secure and low-cost replacement for wired communications. The connection range of IrDA is about one metre, but it should be noted that IrDA is only one way to use infrared connectivity and it is possible to have longer ranges with other infrared based technologies.

Compared to RF-based systems, an IR-based implementation offers a different selection method for selecting the target object. The spatial resolution of the infrared beam enables the user to select the target application item by pointing the terminal towards the application item [Ailisto *et al.*, 2003a] and typically pushing one button in the terminal. For example, in IrDA, the connection angle is about thirty degrees, and thus requires deliberately aiming the communicating devices towards each other. After the selection, the object can send its user interface to the terminal [Strömmer and Suojanen, 2003] or some other action may follow.

A drawback of spatial resolution is the line-of-sight requirement between the communicating devices [Ailisto *et al.*, 2003a]. Commonly, further communication between the target object and the terminal also happens using infrared, which requires the user to keep the devices aligned towards each other. On the other hand, the required visibility of the target may be an advantage as the user can see with which device the communication occurs.

Strömmer and Suojanen [2003] have developed a low-power infrared tag and a compatible Terminal Interface Unit (TIU) as a reader. Their goal was to develop a "system for interconnections between smart objects and humans that responds to the challenge of ultra-low power consumption". The TIU can be integrated into mobile terminals and the tag can be integrated into various electronic devices, and it can also operate alone. The latter operation mode allows the tag to act as a memory unit or sensor unit.

An important difference between infrared and RF based transmission is that IR does not penetrate walls [Stallings, 2000]. The security of infrared communications is additionally due to the directionality of the infrared beam, for example thirty

---

[5] http://www.irda.org.

degrees in case of IrDA. Eavesdropping directional communication generally requires the third party to be on the same axis as the communication partners, and thus eavesdropping is more difficult than with non-directional broadcasts.

3.5 Wireless Personal Area Networking

Wireless Personal Area Networks (WPAN) include RF-based local communication technologies that are suitable for connecting various devices with each other and with the personal mobile terminal of the user. WPAN technologies can act as implementation technologies for physical browsing and tags, but they can also benefit from physical selection. Examples of standardised WPAN technologies include Bluetooth [Haartsen *et al.*, 1998] and ZigBee. In addition to them, SoapBox [Tuulari and Ylisaukko-Oja, 2002], a proprietary WPAN device, which we have used for emulating interaction with passive long-range sensor-equipped RFID tags, is described briefly.

Bluetooth was originally intended for low-cost wireless cable-replacement between computers and small portable devices like PDAs and mobile phones. The vision of Bluetooth is to enable communications in an ad-hoc fashion, based on the proximity of the devices [Haartsen *et al.*, 1998]. Compared to infrared communication, Bluetooth allows interaction in which the user does not have to know where to aim the communication beam [Ailisto et al., 2003a]. The operating range of Bluetooth depends on the device class and can be up to one metre, ten metres (typically in mobile devices) or one hundred metres (typically in industrial devices). ZigBee[6] is another RF-based WPAN technology. Like Bluetooth, it aims for a range of 1–100 metres but at lower power consumption and slower data transmission rate.

By default, WPAN RF technologies enable physical selection by scanning. Additionally, physical selection by touching or pointing may be used to select the correct WPAN device if the tag or reader can sense the position and alignment of the reader relative to the tag. A commercial example of implementing touch-based interaction to ZigBee has been released by

---

[6] See www.zigbee.org.

Cambridge Consultants [2006]. They take advantage of the received signal strength indication to measure the distance between two ZigBee nodes.
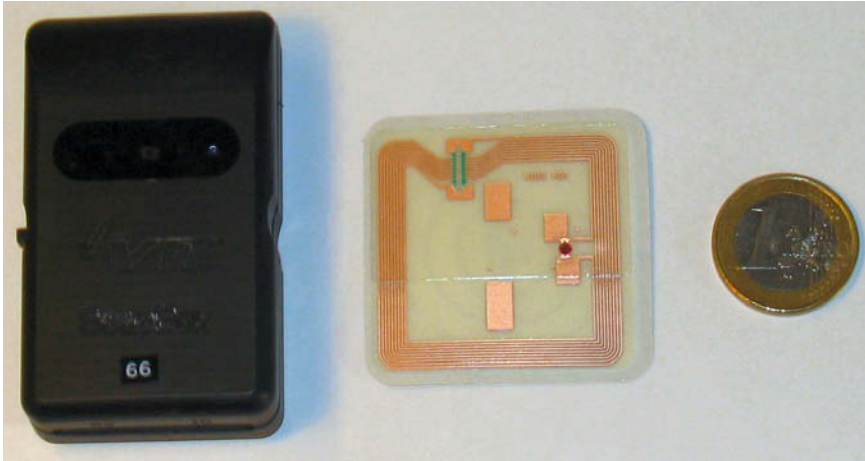
Figure 5. From left to right: A SoapBox, an RFID tag and a one euro coin.

SoapBox [Tuulari and Ylisaukko-Oja, 2002] is a low power proprietary WPAN device (see Figure 5). It is a configurable platform and offers low data rate services, intended mainly for ubiquitous computing research. The SoapBox network architecture includes one central SoapBox and zero or more remote SoapBoxes that send their measurement data to the central one. A SoapBox can contain a set of sensors, such as acceleration, illumination, temperature or moisture. We have used SoapBoxes to implement pointable (and touchable and scannable) tag emulators using photosensors in the SoapBoxes and illuminating them with a laser or infrared beam while pointing [Tuomisto *et al.*, 2005].

3.6 Comparison of Tagging Technologies

The different technologies presented above have fundamental differences and they enable different physical selection techniques. In this subsection, the main differences, considering their use as tagging technologies for physical selection, are discussed. Want *et al.* [1999] state that a tagging system must consider not only the affordances of the tag, but also the affordances of the tag reader. All the discussed technologies are at least to some degree usable in mobile terminals

now that RFID-enabled mobile phones have been released and commercial off-the-shelf readers are available for PDAs as well. The readers for the discussed technologies are also relatively inexpensive and for example cameras in mobile phones come essentially free for physical browsing purposes because they are typically bought for taking photographs.

In Table 1 a selection of tagging technologies is summarised according to their characteristics (adapted from Ailisto *et al.*, 2003b and Paper III). The summary contains technical features of the tags (for example possibility for sensors or internal memory) and features affecting the interaction (for example alignment precision of the reader and supported selection methods).

The values in the table cells are approximations, but they should give some idea about the capabilities and differences of the available technologies. The capabilities of the technologies can thus be compared with each other instead of providing absolute values.

Unobtrusiveness of the tag is generally an advantage, but unless the tag is visualised well, it may turn into a disadvantage. Visual codes are always visible so the user knows at least that some information or services are available, but RFID tags can be hidden and then may require some visualisation methods to announce their presence and other features.

Associating functionality to the tags is one important challenge [Want *et al.*, 1999]. If the tag has enough onboard memory to contain for example URLs, the ID does not have to be necessarily associated to any action, but of course the content must still be written into the tag.

Table 1. Comparison of tagging technologies.

	Visual codes	Near-Field RFID	Far-Field RFID	IR	WPAN
Alignment precision required	Yes	No	No	Yes	No
Unobtrusive	No	Yes	Yes	Yes	Yes
Cost	Very inexpensive	Inexpensive	Inexpensive	More expensive	More expensive
Printable	Yes	Possibly soon	Possibly soon	No	No
Selection methods	Touching or pointing	Touching	Possibly all	Pointing	Scanning
Internal memory	No	Yes	Yes	Yes	Yes
Line-of-sight required	Yes	No	No	Yes	No
Batteries required	No	No	No	Yes	Yes
Robustness and maintainability	High	High	High	Low	Low
Rewritable data	No	Yes	Yes	Yes	Yes
Detect position and alignment of tag	Yes	Position	No	Usually no	No
Sensor interface	No	Possible	Possible	Possible	Possible
Read multiple simultaneously	No	Yes	Yes	Possibly	Yes

3.7 Selection and Communication

As seen from the previous subsection, different communication technologies have different advantages and disadvantages. An interesting note regarding the choice of technology is that selection and further communication do not have to use the same communication technologies. Selection can be done with a directional technology such as IR or a touch range technology such as NFC but the further communications can be by other wireless means such as Bluetooth, WLAN or cellular networks. In that case, the selection transfers the communication parameters to the mobile terminal, which then connects to the target device using these parameters.

From the user's point of view, it is easy to create a temporary connection by touching or pointing at a target, but keeping the mobile terminal next to the target or aligned towards it gets harder if the connection is to remain long [Swindells *et al.*, 2002]. Therefore the concept of using one technology for selection and another for further connections is feasible from the user's point of view. This is also another reason for separating selection and action from each other.

4. From Graphical to Tangible User Interfaces

To better illustrate the evolution of physical selection, the road of digital information from the purely digital form in desktop computers to mobile terminal readable physical containers and tokens is traced through a few well-known and interesting example systems. The beginnings of the desktop metaphor are first described briefly, and then how computing was brought from the virtual desktop to the real desktop, and to the physical world through concepts such as tangible user interfaces and computer-augmented environments. Eventually, these ideas led to bridging the physical and virtual worlds by identifying physical objects, and with mobile computing, also to the concept of physical browsing.

4.1 Dynabook

Kay and Goldberg [1977] presented the idea of Dynabook – a general-purpose notebook-sized computer, which could be used by anyone from educators and business people to poets and children. In the vision of Kay and Goldberg, each person has their own Dynabook. They envisioned a device as small and portable as possible, which could both take and give out information in quantities approaching that of human sensory systems. One of the metaphors they used when designing such a system was that of a musical instrument, such as a flute, which the user of the flute owns. The flute responds instantly and consistently to the wishes of the owner, and each user has an own flute instead of time-sharing a common flute.

As a step towards their vision, Kay and Goldberg designed an interim desktop version of the Dynabook. The Dynabook vision led eventually to building the Xerox Star [Smith *et al.*, 1982] with graphical user interface and to the desktop metaphor of the Star. On the other hand, the original vision of Dynabook was an early idea of a mobile terminal, making the Dynabook vision even more significant in the evolution of physical browsing.

4.2 The Star User Interface and the Desktop Metaphor

The desktop metaphor refers to a user interface metaphor in which the UI is designed to resemble the physical desktop with familiar office objects. The beginning of the desktop metaphor was Xerox Star [Smith *et al.*, 1982], a personal computer designed for offices. The designers of Xerox Star hoped that with the similarity of the graphical desktop user interface and the physical office, the users would find the user interface familiar and intuitive.

Wellner [1993] stated that the electronic world of the workstation and the physical world of the desk are separate, but each "has advantages and constraints that lead us to choose one or the other for particular task" and that the desktop metaphor is an approach to solving the problem of choosing between these two alternatives.

Like the interim Dynabook, the Star user interface hardware architecture included a bit mapped display screen and a mouse [English *et al.*, 1967]. Smith *et al.* [1982] claim that pointing with the mouse is as quick and easy as pointing with the tip of the finger. However, the mouse is not a direct manipulation device [Wellner, 1993]; the movements of the mouse on a horizontal plane are transformed into cursor movement on a typically vertical plane some distance away from the mouse itself.

4.3 From Desktop Metaphor to the metaDESK

The physical desk led to the desktop metaphor, taking advantage of the strengths of the digital world, but at the same time ignoring the skills the users had already developed for interacting with the real world, and separating the information in digital form from the physical counterparts of the same information. It is also easy to see that the desktop metaphor suits best office tasks and not for example mobile computing. These shortcomings have led to two directions in bridging the physical and virtual worlds:

1. Integrating the physical objects back into the user interface, as is done in the tangible user interfaces approach;

2. Augmenting physical objects with digital information and accessing that information with a computational device, such as a mobile terminal.

In the next subsections, the first of the aforementioned approaches – integrating physical objects with the user interface – is explored through a few example systems that illustrate well the ideas of bridging the gap between the physical and virtual worlds. The DigitalDesk [Wellner, 1993] is an early system for merging the physical and digital worlds in an office desktop setting. It is also one of the pioneering augmented reality systems. Although Wellner's focus was on office systems, his ideas of combining the physical and digital worlds, while taking advantages from each, are still valid in the ubiquitous computing environments. Bricks [Fitzmaurice *et al.*, 1995] and Active Desk is another physical desktop-based system in which physical and digital objects are connected.

4.3.1 DigitalDesk

Wellner [1993] states that we interact with documents in two separate worlds: the electronic world of workstation (using the desktop metaphor), and the physical world of the desk. Both worlds have advantages and constraints that lead us to choose one or the other for particular tasks in the office setting. Although activities on the two desks are often related, the two are unfortunately isolated from each other. Wellner suggests that "*instead of putting the user in the virtual world of the computer, we could do the opposite: add computer to the real world of the user.*" Following this thought, instead of making the electronic workstation more like the physical desk, the DigitalDesk makes the desk more like a workstation and supports computer-based interaction with paper documents. According to Wellner, the difference between integrating the world into computers and integrating computers into the world lies in our perspective: "*do we think of ourselves working primarily in the computer but with access to physical world functionality, or do we think of ourselves as working primarily in the physical world but with access to computer functionality?*"

The DigitalDesk system [Wellner, 1993] is based on a real physical desk, which is enhanced to provide the user some computational capabilities. The desk includes a camera that projects the computer display onto the desk, and video cameras that track the actions of the user and can read physical documents on the desk. The interaction style in DigitalDesk is more tactile than the non-direct manipulation with a mouse. The user does not need any special input devices in addition to the camera: pointing and selecting with just fingers or a pen tip is

sufficient. The projected images can be purely digital documents, user interface components, and images and text that are superimposed onto paper documents with the existing contents of the paper.

4.3.2 Bricks

Fitzmaurice *et al.* [1995] took another step towards integrating the physical and virtual worlds with Graspable User Interfaces, a new user interaction paradigm. They described their concept: "the Graspable UIs allow direct control of electronic or virtual objects through physical artefacts which act as handles for control." Graspable UIs are a blend of virtual and physical artefacts, each offering affordances in their respective instantiation. A graspable object in the context of graspable UIs is an object that is composed of both a physical handle, and a corresponding virtual object.

The prototype system in Bricks [Fitzmaurice *et al.*, 1995] consists of Active Desk, the Bricks (roughly one inch sized cubes) and a simple drawing application. As in DigitalDesk [Wellner, 1993], the Active Desk uses a video projector to display the graphical user interface parts on the surface of the desk. Instead of tracking the fingers of the user or a pen tip with a camera, the Bricks system uses six-degrees-of-freedom sensors and wireless transmitters inside the Bricks. The location and orientation data is transmitted to the workstation that controls the application and the display. Although it offers a new user interaction paradigm, Bricks still builds upon the conventions of the graphical user interface in a desktop setting.

4.3.3 The metaDESK and Tangible Geospace

The metaDESK [Ullmer and Ishii, 1997] is an example of a tangible user interface (TUI, further explored in the next subsection), which brings familiar metaphorical objects from the computer desktop back to the physical world. The metaDESK is an augmented physical desktop on which phicons – physical icons – can be manipulated. The use of physical objects as interfaces to digital information forms the basis for TUIs. The metaDESK was built to instantiate physically the graphical user interface metaphor. The desktop metaphor drew itself from the

physical desktop and in a manner of speaking, metaDESK again realised physically the GUI components. In metaDESK, the interface elements from the GUI paradigm are instantiated physically: windows, icons, handles, menus and controls each are given a physical form. For example, the menus and handles from GUI are instantiated as "trays" and "physical handles" in TUI.

Tangible Geospace [Ullmer and Ishii, 1997] is a prototype application built on the metaDESK platform. In Tangible Geospace, a set of objects was designed to physically resemble different buildings appearing on a digital map of the MIT campus. By placing a model of a building on the display surface, the user could bring up the relevant portion of the map. Manipulating the building on the metaDESK surface controlled the position and rotation of the map. The physical form of the objects would serve as a cognitive aid for the user in finding the right part of the map to display. For example a model of a familiar landmark such as the Great Dome of the MIT is easy to recognise on the metaDESK surface. The model thus acts as a container for the digital information and as a handle for manipulating the map. Adding a second phicon on the map rotates and scales the map so that both phicons are over correct locations on the digital map. Now the user has two physical handles for rotating and scaling the map by moving one or both objects with respect to each other, which is very similar to interaction with Bricks.

The Tangible Geospace application already resembles the basic idea of physical browsing. Although the "terminal" is a desk surface and instead of physically manipulating the terminal, the user manipulates tagged objects. The tagged objects act as links to digital information and bringing the "terminal" and the tagged objects together, the digital information can be displayed to the user.

4.4 Tangible Bits Vision

According to Ishii and Ullmer [1997], the GUI approach falls short in many respects, particularly in embracing the rich interface modalities between people and the physical environments. Their approach to this problem is a tangible user interface (TUI), part of which was demonstrated earlier with graspable user interfaces and the metaDESK platform. TUIs are user interfaces employing real physical objects, instruments, surfaces, and spaces as physical interfaces to

digital information, and the user can physically interact with digital information through the manipulation of physical objects. The use of tangible objects – "real physical entities which can be touched and grasped" – as interfaces to digital information forms the basis for TUIs. This use of physical objects as containers for digital information makes tangible user interfaces important to physical browsing.

As a part of their work with tangible user interfaces, Ishii and Ullmer introduced the Tangible Bits vision [Ullmer and Ishii, 1997]. The Tangible Bits vision includes three platforms: metaDESK, transBOARD and ambientROOM. Together these three platforms explore graspable physical objects and ambient environmental displays. In the Tangible Bits vision, people, digital information and the physical environment are seamlessly coupled – an idea very similar to the physical browsing systems of Want *et al.* [1999] and to the Cooltown [Kindberg *et al.*, 2002]. An important topic in their work is exploring the use of physical affordances within TUI design. The ambientROOM explores ambient, peripheral media and is outside the scope of this dissertation.

The transBOARD [Ishii and Ullmer, 1997] is an interactive surface in spirit of both the vision of Tangible Bits and Weiser's vision of boards as one class of ubiquitous computing devices. This kind of interactive surface absorbs information from the physical world and transforms it into digital information what can be distributed to other computers in the network. The transBOARD uses hyperCARDS, which are paper cards with barcodes to identify and store the strokes on the physical board as digital strokes. The cards can be attached onto the transBOARD and the when the strokes are recorded and stored, the barcode of the card is associated with the location of the stored data. The board contents can this way be saved, taken to other computers and replayed when the card is introduced to a suitably equipped computer again. Whereas the metaDESK is an interactive surface, which can also alter its contents, the transBOARD is a simple recording device. The interesting idea here from the point of view of physical browsing is "saving the contents of the board into a card", making the card a container for or a link to the information.

4.5 Closely Coupling Physical Objects with Digital Information

In their further work, Ullmer and Ishii [2001] re-defined tangible user interfaces to have no distinction between input and output. According to their definition, physical objects act both as physical representations and controls for digital information. With the new definition, tangible interfaces give physical form to digital information instead of just associating physical objects and digital information, employing physical artefacts both as representation and controls for computational media. The important distinction between tangible user interfaces and traditional input devices that have physical form – such as keyboards and mice – lies in that the traditional input devices hold little representational significance. In graphical user interfaces the information representation is separated to displays.

In physical selection, the control and representation (input and output) are not integrated as tightly as in this model for tangible user interfaces. This definition of tangible user interfaces separates TUIs somewhat from physical browsing. In physical browsing, the links are rarely controls, and the physical objects with tags only represent the information, but do not dynamically display it. The physical object acts only as an input token to the mobile terminal.

4.5.1 mediaBlocks

MediaBlocks [Ullmer *et al.*, 1998] are small, electronically tagged wooden blocks that serve as phicons (physical icons) for the containment, transport and manipulation of offline media. They allow digital media to be rapidly stored into them from a media source such as a camera or a whiteboard and accessed later with a media display such as a printer or a projector. MediaBlocks thus allow "physical copy and paste" functionality. MediaBlocks do not store the media internally but instead they are augmented with tags that identify them, and the online information is accessed by referencing to it with a URL. MediaBlocks function as containers for online content and they can be understood as a physically embodied online media. Ullmer *et al.* see mediaBlocks as filling the user interface gap between physical devices, digital media and online content. They intended mediaBlocks as an interface for the exchange and manipulation of online content between diverse media devices and people.

Several tangible user interfaces described earlier have influenced the design of the mediaBlocks [Ullmer *et al.*, 1998]. Bricks [Fitzmaurice *et al.*, 1995] were among the first phicons, although Bricks were not containers for digital content but instead were used to manipulate digital objects inside a single area, the Active Desk. In metaDESK [Ullmer and Ishii, 1997], phicons were used, not only as short cuts to digital information, but also as physical controls – for example rotating a phicon on the metaDESK rotated the displayed map. The functionality of mediaBlocks as storage devices for whiteboards draws from the transBOARD [Ishii and Ullmer, 1997], but instead of barcodes, electronic tags are used to link to the contents of the board. Ullmer and Ishii [1997] see RFID tags as a promising technology for realising the physical/digital bindings.

The contents of mediaBlocks remain online and that makes mediaBlock seem to have unlimited data storage capacity and rapid transfer speed when the block itself is moved around or the contents are copied just by copying the link to the online content [Ullmer *et al.*, 1998]. MediaBlocks can also contain streaming media. One role of the mediaBlocks is to support simple physical transport of media between different devices. Copying and pasting information is a commonly used function in graphical user interfaces and mediaBlocks are intended to provide the same functionality to physical media. Ullmer et al. have built slots for mediaBlocks in different devices such as whiteboards and printers but also on desktop computers.

In addition to adding mediaBlock interfaces to various existing devices, Ullmer *et al.* [1998] have built special devices for mediaBlocks. The media browser is used to navigate sequences of media elements stored in mediaBlocks. The media sequencer allows sequencing media by arranging mediaBlocks on its racks. This extended functionality is beyond the scope of this dissertation.

4.5.2 Other Removable Media Devices

In addition to mediaBlocks, other removable media devices exist, from floppy disks to more current DVDs and USB sticks, which were not ubiquitous technologies at the time of the development of mediaBlocks. Ullmer *et al.* [1998] claim that an important difference between these technologies and mediaBlocks is that mediaBlocks store only a link to the online media instead of recording the actual content onto the storage device.

However, nothing prevents us from storing links to online media on the other removable storage media, even onto floppy disks if we so desire. This way any storage device can support almost infinite space and varying bandwidths, just as Ullmer *et al.* [1998] describe MediaBlocks. They claim that other media transport devices are accessed indirectly through graphical or textual interaction. But what prevents us from "auto playing" for example video files from a USB stick when it is inserted into a projector? Ullmer *et al.* also mention the lack of disk drives on the different media sources and targets, but neither are there mediaBlock slots on commercial devices. Granted, it is not feasible to have for example DVD drives on many devices simply because of the physical dimensions and power requirements. However, many current media devices have USB ports and can record content to USB disks and read from them. Additionally, the devices Ullmer *et al.* augmented with mediaBlock slots, had only one such slot and only their custom-built browsers and sequencers took advantage of the possibility to contain many blocks at the same time. So, looking briefly, it seems that the mediaBlock concept would not be valid any more.

Still, mediaBlocks have a property that is extremely useful for physical interaction. They contained electronic tags that are small and cheap compared to current storage devices, allowing the use of one block for one link, thus making it possible to physically sort the blocks and extend the manipulation and sorting of digital content into the physical world just as Ullmer *et al.* [1998] intended. This is a powerful interaction paradigm and the mediaBlocks demonstrate it well.

4.6 Token-Based Access to Digital Information

Token-based access to digital information means accessing virtual data through a physical object. The paper of Holmquist *et al.* [1999] is among the first systematic analyses of systems that link physical objects with digital information. They defined token-based access to digital information as follows:

"*A system where a physical object (token) is used to access some digital information that is stored outside the object, and where the physical representation in some way reflects the nature of the digital information it is associated with.*"

Holmquist *et al.* [1999] enumerate the two components in a token-based interaction system: tokens and information faucets. Tokens are physical objects, which are used as representation of some digital information. In physical selection, the tokens correspond to the tagged physical objects and provide links to digital information related to the objects. Information faucets or displays are access points for the digital information associated with tokens. In physical selection, the faucet corresponds to the mobile terminal, but in theory, it can be any device capable of reading the tag and presenting the information it links to.

The physical objects (tokens in previous paragraph) are further classified into containers, tokens and tools [Holmquist *et al.,* 1999]. Tools are physical objects that are used to actively manipulate digital information. They usually represent some computational function. For example, in the Bricks system [Fitzmaurice *et al.*, 1995], the physical bricks could be used as tools by attaching them onto virtual handles on a drawing application. The lenses in the metaDESK system [Ishii and Ullmer, 1997; Ullmer and Ishii, 1997] also correspond to tools. Tools do not have a direct counterpart in physical selection.

Containers are generic objects that can be associated with any type of digital information [Holmquist *et al.*, 1999]. They can be used to move information between different devices or platforms. The physical properties of a container do not reflect the nature of the digital information associated with it. For example, mediaBlocks [Ullmer *et al.*, 1998] are containers, because by merely examining the physical form of a mediaBlock, it cannot be known what kind of media it contains.

Tokens are objects that physically resemble in some way the digital information they are associated with [Holmquist *et al.*, 1999]. That way the token is more closely tied to the information in represents than a container is. The models of buildings in Tangible GeoSpace [Ishii and Ullmer, 1997; Ullmer and Ishii, 1997] are an example of tokens.

In physical selection, it does not matter (technologically) whether the object is a container or token. In an ideal case the information and the object are connected, but nothing prevents a user from sticking completely unconnected tags and objects together.

The two most important interactions in a token-based system are access and association [Holmquist *et al.*, 1999]. The user has to be able to access the information contained in the token by presenting the token to an information faucet. Association means creating a link to the digital information and storing that link in the tag of the token so that it can be accessed later. Holmquist *et al.* note that it may be useful to allow associating more than one piece of information to a single token and they call this method overloading. When the token is brought to a faucet, the information presented to the user may then vary according to the context, or the user may get a list of the pieces of information stored in the token. This may present problems to the physical hyperlink visualisation, as is shown later.

Holmquist *et al.* [1999] note that it is important to design the tokens in a way that clearly displays what they represent and what can be done with them. This refers to taking into account the existing affordances of the existing physical object in question when linking it to some digital information, but it can also be applied if a specific "link container" is designed for a link.

Later, Ullmer and Ishii [2001] chose to describe the physical elements of tangible user interfaces in terms of tokens and reference frames. They consider a token a physically manipulatable element of a tangible interface, such as a metaDESK phicon [Ullmer and Ishii, 1997; Ishii and Ullmer, 1997]. A reference frame is a physical interaction space in which these objects are used, such as the metaDESK surface. Ullmer and Ishii [2001] accept the term container for a symbolic token that contains some media (again, as in mediaBlocks [Ullmer *et al.*, 1998]) and the term tool for a token that is used to represent digital operations or function. Considering physical selection, we are mostly interested in the terms container and token, which take approximately the same meaning as defined by Holmquist *et al.* [1999].

4.7 A Taxonomy for Tangible Interfaces

4.7.1 Definitions for Tangible User Interfaces

The term "tangible user interface" surfaced in *Tangible Bits*, in which Ishii and Ullmer [1997] defined it as a user interface that "augments the real physical world by coupling digital information to everyday physical objects and environments".

In *Emerging Frameworks for Tangible User Interfaces*, Ullmer and Ishii [2001] re-defined tangible interfaces as a user interface that eliminates the distinction between input and output devices. However, they were willing to relax the definition to highlight some interaction methods.

Fishkin [2004] describes the basic paradigm of tangible user interfaces as follows: "*a user uses their hands to manipulate some physical object(s) via physical gestures; a computer system detects this, alters its state, and gives feedback accordingly*". According with his definition, Fishkin created a script that characterises TUIs:

1. Some input event occurs, typically a manipulation on some physical object by the user, and most often it is moving the object.

2. A computer senses the event and alters its state.

3. An output event occurs via a change in the physical nature of the object.

Fishkin [2004] describes how the script applies to metaDESK [Ullmer and Ishii, 1997]. The user moves a physical model of a building on the surface of the metaDESK. The system senses the movement of the model and alters its internal state of the map. As output, it projects the new state of the map onto the display surface. Another of the examples Fishkin gives, is the photo cube by Want *et al.* [1999]. Bringing the cube a specific face down onto the RFID reader is the input event. The computer reads the tag on the cube in the second phase of the script and in the third phase, displays the associated WWW page as an output event. The output event in Fishkin's script does not thus happen in the physical object that contains the tag, but it can occur in another object, the display terminal in the photo cube case.

Similarly, we can see that physical selection is a user interaction method in a tangible user interface. In the first phase of the script, the user manipulates the mobile terminal for example by bringing it close to a tag. The tag reader in the terminal reads ("senses") the tag and alters its state according to what it is programmed to do when the tag in question is read. In the output phase, an action linked to the tag is activated and the action is visible on the screen of the phone (for example a WWW page) or can be sensed in the environment (for example an electronic lock has opened).

As Fishkin [2004] himself notes, any input device can fit into this script. Even a keyboard in a desktop computer is a physical object. Manipulating it causes an input event to occur and the computer senses the event, altering its state and produces an output event on the computer screen. Therefore Fishkin does not characterise an interface as "tangible" or "not tangible", but introduces varying degrees of "tangibility".

4.7.2 The Taxonomy

Fishkin proposes a two-dimensional taxonomy for tangible interfaces. The axes of the taxonomy are embodiment and metaphor. Embodiment describes to what extent the user thinks of the state of computation as being inside the physical object, that is, how closely the input and output are tied together. Fishkin presents four levels of embodiment:

1. Full: the output device is the input device and the state of the device is fully embodied in the device. For example, in clay sculpting, any manipulation of the clay is immediately present in the clay itself.

2. Nearby: the output takes place near the input object. Fishkin mentions the Bricks [Fitzmaurice *et al.*, 1995], metaDESK [Ullmer and Ishii, 1997] and photo cube [Want *et al.*, 1999] as examples of this level.

3. Environmental: the output is around the user. For example, ambient media [Ishii and Ullmer, 1997] corresponds to environmental embodiment.

4. Distant: the output is away from the user, for example on another screen or in another room. Fishkin mentions a TV remote control as an example of this level.

Physical selection typically has embodiment levels from Full to Environmental. Often the output occurs in the mobile terminal itself (Full), but if the output device is rather seen to be the object the user selects with the terminal, the embodiment level is then Nearby. Physical selection can also cause actions around the user, in the environment. As the photo cube [Want *et al.*, 1999] is very closely related to physical selection, we should probably take Fishkin's classification of the photo cube to correspond to the classification of physical selection, making therefore it Nearby.

Fishkin defines the second axis, metaphor, as "is the system effect of a user action analogous to the real-world effect of similar actions?" Fishkin divides his metaphor axis to two components: the metaphors of noun and verb. Thus, there are five levels of metaphor:

1. No Metaphor: the shape and manipulation of the object in TUI does not resemble an object in the real world. Fishkin mentions the command line interface as an example of this level.

2. Noun: the shape of input object in TUI is similar to an object in a real world. The tagged objects Want *et al.* [1999] developed correspond to this level of metaphor. For example, their augmented bookmarks resemble real bookmarks.

3. Verb: the input object in TUI acts like the object in the real world. The shapes of the objects are irrelevant.

4. Noun and Verb: the object looks and acts like the real world object, but they are still different objects. In traditional HCI, an example of this level is the drag and drop operation in the desktop metaphor.

5. Full: the virtual system is the physical system.

Physical selection can be seen to correspond roughly to the Noun metaphor. Again, we can safely assume Fishkin's classification of Want *et al.*'s examples as guidelines.

Advancing on the metaphor scale means less cognitive overhead as the object itself contains in its shape and function information about how it can be used in a tangible interface. However, decreasing the level of metaphor makes the object more generic and re-usable. Therefore, the level of metaphor should, if possible, be designed consciously to suit the task [Fishkin, 2004]. For example, among the strengths of Bricks [Fitzmaurice *et al.*, 1995] and transient WebStickers [Ljungstrand and Holmquist, 1999; Ljungstrand *et al.*, 2000] are the possibilities to contain any information, and to act as operators to any virtual functions.

4.7.3 Comparison to Containers, Tokens and Tools

Fishkin [2004] compares the containers, tokens and tools of Holmquist *et al.* [1999] taxonomy to his own. Containers are fully embodied in the Fishkin taxonomy and use the verb metaphor. The information is considered to be inside the container and moving the container moves the information. As long as a container does not employ the noun metaphor (the shape does not resemble the data), the container retains its generic and flexible nature and can contain any information.

Tokens are objects that physically resemble the data they contain and thus correspond to the noun metaphor [Fishkin, 2004; Holmquist *et al.*, 1999]. Like containers, they can also be used to move information around and therefore also correspond to the verb metaphor, making them span the metaphor scale from Noun and Verb to Full.

As physical selection is mostly about containers and tokens, it can be seen as having the embodiment of any level, but particularly from Full to Environmental, as noted earlier. The metaphor level of containers and tokens and thus tagged objects is something between only Noun or Verb, and Full. Fishkin's own analysis of how the taxonomy of Holmquist *et al.* maps to his taxonomy seems to be slightly ambiguous. Perhaps we can say that even the steps in Fishkin scales are not binary but different tangible interfaces can be seen as having different degrees of "Noun-ness" or "Nearbyness".

5. Physical Browsing Systems

In this Chapter some example systems that implement physical selection are covered. Especially after the release of commercial RFID readers for mobile phones and PDAs, the amount of systems and applications has become so vast that all related work is impossible to cover in reasonable space. Therefore some examples that have either been pioneering in the development of the concept, or have contributed significantly to the user interaction development, or to this dissertation, have been chosen. A discussion of relevant user studies can be found in "Evaluating Touching and Pointing for Physical Browsing" (Paper V).

5.1 Architectures

5.1.1 Cooltown

The goal of Cooltown project [Kindberg *et al.*, 2002] was to create a bridge between the World Wide Web and the physical world we inhabit – people, places and things. The services in Cooltown were common web-based services that are available anywhere there is connectivity, and also services integrated into the everyday physical world. To access these services, the users had to carry a wirelessly connected and sensor-equipped mobile terminal. IR beacons, RFID tags and location-awareness with Global Positioning System (GPS) were used to link the web services to people, places and things.

Some Cooltown applications included a museum and a meeting room. In the museum, the rooms contained IR beacons that broadcasted URLs to the mobile terminal of the user. Paintings in the museum contained RFID tags that could be read with the mobile terminal. Cooltown used "direct sensing" with the IR tags, meaning that the tags transmitted the URL itself to the reader. With RFID, the transmitted information was the tag ID, which was separately mapped to a URL, and Kindberg *et al.* [2002] used the term "indirect sensing" for this technique.

The meeting room demonstrated how different devices could communicate with each other. For this purpose Kindberg *et al.* [2002] developed the eSquirt protocol based on IR communications. The user could read URLs with the

mobile terminal in the same manner as in the museum example, but with eSquirt, the URLs could be transmitted to other devices. A user could for example pick up the URL of a document and eSquirt it to a projector or a printer to display or print it.

5.1.2 Tag Manager

Tag Manager [Keränen *et al.*, 2005; Pohjanheimo *et al.*, 2005] is an example of a software architecture built for mobile phones running the Symbian Series 60 operating system. Whereas Cooltown is a system infrastructure with emphasis on the communication between the different devices and building blocks, Tag Manager focuses on creating a middleware component between various tag readers and applications.

The Tag Manager supports among others RFID tags and bar codes. Pohjanheimo *et al.* [2005] and Keränen *et al.* [2005] demonstrated several applications in which they connected tag readers to a mobile phone and used custom-made links or utilised existing coding systems found in products, such as ISBN codes in books. The difference between this dissertation work and the work of Keränen *et al.* is that they see tag as an identifier and the resulting action dependent on the context of use, whereas in this work, the focus has been on providing a constant hyperlink to digital information from the tag. However, in both works it is agreed that the mobile terminal should take into account the active application so that touching a tag can be used as input into a dialogue between the system and the user [Kaasinen *et al.*, 2005; Kaasinen *et al.*, 2006].

5.2 RFID and IR Systems

5.2.1 Bridging Physical and Virtual Worlds with Electronic Tags

Want *et al.* [1999] used RFID tags to connect physical objects with virtual representations or computational functionality. This way the physical object was linked to digital information related to it and the object acted as a pointer to the digital information.

A tag reader was attached to a mobile computer, which also had wireless networking capabilities. Want *et al.* demonstrated how to link invisibly, seamlessly and portably physical objects to networked electronic services and actions associated to the objects and their form. They see bridging physical and virtual worlds a key path for the design of future user interfaces.

The goal of Want *et al.* was to "support everyday tools and objects, the affordances of these objects, and to computationally augment them to support casual interaction using natural manipulations and associations." Want *et al.* took everyday objects that already have some purpose independent of any electronic system and they then augmented those objects with unobtrusive RFID tags. The simple and inexpensive infrastructure allows a large number of everyday objects be easily tagged and used in multiple locales.

When the reader reads a tag, the application interprets the tag ID, determines the current application context and provides an appropriate action. Their selection method corresponds to "Direct pick of tagged objects" in the taxonomy of Ballagas *et al.* [2006].

Want *et al.* augmented physical documents and books with RFID tags, thereby introducing a link between the physical document and its corresponding electronic document. A business card included a link to the home page of the person whose card it was. Another function in business cards was to generate email messages with the addressing information already filled in.

They also created a bookmark application similar to the WebStickers, but added two tags onto a physical bookmark. Touching the other end of the bookmark associated the currently open page of a document to the physical bookmark. Touching the other end opened the previously associated bookmark in a document reader.

They implemented linking to services by augmenting a book with a link to the corresponding web page in a web-based bookstore so that a copy of the book could be purchased by reading the tag. They also created a "translation tag" by attaching a tag onto a dictionary and associating the tag with an application on the mobile computer to translate the currently open electronic document. A wristwatch was augmented with a tag so that when the watch was brought near the reader, the mobile computer opened the calendar for the identified person.

Tags could also set context. In some applications, Want *et al.* have used the IrDA ports on the mobile computers to receive room identification from an IR beacon in the room in a similar manner to their earlier *Active badge location system* [Want *et al.*, 1992]. This allowed them to refine more accurately the context of tagged objects. They thus demonstrated two location and tag related functions: 1) the location information from some other source provides further context for the tag and possibly modifies the action, and 2) the tag itself sets the location information, possibly for some other application that depends on the location information.

The Photo Cube was an example of creating a physical container for several virtual links. The cube had a picture of one person on each of its sides and under each picture was an RFID tag. The tag acted as a link to the homepage of the person and the picture acted as a visualisation of the link.

5.2.2 A 2-Way Laser-Assisted Selection Scheme

Patel and Abowd [2003] explored direct interactions with devices in the physical world mediated through a mobile terminal and implemented a physical selection system with pointing. They found out that laser provides a simple means of visual feedback as well as at-a-distance interaction.

The prototype implementation of Patel and Abowd [2003] consisted of a hand-held device with a laser pointer, and a photosensitive tag with light-emitting diodes to show its location. The hand-held device uses a modulated laser signal to communicate its identity to the tag. The tag can then communicate back to the hand-held, establishing a two-way link. Important human factors they found were: 1) acquisition time to locate a tag with the laser pointer (1 second, consistent with Myers *et al.* [2002]), 2) how long the user can comfortably keep the beam steady on the sensor to ensure a hit and 3) what feedback tells the user that the target is hit. The tag may be triggered accidentally because the user does not know where exactly the beam will hit when turned on, so it may hit a wrong target initially, or on the way to the correct target.

5.2.3 RFIG Lamps

In RFIG Lamps, Raskar *et al.* [2004] have extended tag-readers that operate in broadcast mode to allow selection of individual tags. Other aspects of their work have been to create a three-dimensional coordinate frame for the tags and to create an interactive projection. To achieve these functionalities, they have combined a tag reader with a projector and augmented each tag with a photo-sensor.

Raskar *et al.* [2004] used battery-powered tags but their ultimate aim was to create a passive RFID system supporting the same functionality and therefore they limited themselves to power levels that they suspected were available for tags without batteries. For this reason they could not use solutions by Patel and Abowd [2003], who used light-emitting diodes on the tags to display the location of the tag, and power-hungry communication protocols. The primary purpose of their projector is to use it for communicating with the RFID tags, but it also allows projecting augmented reality data over the tagged surfaces.

Raskar *et al.* [2004] state that accurate pointing required in laser-pointer systems is difficult when the target (tag) is visually imperceptible. In contrast, they use a "casually-directed" projector beam to select all tags within the beam for further interaction. This is very similar to our IR selection idea, used in our pointing prototype [Tuomisto *et al.*, 2005] (Paper IV).

According to Raskar *et al.* [2004], traditional tags can not use visible or infrared light communication because occlusion and ambient illumination make light unsuitable for data transfer. They see however light as an option for directional selection of tags. The hand-held projector of Raskar *et al.* first transmits a conventional RF broadcast, which wakes up each tag in range. The photo-sensor of each tag takes a reading of the ambient light. After that, the projector illumination is turned on and each tag that detects an increase in the light measurement, sends a response to indicate that it is in the beam of the projector.

To display a tag in the screen of the mobile terminal, the projector sends a Gray code in each pixel when it is illuminating the tags. The tags respond to the reading request by transmitting their identification and the received code from the projector. Since the terminal device has a camera that can also detect the projected code, the terminal knows where each unique code was projected.

When a tag responds with a code it received, it is straightforward for the terminal to add visual markers for the locations of the tags in its projection image. Even without this extra functionality, RFIG lamps illustrates well how the combination of sensors and passive RFID tags can be used for pointing.

5.2.4 RFID Tag Visualisations

Riekki *et al.* [2006] have built a system that allows requesting pervasive services by touching RFID tags. The RFID reader in their system is an external handheld device and connected via Bluetooth to a mobile phone. The work of Riekki *et al.* is among the first ones in which the tag visualisation has been studied; in earlier systems the visualisations have been either non-existent or designed in an *ad-hoc* manner.

Riekki *et al*. have two categories of tag visualisations and matching functionalities for the categories. A general tag is attached to an object and it identifies the object. Touching a general tag brings to the terminal a list of services related to the object and the user will have to select the desired service from the GUI of the mobile terminal. A special tag identifies an object like a general tag, but it also presents additional visual information related to the action to be performed. Touching a special tag accesses the special service directly instead of bringing to the terminal a list of all appropriate services related to the object.

5.2.5 Touching, Pointing and Scanning prototypes

Rukzio *et al.* [2006] studied touching, pointing and scanning and built one prototype system for each selection method. For touching, they used the Nokia 3220 NFC phone. The pointing system was based on a Nokia N70 mobile phone augmented with a laser pointer and the tags used light sensors to detect pointing and a radio to communicate the pointing event. For scanning, they used Bluetooth in Nokia N70 mobile phone and the users could scan for nearby Bluetooth devices.

Rukzio *et al.* conducted user experiments with all three selection methods and analysed the selection methods and in which situations they were chosen. They found touching and pointing to be natural methods and touching especially unambiguous with a very low cognitive load. Scanning had the lowest physical effort.

5.2.6 GesturePen

Swindells *et al.* [2002] have built gesturePen, which allows identifying devices through pointing. The gesturePen system includes a custom stylus and custom tags. The idea behind gesturePen is that when an environment contains an increasing amount of devices, it is difficult for human users to establish the identity of a certain device. Therefore Swindells *et al.* wanted to create a system that allows the user, who sees a desired device in the physical world, to connect with that device in the physical world, selecting "that one there". Another advantage of gesturePen is that it separates selection and further communication: the selection is done with a pointing gesture, but the user does not have to hold the gesturePen stationary during the data transfer, because the selection transmits wireless communication parameters that help establishing a Bluetooth or WLAN connection.

The gesturePen itself is a stylus with an IrDA transceiver and it can be connected to a portable device such as a PDA. It communicates with infra-red tags that are attached to external devices such as a printer. The tags communicate the connection parameters of the devices they are associated with to the PDA, allowing selection by pointing.

Swindells *et al*. compared their approach to the traditional scanning-style selection method involving selecting the target device from a menu. They found out that their gesturePen is significantly faster for connecting devices than menu-based methods because identifying the target in the real world was easier than identifying it by a name in a list.

5.3 Visual Codes

The interaction with visual codes is somewhat different from the interaction with electromagnetic or IR tags. The selection method is something between touching and pointing, depending heavily on the available camera hardware and tag decoding algorithm. However, similar connections between physical and digital worlds are possible as with RFID and other electromagnetic systems.

5.3.1 NaviCam

Rekimoto and Nagao [1995] propose a method of building computer augmented environments using NaviCam, a context-aware portable device. NaviCam recognises

the user's context by detecting colour-code identifiers in real world environments without the user's explicit input and displays information on a transparent display so that the user can look through the display. For example, in an augmented museum, the NaviCam can automatically detect an identifier next to a painting and generate a description of it on the screen of the NaviCam so that the description looks like it is next to the painting. Rekimoto and Nagao mention the advantage of this approach to be that the information can be personalised.

Later, Rekimoto and Ayatsuka [2000] criticise their earlier NaviCam system, for the use of coloured visual tags. They noticed that using colour in real-world applications is too unreliable because of changing lighting conditions and differing colour sensitivity of mobile terminal cameras. They therefore abandoned the coloured visual tags and used monochrome tags in CyberCode.

5.3.2 CyberCode

CyberCode is a tagging system based on two-dimensional visual tags [Rekimoto and Ayatsuka, 2000]. CyberCode tags can be read using low-cost cameras found in mobile devices. In addition to the tag ID, the system can recognise the alignment and 3D position of the tag. Although CyberCode is intended primarily for use with mobile devices, it can also be recognised with cameras embedded into the environment. In addition to printing CyberCode tags on paper, they can be displayed on computer and TV screens, further crossing the boundaries between the physical and digital worlds.

Rekimoto and Ayatsuka [2000] have also created a 'physically' extended concept of drag-and-drop by using CyberCode tags as operands for direct-manipulation operations. InfoPoint is a direct manipulation device for CyberCode tags. The user points at a tag with the InfoPoint wand, and then selects one of available actions within the wand display. Rekimoto and Ayatsuka describe InfoPoint as a universal commander for digital appliances. It can also be used to drag and drop digital items in the physical world. The user points at the source tag, moves towards the destination tag and drops the source tag by releasing the InfoPoint button.

5.3.3 SpotCode

SpotCode [Toye *et al.*, 2004] is a special circular visual code. Toye *et al.* have developed the SpotCode as a part of their infrastructure for connecting a mobile phone to Bluetooth services via visual codes. They have built the Mobile Service Explorer application that is used to detect SpotCodes and to emphasise the tag on the screen. If the user then selects the tag, the phone connects to the associated Bluetooth device.

Toye *et al.* have expanded their user interface from the mobile terminal to environmental displays. The information or service may be used on the mobile phone, but a nearby large display may as well be used for displaying graphical information. This way the mobile phone ceases to be an output device and it is only used to select tags.

In addition to selection, the SpotCode system supports other interactions with visual tags. The user can use the tags as controllers for the service displayed on the large screen, for example by rotating the mobile phone the tag acts as a kind of rotating knob for inputting numerical data. This demonstrates one of the differences between the capabilities of visual and electromagnetic tagging technologies.

5.4 Summary

In this Chapter some influential physical browsing and physical selection systems have been described. In Chapter 8 the evaluation results from these systems are compared to our results and discussed.

The main difference between this work and the systems described in this Chapter is that this work takes into account all three selection methods within one user interface. Previously, typically only one selection method using one technology has been used. In this dissertation I present a framework for three selection methods that are somewhat independent from the implementation technologies and application areas.

6. Selection as an Interaction Task

To better understand physical selection, it is necessary to look at selection in other, more established computing environments. Selection of an object for interaction is one of the basic interaction tasks found in many environments. In this Chapter, selection in desktop computer systems, immersive virtual environments and selection with a laser pointer with environmental displays, are introduced.

6.1 Selection in Desktop Computer Systems

Foley *et al.* [1984] proposed an organisation of interaction techniques based on the interaction tasks, for which the techniques are used. Their emphasis is in graphics applications, but their organisation is still at least somewhat applicable even for ubiquitous computing applications. The basis of the organisation was an analysis of the contemporary input devices and what input techniques they were able to perform. Although Card *et al.* [1990] criticise the organisation of Foley *et al.* as "the limitation of the scheme is that the categories are somewhat ad hoc and there is no attempt at defining a notion of completeness for the design space", the organisation of interaction techniques into interaction tasks is still a useful one.

In their analysis, Foley *et al.* [1984] found six fundamental interaction tasks:

- select (the user makes a selection from a set of alternatives),
- position (the user specifies a position, for example screen co-ordinates),
- orient (the user specifies an angle),
- path (the user specifies a set of positions and orientations),
- quantify (the user inputs a numeric value), and
- text entry (the user inputs a string value).

Of these tasks, selection is the most interesting one for this dissertation. Foley *et al.* [1984] further divided the selection task into different selection techniques. 'Direct pick' is a selection technique in which the selectable object is displayed and the user can pick it directly from the screen. In graphical user interfaces direct pick technologies are for example light pen and touch sensitive screen. In 'simulated pick with cursor match', a cursor is positioned over the desired visible object using some locator device like a mouse.

Direct pick can be seen to roughly correspond to physical selection by touching. Simulated pick roughly corresponds to physical selection by pointing. From this analysis can also be seen that selecting an object by touching is more direct than selecting it by pointing, which in turn is more direct than selection by scanning.

6.2 About the Choice of Selection Technique

English *et al.* [1967] state that important factors in the choice of selection techniques are 1) the mix of other operations required of the select-operation hand, 2) the ease of getting the hand to and gaining control of a given selection device, and 3) the fatigue effects of its associated operating posture. Although their experiments were based on an early desktop computer and different selection devices for it, the basic principles can be thought to be valid in physical selection techniques as well. Especially the fatigue effects may prove important with touching and pointing. Additionally, the first item in the list – the mix of operations – could be presumed to affect the selection techniques. For example, the basic pointing sequence includes several operations: first aligning the terminal towards the target, then pressing a button to trigger the selection, and finally bringing the terminal back close enough to see the result of the selection.

Foley *et al.* [1984] see time, accuracy and pleasantness as the primary criteria for the quality of interaction design. These criteria will be met in different degrees by the same interaction technique, depending on 1) the context of the task, 2) the experience and knowledge of the user, and 3) the physical characteristics of the interaction devices. Therefore interaction techniques "should not be selected in isolation from knowledge of the other techniques at use at approximately the same time. The performance of a complete action generally involves a series of tasks to be carried out almost as a single unit" [Foley *et al.* 1984]. The context of interaction techniques will be especially relevant in ubiquitous computing environments in which 1) the interaction is intended to be as natural as possible, and 2) the interaction is more unpredictable than in constrained desktop environments.

6.3 Selection in Immersive Virtual Environments

Selecting an object for further interaction is one of the most fundamental actions in a virtual environment [Robinett and Holloway, 1992; Mine, 1995; Bowman and Hodges, 1997]. Interaction with virtual objects requires some way to indicate to the system the target of the interaction.

Virtual environments typically support direct user interaction, which allows the use of interaction techniques such as hand tracking, gesture recognition, pointing and gaze direction to specify the parameters of the interaction task, including object selection [Robinett and Holloway, 1992; Mine, 1995]. This interaction paradigm supports well natural mapping between the actions of the user and the results of the action in the virtual environment. In these kinds of environments, object selection can be performed directly: if an object is within the reach of the user, the user can select the object by extending a hand and touching it, in a similar manner to touching objects in physical selection. Selection in virtual environments typically involves using a button or a gesture to signal the system that the chosen object is the one the user wants to select.

The two primary selection technique categories in virtual environments are local and at-a-distance [Mine, 1995; Bowman and Hodges, 1997]. If the object is out of the reach, an at-a-distance selection technique is needed. Selecting a remote object for manipulation in a virtual environment bears some resemblance to selecting a remote object for interaction in a physical environment. Remote object selection can be done by an arm-extension technique or a ray-casting technique. Arm extension requires a graphical system to show how far the virtual hand of the user is extended and thus is not suitable for physical environments without augmented reality displays. Ray-casting, on the other hand, is very similar to physical selection by pointing. In ray-casting, a light ray (such as a virtual laser beam or spotlight) projects from the hand of the user, typically when a specific button is pressed. By intersecting the ray with the desired object, and releasing the button, the object is attached to the ray and is ready for manipulation. Other techniques for selecting objects in virtual environments include gaze direction, head orientation, voice input and list selection, which is similar to physical selection by scanning.

According to Mine [1995], feedback is essential for selection in virtual environments. The user must know when the object is ready for selection, that is, when for example the pointing beam is intersecting with an object. The user must also get feedback about a successful selection. Bowman and Hodges [1997] state that it is important to consider selection and manipulation as separate issues, which supports well our view of selection and action as orthogonal phases of interaction.

6.4 Selection with Laser Pointers

In addition to physical selection by pointing, laser pointers have been used as input devices for large projected screens. As ubiquitous computing becomes more common, rooms will contain computer-controlled devices, appliances and displays [Weiser, 1993]. In these non-desk situations, users need to interact with display surfaces at a distance [Olsen and Nielsen, 2001]. For people who are distant from the screen, the most natural way to select objects is to point at them [Myers *et al*., 2001]. A popular interaction scheme is to target the projection with a camera to track a laser pointer beam on the display surface. The drawback of this approach with regards to tag-based physical selection is that the target area must be tracked with a camera, meaning it can be used only in suitably instrumented environments.

Kirsten and Müller [1998] have developed a system in which a common laser pointer can be tracked by video camera. The pointer works as an input device and frees the user from the location of the traditional input devices. The user uses the laser pointer by turning it on, off and moving the point on the display area while the pointer is turned on. A camera is used for detecting the point induced by the laser pointer on the projection screen. The image is analysed and mouse control signals (up, down, move) are derived from the detected beam and sent to the computer. The applications in the computer do not notice any difference whether a regular mouse or a laser pointer is used.

Olsen and Nielsen's [2001] approach is similar to Kirsten and Müller's approach, but they argue that such a simple mapping (mouse up/down, move) is not sufficient for general information manipulation. They describe a full suite of interaction techniques and an implementation to accomplish the interactions.

Similarly to the system by Kirsten and Müller, users have laser pointers and can use them to interact with information on a large projected display. The difference is that Olsen and Nielsen do not try to directly translate laser pointer events to mouse events. In addition to laser on, laser off and move events they also have events for the laser being off for an extended period of time and for the laser being held in a place for an extended period. They have developed laser pointer specific user interface widgets that respond to these new events. Olsen and Nielsen divide the widgets into button, enumeration, scrollable (numbers, dates and times), text and list categories. Their view is thus that interaction with a laser pointer requires a specific user interface from the application on the projected display. Olsen and Nielsen performed a user study. Their results indicate that laser pointer based interactions on a projected screen are clearly slower than using a mouse.

Myers *et al*. [2002] state that interaction techniques using laser pointers tend to be imprecise, error-prone and slow. In addition to technological problems (for example there is no mouse button in a common laser pointer), inherent human limitations cause difficulties for laser pointer interaction:

1. users will not know exactly where the laser beam will hit when they turn it on, and it takes about one second to move it into the desired position,

2. the hands are unsteady, which causes the beam to wiggle on the magnitude of 0.10–0.17 degrees (mostly vertically) depending on the design of the pointing device (a PDA was found to be the most stable pointing device), and

3. when the button is released, the beam often shifts away from the target before going off.

Therefore, a new interaction style is needed. Myers *et al*. [2001] suggest a different approach to previous work. Their idea is to use pointing for referencing a broad area of interest. The laser pointer is used to indicate the region of interest and the item in the region is copied to the user's hand-held device. They call this interaction style *semantic snarfing*. Myers *et al*. use the word "snarfing" [Raymond, 1996] to refer to grabbing the contents of a large screen to a hand-held device. Their use of the word 'semantic' refers to the fact that the meaning, or semantics, of the grabbed objects is needed often instead of a picture or an

exact copy of the interface, because the interaction capabilities of the hand-held devices are limited. In addition to interacting with displays, the users can also "snarf" the user interface of a remote appliance to their mobile device, which is similar to physical selection by pointing.

The human limitations Myers *et al*. describe also apply to physical selection by pointing. It does not matter whether the user is trying to point at a widget on a projected display or a physical object, they do not know where the laser beam will be when they first turn it on (at least until they have had considerable practice with their personal device). If the laser beam alone is used to trigger a pointable tag, the problem of shaking hands makes the selection more difficult, as does the potential shift when the pointing button is released. Myers *et al*. state that because of the wiggle, the graphical user interface widgets designed for laser pointer interaction should be fairly big. This will hold true also with pointable tags if only laser triggering is used, meaning that the tag should for example have an array of photosensitive sensors. To counteract some of the problems Myers *et al.* describe, we decided to trigger a tag in our pointing implementation as soon as the laser beam hit it. The initial location of the beam is still somewhat problematic, because if there happens to be a tag under the beam when it is still in a wrong location, the wrong tag will be triggered.

6.5 Mobile Terminal as an Input Device

Ballagas *et al.* [2006] use the taxonomy of Foley *et al*. [1984] as a framework for their analysis of mobile phones as ubiquitous computing input devices. Ballagas *et al.* note that although Foley *et al.* performed their analysis in the desktop graphical user interface context, the interaction tasks they found, can be applied to ubiquitous computing. The most interesting interaction task in the analysis is that of selection. In ubiquitous computing, selection is used to select a physical object to operate upon. Ballagas *et al.* [2006] divide the selection task into following interaction techniques:

- direct pick of tagged objects (the user selects a tag with a tag reader equipped mobile terminal),

- direct pick with camera (the user selects the target using a camera embedded in the mobile terminal),

- direct pick by laser pointer (the user selects the target by pointing at it with a laser pointer),

- voice recognition (the target is selected based on the utterance of the user), and

- gesture recognition (the target is selected based on some physical gesture of the user).

The direct pick techniques correspond most closely to physical selection as they are defined in this dissertation. Ballagas *et al.* [2006] define the mobile RFID reader system by Want *et al.* [1999], the laser pointer activated photosensitive tags [Patel and Abowd, 2003] and the RFIG Lamps [Raskar *et al.*, 2004] as 'direct pick of tagged objects'. Common to all these systems is that physical objects contain identification tags and these tags can be read with a mobile device, effectively selecting the physical object via the tag.

The 'direct pick with camera' technique is similar to 'direct pick of tagged objects', but Ballagas *et al.* [2004] choose to differentiate between visual and electronically readable tags. Although they do not mention it in the article, the tag reading with camera differs somewhat from that with an RFID tag reader. A camera-based interaction typically shows on the mobile terminal screen the visual tag to help the user align the tag to the camera field of vision. RFID tag readers are somewhat less sensitive to alignment of the reader relative to the camera, although the readers generally work better if the tag antenna and the reader antenna are aligned parallel to each other.

The third direct pick technique, 'direct pick by laser pointer' considers such systems as semantic snarfing by Myers *et al.* [2002] (see previous subsection). Typical for this technique is that the laser pointer beam is monitored with a camera in the ubiquitous computing environment.

7. Introduction to the Themes of the Publications

This dissertation consists of nine publications on physical browsing, and especially on physical selection within ubiquitous computing environments. There are three themes within the dissertation: the selection interaction task itself, visualising physical hyperlinks to support the selection, and physical selection as an interaction task in a set of ubiquitous computing interactions. The relation of the papers is illustrated in Figure 6.
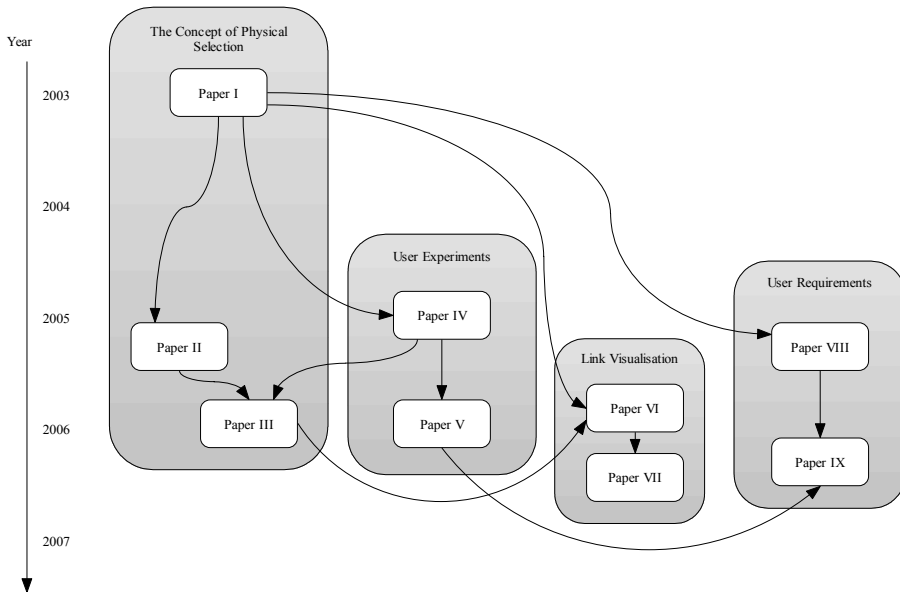
Figure 6. The relation of the papers to each other and to the themes of this dissertation.

7.1 Physical Browsing Concept

In Paper I, "A User Interaction Paradigm for Physical Browsing and Near-Object Control Based on Tags", we present our vision of physical browsing and three physical selection methods, touching, pointing and scanning, which are described in this dissertation in section 2.5. This paper gives initial answers to the question on how to provide the user with a usable way to select an object in a physical environment in order to access the digital counterpart of the physical object.

The main contributions of the paper are the description of physical browsing and the selection methods, and an analysis of implementation possibilities. We analyse the most important features of passive RFID technology in relation to how to implement the selection methods and show that all three selection methods can be implemented with RFID technology. Touching is already available in near-field based RFID implementations. Scanning can be implemented with long-range RFID technology and pointing with long-range RFID technology augmented with a sensor interface on the tags.

In addition to the selection methods, we explore the possibility to implement a mobile phone based universal remote control with RFID technology. Two other scenarios show the potential of augmenting physical objects with digital information: information retrieval and a shopping assistant. We also identify the initial usability questions about visualising physical hyperlinks:

- How can the user find out if there are tags in the environment and recognise tagged objects from those not tagged?

- What would the tag do if it was selected?

- How do we communicate these issues to the user?

These questions created the need to study visualising physical hyperlinks. The questions are studied in papers VI and VII.

We expand the theme of the first paper in Paper II, "Bridging the Physical and Virtual Worlds by Local Connectivity-Based Physical Selection" and provide more depth to the question on how to create a natural and direct mapping between a physical object and its digital counterpart. We further define the selection methods and analyse the implementation possibilities in more depth. We also look at the actions that follow physical selection: getting information from physical objects and connecting devices to each other using the existing local connectivity functionality within mobile devices. In this paper, we recognise the independence of selection method and the action following selection. My work in this paper was concentrated on analysing the selection methods.

In Paper III, "Physical Browsing", we further expand on the themes of the previous two papers and define central terms of physical browsing. In our terms,

the user uses *physical selection* to select a *link* in a physical *object*, and the mobile terminal reads an *information tag* attached to the object and launches some *action*. Physical selection can be initiated by the user, which is the case in *touching*, *pointing* and *scanning* selection methods, or it can be initiated by the tag, as is the case in the *notifying* method.

In this paper, we also examine the relation of physical browsing and tangible user interfaces and context-aware interfaces and show that physical selection is a form of a tangible user interface as discussed in Chapter 4. We show that physical selection can be used to set a part of the context in context-aware systems, and context-awareness can be used to enhance physical selection. We also describe in more detail the various tagging technologies that enable physical selection and browsing. These technologies include RFID, visual codes, infrared and Bluetooth. The technologies were described in this dissertation in Chapter 3.

7.2 User Experiments

In Paper IV, "RFID Tag Reader System Emulator to Support Touching, Pointing and Scanning", we present our mobile terminal and tags that support the three selection methods within the same device. The previous systems had supported touch-only selection for example with RFID-enabled mobile phones, point-only selection for example with infrared technology or scan-only selection for example with Bluetooth. The main contribution of the paper is a description of a physical browsing system that supports all three selection methods in the same device allowing the user to choose the selection method most appropriate to the situation. The motivation for building the system was to be able to study experimentally the usability of physical selection. The system was implemented by Timo Tuomisto and my main contribution was in its interaction design.

The system uses active SoapBox devices as sensor-equipped tags, but we built the system so that it emulates the use of passive RFID tags from the user's point of view. To study the different touching techniques, the system supports touching by either bringing the terminal close to the tag, or pressing a button at the same time to indicate intention to select the tag. To study the different pointing techniques, the system supports pointing with a narrow and visible laser beam, wide and invisible infrared beam, and a combination of both. The

selection could be configured to ask for confirmation to activate the link after selection to study the user preference for confirmations.

The system consists of a PDA as the mobile terminal, a central SoapBox connected to the PDA as tag reader, and several remote SoapBoxes as tags. The remote SoapBoxes were equipped with a light sensor to detect visible light (laser) and an infrared transmitter and receiver. The central SoapBox could emit a brief infrared pulse, which the IR receiver of the remote SoapBox detected. The infrared receiver with a transmitter was used as a proximity sensor to detect touching.

The tag emulator system was used for user experiments and the results are presented in Paper V, "Evaluating Touching and Pointing with a Mobile Terminal for Physical Browsing". The specific research questions we studied in the experiment were

- Are touching and pointing usable selection methods and are both of them needed?

- If both selection methods are needed, what is the threshold distance at which the user chooses between touching and pointing?

- Should touching use an explicit action such as a button press in order to tell the terminal the user intention of selecting a link in addition to just bringing the terminal close to the tag?

- Which pointing technique, laser, infrared or the combination of both, is most usable?

- Are confirmations needed for activating the link?

In that paper, we found touching and pointing to be easy to learn and use, and to complement each other. The threshold distance was 1.1 metres, that is, below that distance the users choose touching and for targets beyond that distance they choose pointing. The distance did not depend on whether the user was sitting or standing.

The users preferred touching to be continuously on compared to touching activated with a button press. Effortlessness was the most important difference between the two configurations. A button press in conjunction with the touching

gesture increased the sense of security of touching. Otherwise their usability qualities were perceived similar.

The combination of infrared and laser pointing was preferred by the users. Laser-only pointing required very accurate aiming and the infrared-only pointing technique was more difficult because of lack of visual feedback.

The users wanted to open simple information links as effortlessly as possible without confirmations and intra-terminal operations. Therefore the tag visualisations should support this single most important case: the simple free-of-charge information content must open with a minimum of confirmations or menu selections.

7.3 Visualising Links

During our studies, we found visualising the links in the physical environment an important research theme. We describe the problem domain and preliminary suggestions in Paper VI, "Suggestions for Visualising Physical Hyperlinks". The initial research questions for visualisations from Paper I are refined in Paper VI as how can we communicate to the user

1. the existence of a link,

2. the precise location of the link,

3. the supported selection methods, if the link does not support all selection methods, and

4. the action of the link?

One motivation for this work was that in the evaluation described in Paper V, the users wanted to open simple information links as effortlessly as possible without confirmations and intra-terminal operations. This would require the visualisation of the link to include enough information for the user to be able to make the decision whether the link contains the desired information.

The contribution of this paper includes the formulation of the research questions and initial analysis of the usability issues, based on our previous work

(especially papers III and VIII) and link visualisation in the World Wide Web. We introduce the idea of hovering with a mobile terminal over the link to effortlessly display information about the link before actually activating it.

The main suggestions in the paper are that the tag visualisations would benefit more from visualising the selection method as opposed to selection technology as is done in the current systems, and from visualising also the action. Some example visualisations are presented to illustrate the ideas on how to combine the selection method and action into an icon.

In Paper VII, "Hovering: Visualising Physical Hyperlinks in a Mobile Phone", I describe my implementation of the hovering visualisation concept in a mobile phone. Hovering gives the user pre-selection information about the action of the tag before actual selection. In this paper I describe the concept and how it was implemented using NFC-enabled mobile phone and RFID tags. The system allows including in the visualisation information that would be impractical or hard to include in the physical visualisation of the tag, such as a URL and a title for the link.

7.4 Physical Selection in an Ambient Intelligence Architecture

This physical selection work has been a part of a larger user interaction concept of ubiquitous computing architecture developed in EU projects Mimosa and Minami. In Paper VIII, "Ambient Functionality – Use Cases" we describe our work with the use cases and user requirements of the architecture. By analysing and evaluating ubiquitous computing scenarios created for the architectures, we found that the mobile applications that use local connectivity share many common user interaction patterns. Physical selection was identified as a recurring use case in the scenarios, which gave us an overview of the possible ways of using physical selection as a user interaction task.

Based on evaluation of the scenarios we present recurring use cases for the scenarios and user requirements for the use cases. These classes of use cases as actions also provide for basis of link visualisations in Paper VI. The physical

selection methods are presented as separate use cases. The classes of action use cases are:

- viewing simple information content,
- activating applications,
- collecting sensor data, and
- context-awareness.

User requirements for physical selection were initially found by evaluating the application scenarios with users. The user requirements we found from the scenario evaluations were:

- The tags and sensor radio nodes should support all three physical selection methods and behave consistently in the interaction to support users adopting the new user interaction concept.

- The tags should be marked in a consistent way so that the user can identify them. The appearance of the tag should indicate the functions that are included in it.

- The reading should be quick and reliable without the user having to wave the mobile terminal back and forth.

- Taking applications and equipment into use should be done easily with as little manual configuration as possible.

These user requirements for physical selection are very broad but they provided us a basis for designing the user interaction and formulating the research questions that were studied in Paper V.

The user requirements work continues in Paper IX, "Identifying User Requirements for a Mobile Terminal Centric Ubiquitous Computing Architecture". In this paper we further refine our user requirements for the architecture, including user requirements for physical selection. The requirements in this paper are based on the user experiments described in Paper V and in scenario evaluations with users.

We identify two classes of user requirements for physical selection, general and selection method specific. The specific user requirements for touching and

pointing are based on the results of the evaluations from Paper V and presented in this dissertation in section 7.2. From the scenario analysis, we identify the following general user requirements for physical selection:

- All three selection methods are needed. In an environment with many tags, it will be difficult to select the correct one from the list presented after scanning. Therefore there is a need for touching and pointing.

- For touching and pointing the user needs to know the location of the tag. Therefore the location of the tag needs to be made known to the user.

- Physical selection should allow actions be activated without the user first starting an application to interpret the tag contents. There should be a default action depending on the content type of the tag. For example, if the tag contains a URL, the terminal should open the web browser and display the page.

My main contribution in these two papers has been the sections about physical selection. Other requirements were identified for wireless measurements, context-awareness, taking applications into use and ethical issues.

8. Discussion

In the previous Chapters, physical selection in the contexts of ubiquitous computing and computer-augmented environments has been discussed. In addition, how physical selection compares to physically-based and tangible user interfaces has been explored, and how physical selection relates to the selection interaction task in other domains such as desktop GUIs, virtual environments and laser pointer interaction. Physical selection is presented in this dissertation in a manner that combines the different selection methods and technologies into a coherent aggregate, suitable to be used in a broader ubiquitous computing context, as described in papers VIII and IX.

The main contributions of this dissertation are:

1. a physical browsing and physical selection framework including different selection methods, actions and technologies, suitable for a variety of tasks in ubiquitous computing,

2. analysis on how to implement the different selection methods,

3. comparison of different techniques for touching and pointing,

4. suggestions for visualising the links in both the physical environment and in the mobile terminal using the hovering concept, and

5. user requirements for physical selection within a ubiquitous computing architecture.

8.1 The Need for Physical Selection Methods

The first research question presented in the Introduction is "*How to provide the user with a usable way to select an object in a physical environment in order to access the digital counterpart of the physical object?*". In Papers I–III we suggested physical browsing as a means for linking digital information to physical objects in similar way as Kindberg *et al.* [2002] used physical browsing in the Cooltown system and Want *et al.* [1999] in their RFID prototypes. We suggested *physical selection for the selection task for mobile terminal based ubiquitous computing*.

The second research question is "*What kinds of selection methods are needed in physical selection?*" The previous systems have typically included only one selection method. For example, Want *et al.* [1999] used touch-based RFID interactions, Swindells *et al.* [2002] used infrared for pointing, and WPAN technologies such as Bluetooth support selection by scanning. We see a need for different selection methods for different situations. Scanning is useful when the user does not know of the existence or the locations of the links, or only wants to check what is available in an environment. Touching is best for selecting a single link whose location is known and which is near the user. Pointing is useful for selecting a link in a similar way to touching but from a longer range.

The user requirements regarding the need of physical selection methods in Papers VIII and IX are as follows:

- All three selection methods are needed. In an environment with many tags, it will be difficult to select the correct one from the list presented after scanning. Therefore there is a need for touching and pointing

- The tags and sensor radio nodes should support all three physical selection methods and behave consistently in the interaction to support users adopting the new user interaction concept

- Taking applications and equipment into use should be done easily with as little manual configuration as possible. The implication of this requirement is that there should be a method of easily pairing equipment with the mobile terminal and that downloading applications into the terminal should be easy. Physical selection of remote devices and download links is a simple way to provide this functionality to the user.

Other studies have come to the conclusion that different selection methods are usable and needed. Rukzio *et al.* [2006] found out that the users preferred touching and pointing over scanning because of the directness of touching and pointing, which supports our view on the purposes of the single link selection methods, and the varying directness of the selection methods, as described in Chapter 6. Riekki *et al.* [2006] have shown touching to be a quick, easy and comfortable method for selecting a target for interaction. Pohjanheimo *et al.* [2004] and Swindells *et al.* [2002] showed pointing to be preferred by the users over typing the address of the service or even selecting the target from a menu if

the menu contained several items. This supports our view that of the long-range selection methods, pointing is suitable for situations in which the user already knows which link is to be selected. Scanning, on the other hand, is not sensitive to direction and works well when the user is just exploring a physical space and does not yet know what to point at.

In our experiments described in Paper V, we measured that the borderline distances for switching between touching and pointing is 1.1 metres, which shows that the users switch between two selection methods depending on their location. Rukzio *et al.* [2006] studied also other factors on the choice of selection method and provided guidelines on which selection methods would be useful in which context. However, our view is that the user should have the power to choose the selection method that is appropriate for her immediate situation and personal preferences. Based on these results and our own experiments it can be concluded that *touching, pointing and scanning are needed in different situations and all three should be provided for the user*.

8.2 Implementing Physical Selection

The third research question in the Introduction is "*How the selection methods should be implemented from the point of view of the user?*" This question considers the user requirements for the individual selection methods of touching and pointing. Scanning as a more indirect selection method was not studied experimentally in this dissertation, but the need for scanning is recognised, as described in the previous section. Several more defined research questions within this theme were studied and they are presented in the following subsections. The results of these studies are compared to those of others in the following subsections.

8.2.1 Touching

As physical selection in general and touching in particular was already shown to be a useful interaction method (for example Riekki *et al.* [2006]), we studied only how to best implement touching. Our research question regarding touching was "*Should touching use an explicit action such as a button press in order to*

tell the terminal the user intention of selecting a link in addition to just bringing the terminal close to the tag?". The users preferred selection without any other action (button press in our system) and felt that in most cases just bringing the terminal close to the tag was explicit enough. The inclusion of a button press was reported to increase the feeling of security.

However, touching being always on could cause problems, for example if the terminal is put into a purse full of tagged items, the terminal would then continuously read the tags. It seems that there is no clear reason to choose or reject either touching technique, but to let the user configure it, and to let touching without a button to be an easily activated separate mode. Therefore *a touch mode that can be activated and deactivated either manually or automatically would be preferable*.

In addition to the experimental results, the users mentioned in the scenario evaluation that *the reading of tags should be quick and reliable without the user having to wave the mobile terminal back and forth*. To further study touch-based interactions and visualisations related to it, the hovering prototype was built, and user experiments with it will be conducted in the future to find out how the tag visualisations in the physical world in the terminal affect the touching preferences.

8.2.2 Pointing

Our research questions considering pointing were 1) "*Which pointing technique, laser, infrared or the combination of both, is most usable?*" and 2) "*What is the required reading distance for a pointable tag?*". Because we had already established the need for pointing, we did not question the utility of pointing but instead delved deeper into what kind of pointing technique would be user-friendly. The techniques we chose were based on the questions of 3) *whether a narrow or wide pointing beam* (as in RFIG lamps that allows "casual interaction" [Raskar *et al.*, 2004]) *would be preferable* and 4) *whether visual aid is necessary*.

We found that the users preferred the combination of a wide invisible IR beam and a narrow visible laser beam since it provided the casuality for pointing but allowed them so see exactly where they were pointing at. Pohjanheimo *et al.* [2004] noticed in their IR studies that the lack of visual aid hindered the success

of their users in pointing, just as happened with our IR-only technique in the experiment. The laser-only technique suffered from the same problems described by Myers *et al.* [2002], such as difficulty of knowing where the beam will initially hit, the unsteadiness of hands and the shifting of beam when a button is released. The conclusion is that preferably *the pointing system should include a wider beam for casual pointing but also a visible aid to help in aiming* (questions 1, 3 and 4). It might be useful to temporarily switch off this visual aid for privacy purposes in public places.

Because the threshold distance for switching from touching to pointing was measured to be 1.1 metres, *the reading range of a pointable tag and reader combination should significantly exceed 1.1 metres*. In addition, *if the pointing technology requires some kind of distance calibration, it should be calibrated to significantly more than 1.1 metres* (question 2).

8.2.3 Action

In Paper IX, we recognised the need for a "default action" after physical selection based on tag contents. Physical selection should allow actions be activated without the user first starting a specific application to interpret the tag contents. Therefore there should be a default action depending on the content type of the tag. For example, if the tag contains a URL, the terminal should open the web browser and display the page. The implications to the terminal architecture were thus that 1) *there should be a tag-reading application running in the terminal all the time*, and 2) *it should be able to determine an action for the tags the user selects*. An implication to the tag data architecture is thus that *the tags could include meta information about their contents*.

8.3 Visualising Physical Hyperlinks

The fourth research question of this dissertation is "*How the physical hyperlinks should be visualised so that they support user-friendly physical selection?*". Four main characteristics to visualise were found to be:

1.   the existence of a link,

2. the precise location of the link,

3. the supported selection methods, if the link does not support all selection methods, and

4. the action of the link.

Riekki *et al.* [2006] visualised their touch-based tags with icons describing the action and we have extended the visualisation to include the different selection methods since our work included other methods in addition to touching. Arnall [2006] has explored icons describing touching, both for physical selection and other purposes such as smart cards. Belt *et al.* [2006] found that first-time users were surprised when selecting a tag caused an unexpected action when the action was not clear from the context or the visual look of the tag.

Considering these results, to give the user information on all four questions, *both selection method and action should be visualised* if the user is to get a clear idea of how to interact with an object and what to expect from it. This is very different from the current methods, in which generally the selection technology (if anything) is visualised with, for example, NFC logo.

Hovering gives the user pre-selection information about the action of the tag before actual selection. It may be helpful alongside the other forms of link visualisations, and it will be studied in more depth in the future.

8.4 Future Work

I am continuing my work with physical selection in several directions. In the Minami EU project, a passive mass memory RFID tag is being developed. This kind of tag that can contain several megabytes of information or connect to sensors and be read with a mobile phone opens interesting directions for interaction research. One research challenge is how a device whose selection technology and data transmission technology are the same, works from the user's point of view. Touching, as I and others have shown, is a simple and comfortable selection method and it works well with RFID technology. However, for data transmission from the tag, the user needs to keep the mobile phone within reading range for the whole communication duration, which for

large files may be several seconds, even minutes. Additionally, the user interface for a tag that can contain several individual data items is challenging, especially considering the visualisation and data naming issues. This will also link to my visualisation research and I intend to conduct user experiments with the hovering prototype described in paper VII.

My visualisation work will continue to include aural "visualisations" for physical selection. Hovering could include semantic sounds, or "earcons", giving the user additional information about the link being explored. This work will be done together with the University of Jyväskylä and we intend to deploy a large-scale real world physical browsing application to study the earcons, visualisations and physical selection in general.

9. Conclusions

In this dissertation, I have addressed physical selection as an interaction method for ubiquitous computing. Ubiquitous computing is characterised by the embedding of computational devices into the physical environment, and transferring the interaction from the purely virtual world of the desktop to the physical world in which we live and already interact with everyday objects. Physical selection is an interaction task in which the user of a mobile terminal uses a physical gesture – touching, pointing or scanning – to tell the terminal of which object the user is interested.

Physical selection is enabled by various tagging technologies such as RFID or visual codes that can be used to identify the physical objects to the mobile terminal of the user. In addition to simple identification, the tag may contain information about the object such as a web address directing the terminal to a web page about the object. The user has a mobile terminal that is equipped with a reader that can read these tags, for example an RFID reader for RFID tags or a camera for visual tags.

After the user has selected the object or tag for interaction, an action occurs. The action is typically related to the physical object, for example a tag in a movie poster may contain a link to the WWW page of the movie, or a link in a bus timetable may direct the user to an online timetable service. Other possible actions include changing the terminal state (for example to silent mode) or connecting to some local service, for example via Bluetooth. Together physical selection and action form a user interaction paradigm called physical browsing.

Three selection methods have been described in detail: touching, pointing and scanning. Touching and pointing are intended for selecting a single tag from a known location whereas scanning can be used to explore all tags within the reader's range, or when the exact location of a tag is not known. Touching is the most direct and unambiguous selecting method, and requires the user to bring the terminal to virtually touching distance of the target. Pointing is a longer ranged directional method and requires the user to align the terminal towards the target, in a similar way as a TV remote control is used. The need for all three selection methods has been discussed.

The most important local connectivity technologies that either can be used to implement physical selection, or that benefit from added physical selection, have been described. RFID, NFC, infrared and visual codes are technologies that enable physical selection. Wireless personal area networking technologies such as Bluetooth can also be used to implement selection by scanning, but these technologies benefit greatly from added physical selection by touching and/or pointing.

To explore the selection methods, the evolution of physically-based user interfaces and physical selection from the graphical desktop metaphor to the tangible user interfaces has been discussed. Physical selection can be seen as one kind of tangible user interface, with different levels of embodiment and metaphor. The tagged physical objects map to tokens and containers in a taxonomy for token-based tangible user interfaces. Other physical browsing systems were described, illustrating the development of the concept and how it has affected our work.

Selection is a common interaction task in practically all domains. Selection in graphical desktop user interfaces and immersive virtual environments has been analysed to benefit from the lessons learned from these more established domains. Laser pointers have been used to interact with projection screens and their usability issues are very similar to physical selection by pointing. Selection task in more recent mobile phone based ubiquitous computing systems has also been described.

The idea of physical selection, the selection methods and implementation possibilities were built gradually over time and in several publications. The publications one to three reflect this gradual approach in which we refined our vision according to studies conducted by us and by others.

To study in more detail touching and pointing we have built a prototype system that implements all three selection methods and emulates the interaction with passive long-range sensor-equipped RFID tags that can support all the selection methods. We have conducted user experiments with this system and our results have been compared to those of other studies.

Physical hyperlink visualisation is an important issue in the usability of any physical browsing system. In this dissertation it was shown that the links should preferably be visualised in both the physical environment and in the mobile terminal, and that both the selection method and the action of the link should be shown. Additionally, the user could benefit from being able to explore the contents of a link by hovering the mobile terminal near it, giving the user "pre-selection" information about the link.

The mobile phone and PDA are transforming from a remote communication device into a local communication aware device. Most mobile phones can already read visual codes and communicate with local services via Bluetooth or WLAN and with remote service via cellular networks, and RFID readers are finding their way into mobile phones. All this encourages us to continue and development and research in the area of physical browsing.

References

[Abowd and Mynatt, 2000] Gregory D. Abowd and Elizabeth D. Mynatt. Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction 7(1)*, ACM Press, 2000, 29–58.

[Ailisto *et al.*, 2003a] Heikki Ailisto, Aija Kotila and Esko Strömmer. Ubicom applications and technologies. *VTT Research Notes: 2201*, VTT Electronics, 2003. http://www.vtt.fi/inf/pdf/tiedotteet/2003/T2201.pdf.

[Ailisto *et al.*, 2003b] Heikki Ailisto, Ilkka Korhonen, Johan Plomp, Lauri Pohjanheimo and Esko Strömmer. Realising physical selection for mobile devices.

[Arnall, 2006] Timo Arnall. A graphic language for touch-based interactions. *Proc. Mobile Interaction with the Real World*, 2006, 18–22.

[Ballagas *et al.*, 2006] Rafael Ballagas, Jan Borchers, Michael Rohs and Jennifer G. Sheridan. The smart phone: a ubiquitous input device. *Pervasive Computing 5(1)*, IEEE, 2006, 70–77.

[Belt *et al.*, 2006] Sara Belt, Dan Greenblatt, Jonna Häkkilä and Kaj Mäkelä. User perceptions on mobile interaction with visual and RFID tags. *Proc. Mobile Interaction with the Real World*, 2006, 23–26.

[Bowman and Hodges, 1997] Doug A. Bowman and Larry F. Hodges, An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. *Proc. Symposium on Interactive 3D graphics*, ACM Press, 1997, 35–38.

[Cambridge Consultants, 2006] Cambridge Consultants. *Simple Device Pairing by Relative Signal Strengths*. Advertising material, 2006.

[Card *et al.*, 1990] Stuart K. Card, Jock D. Mackinlay and George G. Robertson. The design space of input devices. *Proc. Conference on Human Factors in Computing Systems*, ACM Press, 1990, 117–124.

[Coulouris *et al.*, 2005] George Coulouris, Jean Dollimore and Tim Kindberg, *Distributed Systems: Concepts and Design*, 4th ed. Addison–Wesley, 2005.

[Ecma, 2004] Ecma. *Near Field Communication White Paper*, ECMA/TC32-TG19/2004/1, Ecma International, 2004.

[Elrod *et al.*, 1992] Scott Elrod, Richard Bruce, Rich Gold, David Goldberg, Frank Halasz, William Janssen, David Lee, Kim McCall, Elin Pedersen, Ken Pier, John Tang and Brent Welch. Liveboard: a large interactive display supporting group meetings, presentations and remote collaboration. *Proc. ACM Conference on Human Factors in Computing Systems*, ACM Press, 1992, 599–607.

[English *et al.*, 1967] William K. English, Douglas C. Engelbart and Melvyn L. Berman. Display-selection techniques for text manipulation. *IEEE Transactions on Human Factors in Electronics 8(1)*, IEEE, 1967, 5–15.

[Fishkin, 2004] Kenneth P. Fishkin. A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing, 8(5)*, Springer-Verlag, 2004, 347–358.

[Finkenzeller, 2003] Klaus Finkenzeller. *RFID Handbook*, second edition. John Wiley & Sons Ltd., England, 2003.

[Fitzmaurice *et al.*, 1995] George W. Fitzmaurice, Hiroshi Ishii, and William Buxton, Bricks: Laying the foundations for graspable user interfaces, *Proc. CHI'95*, ACM Press, 1995, 442–449.

[Foley *et al.*, 1984] James D. Foley, Victor L. Wallace and Peggy Chan. The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications, 4(11)*, IEEE Computer Society Press, 1984, 13–48.

[Haartsen *et al.*, 1998] Jaap Haartsen, Mahmoud Naghshineh, Jon Inouye, Olaf J. Joeressen, Warren Allen, Bluetooth: vision, goals, and architecture, *Mobile Computing and Communications Review, 1(2)*, ACM Press, 1998, 38–45.

[Holmquist *et al.*, 1999] Lars-Erik Holmquist, Johan Redström and Peter Ljungstrand, Token-based access to digital information. *Proc. 1st International Symposium on Handheld and Ubiquitous Computing*, Springer-Verlag, 1999, 234–245.

[Ishii and Ullmer, 1997] Hiroshi Ishii and Brygg Ullmer, Tangible bits: towards seamless interfaces between people, bits and atoms. *Proc. SIGCHI Conference on Human Factors in Computing Systems*, ACM Press, 1997, 234–241.

[Kaasinen *et al.*, 2005] Eija Kaasinen, Timo Tuomisto and Pasi Välkkynen, Ambient functionality – use cases, in *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, ACM Press, 2005, 51–56.

[Kaasinen *et al.*, 2006] Eija Kaasinen, Marketta Niemelä, Timo Tuomisto, Pasi Välkkynen and Vladimir Ermolov, User Requirements for a Mobile Terminal Centric Ubiquitous Computing Architecture, in *Proc. Workshop on System Support for Future Mobile Computing Applications*, IEEE, 2006, 9–16.

[Kay and Goldberg, 1977] Alan Kay and Adele Goldberg. Personal dynamic media. *Computer 10(3)*, IEEE, 1977, 31–41.

[Keränen *et al.*, 2005] Heikki Keränen, Lauri Pohjanheimo and Heikki Ailisto. Tag manager: a mobile phone platform for physical selection, *Proc. International Conference on Pervasive Services*, IEEE, 2005, 405–412.

[Kindberg *et al.*, 2002] Tim Kindberg, John Barton, Jeff Morgan, Gene Becker, Debbie Caswell, Philippe Debaty, Gita Gopal, Marcos Frid, Venky Krishnan, Howard Morris, John Schettino, Bill Serra and Mirjana Spasojevic. People, Places, Things. *Mobile Networks and Applications 7(5)*, Kluwer Academic Publishers, 2002, 365–376.

[Kirsten and Müller, 1998] Carsten Kirstein and Heinrich Müller, Interaction with a Projection Screen Using a Camera-tracked Laser Pointer. In *Proc. Multimedia Modeling*, 1998, 191–192.

[Ljungstrand and Holmquist, 1999] Peter Ljungstrand and Lars Erik Holmquist, WebStickers: using physical objects as WWW bookmarks. *Proc. CHI'99 Extended Abstracts on Human Factors in Computing Systems*, ACM Press, 1999, 332–333.

[Ljungstrand *et al.*, 2000] Peter Ljungstrand, Johan Redström and Lars Erik Holmquist, WebStickers: using physical tokens to access, manage and share bookmarks to the web. *Proc. DARE 2000 on Designing Augmented Reality Environments*, ACM Press, 2000, 23–31.

[Mine, 1995] Mark R. Mine, *Virtual Environment Interaction Techniques*. Technical report TR95-018, University of North Carolina at Chapel Hill, 1995.

[Myers *et al*., 2001] Brad A. Myers, Choon Hong Peck, Jeffrey Nichols, Dave Kong, and Robert Miller, Interacting at a Distance Using Semantic Snarfing. In *Proc. 3rd International Conference on Ubiquitous Computing*. Lecture Notes in Computer Science; Vol. 2201. Springer-Verlag, 2001, 305–314.

[Myers *et al*., 2002] Brad A. Myers, Rishi Bhatnagar, Jeffrey Nichols, Choon Hong Peck, Dave Kong, Robert Miller, and A. Chris Long, Interacting at a Distance: Measuring the Performance of Laser Pointers and Other Devices, In *Proc. CHI 2002*, ACM Press, 2002, 33–40.

[Olsen and Nielsen, 2001] Dan R. Olsen Jr. and Travis Nielsen, Laser Pointer Interaction. In *Proc. CHI 2001*, ACM Press, 2001, 17–22.

[Opasjumruskit *et al.*, 2006] Karn Opasjumruskit, Thaweesak Thanthipwan, Ohmmarin Sathusen, Pairote Sirinamarattana, Prachanart Gadmanee, Eakkaphob Pootarapan, Naiyavud Wongkomet, Apinunt Thanachayanont and Manop Thamsirianunt, Self-powered wireless temperature sensors exploit RFID technology. *IEEE Pervasive Computing* 5(1), IEEE, 2006, 54–61.

[Patel and Abowd, 2003] Shwetak N. Patel and Gregory D. Abowd, A 2-way laser-assisted selection scheme for handhelds in a physical environment, *Proc. Ubicomp 2003, Lecture Notes in Computer Science 2864*, Springer-Verlag, 2003, 200–207.

[Philipose *et al.*, 2005] Matthai Philipose, Joshua R. Smith, Bing Jiang, Alexander Mamishev, Sumit Roy and Kishore Sundara-Rajan, Battery-free wireless identification and sensing, *Pervasive Computing* 4(1), IEEE, 2005, 37–45.

[Pohjanheimo *et al.*, 2004] Lauri Pohjanheimo, Heikki Ailisto and Johan Plomp. User experiment with physical pointing for accessing services with a mobile device, *Proc. EUSAI Workshop on Ambient Intelligent Technologies for Wellbeing at Home*, 2004.

[Pohjanheimo *et al.*, 2005] Lauri Pohjanheimo, Heikki Keränen and Heikki Ailisto. Implementing touchme paradigm with a mobile phone, *Proc. 2005 Joint Conference on Smart Objects and Ambient Intelligence*, ACM Press, 2005, 87–92.

[Raskar *et al.*, 2004] Ramesh Raskar, Paul Beardsley, Jeroen van Baar, Yao Wang, Paul Dietz, Johnny Lee, Darren Leigh and Thomas Willwacher. RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors. In *ACM SIGGRAPH 2004 Papers*, ACM Press, 2004, 406–415.

[Raymond, 1996] Eric S. Raymond. *The New Hacker's Dictionary, 3rd ed.*, MIT Press, 1996.

[Rekimoto and Nagao, 1995] Jun Rekimoto and Katashi Nagao. The world through the computer: computer augmented interaction with real world environments. *Proc. 8th Annual ACM Symposium on User Interface and Software Technology*, ACM Press, 1995, 29–36.

[Rekimoto and Ayatsuka, 2000] Jun Rekimoto and Yuji Ayatsuka, CyberCode: Designing Augmented Reality Environments with Visual Tags. *Proc. Designing Augmented Reality Environments* (DARE 2000), 1–10.

[Riekki *et al.*, 2006] Jukka Riekki, Timo Salminen and Ismo Alakärppä. Requesting pervasive services by touching RFID Tags. *Pervasive Computing 5(1)*, IEEE, 2006, 40–46.

[Robinett and Holloway, 1992] Warren Robinett and Richard Holloway, Implementation of flying, scaling, and grabbing in virtual worlds. *Proc. symposium on Interactive 3D graphics,* ACM Press, 1992, 189–192.

[Rohs and Zweifel, 2005] Michael Rohs and Philipp Zweifel, A conceptual framework for camera phone-based interaction techniques. *Proc. Pervasive Computing 2005,* Lecture Notes in Computer Science (LNCS) No. 3468, Springer-Verlag, 2005, 171–189.

[Rukzio *et al.*, 2006] Enrico Rukzio, Karin Leichtenstern, Vic Callaghan, Paul Holleis, Albrecht Schmidt and Jeannette Chin. An experimental comparison of physical mobile interaction techniques: touching, pointing and scanning. *Proc. Ubicomp 2006, LNCS 4206*, Springer-Verlag, 2006, 87–104.

[Smith *et al.*, 1982] D. Smith, C. Irby, R. Kimball, B. Verplank and E. Harslem. Designing the Star user interface, *Byte* 4/1982, 242–282.

[Stallings, 2000] William Stallings. *Data and Computer Communications, 6th Edition*. Prentice Hall, 2000.

[Stanford, 2003] Vince Stanford. Pervasive Computing Goes the Last Hundred Feet with RFID Systems, *Pervasive Computing 2(2)*, IEEE, 2003, 9–14.

[Streitz and Russell, 1998] Norbert A. Streitz and Daniel M. Russell, Basics of integrated information and physical spaces: the state of the art. *CHI 98 Conference Summary on Human Factors in Computing Systems*, ACM Press, 1998, 273–274.

[Streitz *et al.*, 1999] Norbert A. Streitz, Jörg Geiβler, Torsten Holmer, Shin'ichi Konomi, Christian Müller-Tomfelde, Wolfgang Reischl, Petra Rexroth, Peter Seitz and Ralf Steinmetz. i-LAND: an interactive landscape for creativity and innovation. *Proc. ACM Conference on Human Factors in Computing Systems*, ACM Press, 1999, 120–127.

[Strömmer and Suojanen, 2003] Esko Strömmer and Marko Suojanen. Micropower IR tag – a new technology for ad-hoc interconnections between hand-held terminals and smart objects, *Proc. Smart Objects Conference*, 2003.

[Swindells *et al.*, 2002] Colin Swindells, Kori M. Inkpen, John C. Dill and Melanie Tory. That one there! Pointing to establish device identity, *Proc. 15th annual ACM symposium on User interface software and technology*, ACM Press, 2002, 151–160.

[Toye *et al.*, 2004] Eleanor Toye, Anil Madhavapeddy, Richard Sharp, David Scott, Alan Blackwell, Eben Upton. *Using Camera-Phones to Interact with Context-aware Mobile Services*, Technical Report 609, University of Cambridge, 2004.

[Tuomisto *et al.*, 2005] Timo Tuomisto, Pasi Välkkynen and Arto Ylisaukko-Oja, RFID tag reader system emulator to support touching, pointing and scanning, in A. Ferscha, R. Mayrhofer, T. Strang, C. Linnhoff-Popien, A. Dey, A. Butz and A. Schmidt (eds.), *Advances in Pervasive Computing*, *Adjunct Proceedings of Pervasive 2005*, Austrian Computer Society, 2005, 85–88.

[Tuulari and Ylisaukko-Oja, 2002] Esa Tuulari and Arto Ylisaukko-oja. SoapBox: a platform for ubiquitous computing research and applications, *Proc. Pervasive 2002, LNCS 2414*, Springer-Verlag, 2002, 125–138.

[Ullmer and Ishii, 1997] Brygg Ullmer and Hiroshi Ishii. The metaDESK: models and prototypes for tangible user interfaces, *Proc. 10th Annual ACM Symposium on User Interface Software and Technology*, ACM Press, 1997, 223–232.

[Ullmer *et al.*, 1998] Brygg Ullmer, Hiroshi Ishii and Dylan Glas. mediaBlocks: physical containers, transports and controls for online media. *Proc. 25th Annual Conference on Computer Graphics and Interactive Techniques*, ACM Press, 1998, 379–386.

[Ullmer and Ishii, 2001] Brygg Ullmer and Hiroshi Ishii. Emerging frameworks for tangible user interfaces. In John M. Carroll, *Human-Computer Interaction in the New Millennium*, Addison–Wesley, 2001, 579–601.

[Välkkynen *et al.,* 2006] Pasi Välkkynen, Lauri Pohjanheimo, Heikki Ailisto, Physical Browsing, in Thanos Vasilakos and Witold Pedrycz (eds.), *Ambient Intelligence, Wireless Networking, and Ubiquitous Computing*, Artech House, 2006, 61–81.

[Want *et al.*, 1992] Roy Want, Andy Hopper, Veronica Falcão and Jonathan Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1), ACM Press, 1992, Pages 91–102.

[Want *et al.*, 1998] Roy Want, Mark Weiser and Elizabeth Mynatt. Activating everyday objects, *Proc. DARPA/NIST Smart Spaces Workshop*, 1998, 7-140–7-143.

[Want *et al.*, 1999] Roy Want, Kenneth P. Fishkin, Anuj Gujar and Beverly L. Harrison. Bridging physical and virtual worlds with electronic tags. *Proc. SIGCHI Conference on Human factors in Computing Systems*, ACM Press, 1999, 370–377.

[Want, 2004] Roy Want. Enabling ubiquitous sensing with RFID. *IEEE Computer*, 37(4), IEEE, 2004, 84–86.

[Want, 2006] Roy Want. An Introduction to RFID Technology. *Pervasive Computing*. IEEE, 2006, 25–33.

[Weiser, 1991] Mark Weiser. The Computer for the 21st Century. *Scientific American*, 265, 1991, 94–104.

[Weiser, 1993] Mark Weiser. Some Computer Science Issues in Ubiquitous Computing, *Communications of the ACM 36(7)*, ACM Press, 1993, 75–84.

[Wellner, 1993] Pierre Wellner, Interacting with paper on the DigitalDesk, *Communications of the ACM, 36(7)*, ACM Press, 1993, 87–96.

[Wellner *et al.*, 1993] Pierre Wellner, Wendy Mackay and Rich Gold. Computer-augmented environments: back to the real world. *Communications of the ACM, 36(7)*, ACM Press, 1993, 24–26.

Appendix II of this publication is not included in the PDF version.
Please order the printed version to get the complete publication
(http://www.vtt.fi/publications/index.jsp)

A user interaction paradigm for physical browsing and near-object control based on tags

A user interaction paradigm for physical browsing and near-object control based on tags

Pasi Välkkynen, Ilkka Korhonen, Johan Plomp*, Timo Tuomisto,
Luc Cluitmans, Heikki Ailisto*, and Heikki Seppä[?]

VTT Information Technology	[?] VTT Information Technology	*) VTT Electronics
P.O.Box 1206	P.O.Box 1200	P.O.Box 1100
FIN-33101Tampere, Finland	FIN-02044 VTT, Finland	FIN-90571 Oulu, Finland
+358 3 316 3111	+358 9 4561	+358 8 551 2111

Firstname.Surname@vtt.fi

ABSTRACT

In this paper, we present a user interaction paradigm for physical browsing and universal remote control. The paradigm is based on three simple actions for selecting objects: pointing, scanning and touching. We also analyse how RFID technology can be used to implement this paradigm. In a few scenarios, we show the potential of augmenting physical objects and environment with digital information.

Categories and Subject Descriptors

H.5.2. [**Information Systems**]: User Interfaces – *Interaction styles.*

General Terms

Human Factors.

Keywords

Physical browsing, pointing, tangible user interface, mobile phone, PDA, natural UI.

1. INTRODUCTION

Want et al. summarise the goal of *augmented reality* and *physically-based user interfaces*:

"The goal of these projects is to seamlessly blend the affordances and strengths of physically manipulatable objects with virtual environments or artifacts, thereby leveraging the particular strengths of each." [5]

Physical browsing can be defined as getting hyperlink information from physical objects. This can happen if the object has a way to communicate a URL to a user, which requests it. This URL can be transmitted for example with *an information tag* and it can

be read with a mobile device like a cell phone. We define an information tag (hereafter: a tag) as a small and inexpensive unique identifier, which 1) is attached to a physical object but has limited or no interaction with the object itself, 2) contains some information, which is typically related to the object, and 3) can be read from near vicinity.

A tag may be for example a barcode, RFID (radio frequency identifier) tag or an IR (infrared) beacon. Based on the tag information, the user can then for example load the page corresponding to the URL to his device and get electronic information from a physical object. This is a powerful paradigm, which adds the power of World Wide Web to the interaction with physical objects – information signs, consumer goods, etc.

Another aspect of physically based user interfaces is controlling or interacting with physical artefacts using a user interaction device such as a PDA. An example of this is using a PDA as a user interface to a household appliance. This approach can be seen as a *universal remote control*. In this scenario, a universal remote control is a device, which may control or interact with all kinds of objects by using suitable communication mechanisms. A major challenge in this paradigm is the establishment of the communication between the object and the UI device.

In the world of millions of objects to be augmented with digital presence, tags represent a key enabling technology for physically based user-interfaces. Traditionally, RFID tags have been used to track objects and cargo in industry and commerce. In research projects they have also been used for physical browsing and providing services related to for example conference rooms [5]. RFID tag readers are not yet very common in consumer products but as the tags become more widespread, PDAs and cell phones may have readers and there will be a common way to access the tags.

Previously, Want et al. [5] developed *Xerox tags*, a system, which the creators describe as "bridging physical and virtual worlds". The system combines RFID tags and readers, RF networking, infrared beacons and portable computing. They have created several example applications to demonstrate the possibilities of the system. In the *Cooltown* project [3], a method called eSquirt was developed. It allows the users to collect links (URLs) from infrared beacons attached to physical objects like walls, printers, radios, pictures and others. Cooltown's user interaction theme is based on adding hyperlinks on physical locations. In addition, *barcodes* can be used to transfer information between physical objects and mobile devices. The user reads the barcodes with a

wireless reader and the code is sent to a server. The server then transmits the information about tagged object to the user's cell phone, email, or some other information application or device.

Bowman and Hodges [1] have studied similar interactions in virtual environments whereas for example Mazalek et al. [2] have created tangible interfaces. Our paradigm lies somewhere between these two approaches, combining physical and virtual.

In this paper we represent and analyse a paradigm for physical user interaction based on using tags for augmenting physical objects with digital presence. Especially, we will present three paradigms for choosing the object of interest. We also discuss RFID tags as one possibility for implementing this paradigm.

2. INTERACTION METHODS

There are two approaches to using tags in physically based user interfaces: information related approach and control related approach. Essential for both uses is the requirement for choosing the object (tag) of interest. In our concept, there are three methods for choosing tags with readers: 1) scanning, 2) pointing and 3) touching. We suggest that for any tagging technology these paradigms should be supported to provide optimal support for natural interaction with physical objects.

2.1 ScanMe

Scanning is one way to choose the tag of interest. When a user enters an environment, he can use his reader to scan the environment for tags. The services provided by the tags will then be presented on the user's UI device. Thus the presence of the tags is communicated to the user and he can then choose the tag (object) of interest by using his UI device. Effectively, this means choosing a physical object in the digital world. This method can be called ScanMe paradigm (see Figure 1).

Figure 1: ScanMe

Technically ScanMe is supported by methods supporting omnidirectional or at least wide search beam communications, which is true especially for RF based methods. In ScanMe, all tags within reading range would respond to the scan, even if they were behind small objects like packaging[1]. A major issue with ScanMe is,

---

[1] Potentially, with some technologies and in the presence of a multitude of tags, there may be occasions that not all the tags successfully reply to the scan, e.g. due to communication channel overload. This would represent a problem to the UI paradigm unless there is some way of warning the UI device of

however, the universal naming problem — association between virtual and physical objects. The tags must be named somehow so that the user can understand what physical object is associated with the information on the menu.

2.2 PointMe

If the tag is visible, pointing is a natural way to access it. In PointMe paradigm, the user can point and hence choose a tag with a UI device, which has an optical beam, e.g. infra red or laser, for pointing (see Figure 2). Pointing requires direct line of sight to the tag, but it works through transparent surfaces. Like in scanning, the tag can be accessed within the range of the reader. The PointMe paradigm may be typically implemented with IR alone, or by combinations of IR, laser beam, and RF technologies. In the latter case, the optical mechanism is used for choosing the tag while the RF communication is used from tag to UI device communication.

PointMe tags can be accessed directly by pointing and selecting. Depending on the width of the beam there is also a selection problem if there are more than one tag in the place the user points at; in this case, a scanning-like menu of the tags could be presented. In any case, there is an application dependent need for a compromise between the beam width (larger beam leading to more inaccurate selection) and the usability issues (requirement for very exact pointing may lower the usability). Typically, in the PointMe paradigm the tag of interest is chosen without ambiguity and hence the related service may be launched immediately to the UI device if required. For example if the tag responds by sending a URL pointing to product information, it could be loaded into the browser of the device immediately. In more complex situations, a user interface to the tag's services could be presented.

Figure 2: PointMe

2.3 TouchMe

In TouchMe paradigm, the tag (object) of interest is chosen by (virtually) touching it with a UI device. Like pointing, touching requires that the user identify the location of the tag. However, the tag itself does not necessarily have to be visible. RFID tags may be made into TouchMe tags by limiting the reading range. This

---

the unread tags, in which case the scan could be repeated until all tags are successfully read.

can be done either by limiting the power used or by tag antenna design.

Touching is an unambiguous way to select the right tag and object. It eliminates the possibility of multiple tags responding, but the touching range limits its use. Typically, it is the most powerful paradigm in the case where a multitude of objects is close to each other, e.g. in a supermarket for downloading product information.

2.4 Universal remote control concept

The ScanMe, PointMe and TouchMe paradigms may easily be applied in the concept of physical browsing, i.e. in information related applications. However, tags and the above UI paradigms are also powerful in the concept of a *universal remote control.*

In this scenario a generic remote control is a device, which may dynamically control or interact with previously unknown objects by using suitable communication mechanisms. The basic challenges for such universal remote control are:

1. *Discovery*: how to choose the object of interest (in the physical space) by using the UI device (which is functional in the virtual space), or how to provide mapping between the physical and virtual space objects.

2. *Connectivity*: how to establish the communication channel between the object and the UI device in case the communication protocol is not know a priori, or if many communication mechanisms are supported (e.g. IrDA, Bluetooth).

3. *Communication protocol*: how to make the UI device and the object to communicate with the same vocabulary.

4. *User interface*: how to present the information to and allow control by the user on the UI device in an intuitive way.

We suggest that tags can be used as a simple mechanism to address these challenges. A tag attached to the device can hold or provide a pointer to the necessary communication parameters to be used in the control, such as communication mechanism, address, protocol and its parameters. If the tag contains a pointer to these parameters (for example in the Internet), it is possible to take into account the UI device characteristics and to download a proper UI to the device. The usage is as follows:

1. Our UI device (e.g. a PDA) includes a tag reader. In addition, it has some other communication mechanisms.

2. When the user chooses the object of interest, he scans the tag with his UI device by using ScanMe, PointMe or TouchMe paradigm. The most essential feature to the user in this procedure is that the selection is simplified as much as possible and the selection is done primarily in the physical space.

3. The tag replies to the tag reader with information about the necessary communication parameters for further control or communication needs. These may include actual communication parameters, or a URL to download these parameters and/or the device UI.

4. The UI device interprets the communications parameters, downloads (if needed) the drivers and UIs, and starts the communication with the object by using the defined method.

The main advantage from the users perspective is that the only action required from the user is to choose the object in the step 2 – all the rest may be implemented to happen automatically. There are two main advantages from the technological perspective. The first is a simple and standard[2] mechanism for device discovery supporting custom methods for communication. The second advantage is flexibility for supporting multiple devices, languages, etc. (especially in case the returned parameter is the URL of the method).

3. IMPLEMENTATION OF TAGS

The primary feature of tags is their extreme locality: they are only accessible within near vicinity, and hence they are closely related to a certain place or object. Indoor positioning and user identification can be used in similar manner as we suggest tags to be used. However, tags have some advantages over other technologies that can be used to identify a user and her indoor positioning. Some advantages of tags are their efficiency, simplicity and low cost both in computing power and monetary terms.

The most important tagging technologies currently are RFID tags and optically readable tags (barcodes or other kinds of glyphs). Both kinds of tags can be used to easily augment physical objects and the environment on a small scale. The RFID technology is becoming a challenger for barcodes in many applications, and its features allow its usage beyond possibilities of the barcodes.

RFID tags are typically passive components; i.e. they do not have their own power source; they get all the power they need from the device that is reading them. At present, the information content of a tag is typically static but the technology allows dynamic updates to the contents, e.g. updating information or adding some sensor readings from attached sensors. It naturally supports ScanMe and TouchMe concepts (the latter is achieved either by decreasing the reading power to the minimum or by modifying the antenna of the tag to be less sensitive). Support for tag selection by optical methods allowing the PointMe paradigm is being researched.

The central features of RFID tags may be summarised as follows:

1. *Visibility*. RFID tags don't need to be visible so they may be attached below the surface of the object. However, they are not readable through thick materials or metal.

2. *Range*. The maximum range of RFID tags is about four meters with 500 mW reading power [7]. It is possible to use tags that respond to touching or RF requests at very short distances. This kind of tag can be used as a TouchMe tag.

3. *Data storage capacity*. RFID tags usually have greater data storage capacity than barcodes or glyphs. The capacity may beat the range of a few kilobits [5].

4. *Sensors*. RFID tags can be connected to sensors. These sensors can be used as a condition for triggering the tag, or for reading and transmitting sensor data.

5. *Antenna*. The antenna is by far the largest element of the RFID tag, typically about one square inch. It can be made flexible and it may be attached to almost any surfaces.

6. *Price*. The prices of tags are in the order of tens of cents. In large mass production the price may be cut to a few cents.

Different RFID tags respond to different triggers. Still, their basic technology can be the same. This is a major advantage while keeping the price of the tags and their readers low.

---

2 Here it is assumed that an industry standard for a suitable tagging technology becomes accepted and agreed.

4. SCENARIOS

The scenarios in this chapter provide use cases to illustrate the use of tags for physical user interfaces and to emphasise the need for different object selection paradigms.

4.1 Physical browsing

The user notices an interesting advertisement of a new movie (see Figure 2). She points her PDA at the advertisement and presses a button. The tag responds with an URL to a web page of the movie. The PDA immediately launches a web browser and loads and displays the page. The page contains links to the movie's web page, to a local theatre and to a sample video clip. The advertisement could also have direct physical links to aforementioned things. For example, it could have a tag, which would respond directly with the URL of the video clip, whereas the tag at the movie's name would open its web page. Physical objects could act this way like user interfaces to different kinds of information.

4.2 Shopping

The user goes to a shop in which the items are augmented with RFID tags. She sees a new chocolate brand but the trade description of the chocolate bar is not in any language she knows. However, she is very allergic to nuts and must know whether the product contains nuts. So, she touches the TouchMe tag in the chocolate bar with her PDA and gets a link to the page in which all ingredients are described. This page is provided by the shop chain, but it could also be provided by the manufacturer.

4.3 Universal remote control

The user walks into a room and wants to turn on some of the lamps of the room. He notices standard RFID stickers attached to the lamps, points the first lamp with his phone and presses a button. The tag attached to the lamp transmits an URL to the controlling method; i.e. the tag itself does not control anything. As toggle between on/off are the only options for controlling the lamp no specific UI display on the phone is needed.

To identify what controllable devices there are in the room, the user first uses his mobile phone's scan function. The RF reader of the phone sends a scan request to all tags in vicinity. The ScanMe tags respond, in this case with an URL, which is a link to their control and user interface. The mobile device constructs a menu of these responses and displays it to the user. The user then selects the desired item from the menu and his cell phone loads the user interface for that device. It should be noted that the user should not get a list of URLs for choosing. Instead, the mobile device should use these URLs to get a description of the item (i.e. a "link text"). This description would be displayed in the menu and with it a new URL, which points to the user interface of the device, for example the lighting of the room.

5. DISCUSSION

Digital augmentation of everyday objects represents a new powerful paradigm. However, there are some central usability issues involved in making digital augmentation natural. In this paper we have discussed use of tags in physical user interfaces and presented three paradigms for choosing the object. Still, some generic design issues should be kept in mind. First, the users should be able to find out if there are tags in their environment or recognise tagged objects from those not tagged. The users should also understand what the tag would do if it were addressed. This is not always clear from the tag's context. These are the basic issues of *visibility, affordances and mappings*. Visibility means that a user can see what can be done with an object. The term affordances refers to the perceived and actual properties of an object, primarily those fundamental properties that determine how the object could possibly be used. Mapping refers to mapping between control and action, i.e. relationship between doing something and getting a result from it. [4] The question with physical browsing is how do we communicate these issues to the user. Clearly, some standardisation for example in representing different kind of tags would help to solve these issues.

Currently RFID tag readers are not available as embedded in the mobile gadgets. However, especially when the RFID tags extend their range into higher radio frequencies (especially to 2.4GHz) it becomes feasible to integrate the reader with the handsets. This is required for the scenarios presented above to become reality in large term. However, despite the great number of mobile handsets sold so far, the number of potential objects to be tagged and hence augmented outnumbers them by far. Hence, it is especially the price of the tags and only secondarily the price of the reader which will decide which tagging technology is the winning technology in large-scale applications.

To conclude, we have presented a tag-based user interaction paradigm for physical browsing and near-object control. We suggest that a concept of physical user interaction should optimally support object selection by scanning, pointing and touching to fully utilise richness of natural interaction. Finally, we believe that new RFID technology developments are making it a potent technology for implementing physical browsing and digital augmentation.

6. REFERENCES

[1] Bowman, D. & Hodges, L. F. *An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments*. Proceedings of 1997 Symp. on Interactive 3D Graphics, Providence, RI, 1997. pp 35–38.

[2] Mazalek, A., Davenport, G., Ishii, H. *Tangible Viewpoints: a Physical Approach to Multimedia Stories.* Proceedings of the Tenth ACM International Conference on Multimedia, Juan-les-Pins, France, 2002. pp. 153-160.

[3] Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B., Spasojevic, M. *People, Places, Things: Web Presence for the Real World*. Proceedings of third Annual Wireless and Mobile Computer Systems and Applications. Monterey CA, USA, Dec. 2000. p. 19.

[4] Norman, D. *The Psychology of Everyday Things*. 1988. Basic Books, 257 p.

[5] Want, R., Fishkin, K., Gujar, A., Harrison, B.L. *Bridging Physical and Virtual Worlds with Electronic Tags*. Proceedings of 1999 Conference on Human Factors in Computing Systems. pp. 370-377.

[6] Wicol Ltd Oy. 2003. [online]. http://www.wicol.net/

[7] Extending RFID's Reach in Europe. RFID Journal March 10, 2002.http://www.rfidjournal.com/article/articleview/328/1/1/

PAPER III

Physical browsing

Physical Browsing

Pasi Välkkynen, Lauri Pohjanheimo, and Heikki Ailisto

4.1 Introduction

Physical browsing is a means of mapping digital information and physical objects of our environment. It is analogous to the World Wide Web: The user can physically select, or click, links in the nearby environment. In physical browsing, the user can access information or services about an object by physically selecting the object itself. The enabling technology for this are tags that contain the information—for example, a Web addresses—related to the objects to which they are attached. This user interaction paradigm is best introduced with a simple scenario.

> Joe has just arrived on a bus stop on his way home. He touches the bus stop sign with his mobile phone and the phone loads and displays him a Web page, which tells him the expected waiting times for the next buses so he can best decide which one to use and how long he must wait for it. While he is waiting for the next bus, he notices a poster advertising a new interesting movie. Joe points his mobile phone at a link in the poster and his mobile phone displays the Web page of the movie. He decides to go see it in the premiere and clicks another link in the poster, leading him to the ticket reservation service of a local movie theater.

Mobile phones have become ubiquitous, and they have become versatile mobile computing platforms with access to diverse wireless services, especially with their World Wide Web and messaging capabilities. However, these small, portable devices have limited input capabilities, hindering the ease and convenience of use.

Passive radio frequency identification (RFID) tags and visual tags such as barcodes are simple, economic technologies for identifying objects. These tags can store some information—for example, an identification number, or more recently, a universal resource locator (URL). Until recently these tags have required specialized readers that have usually been connected to desktop or portable PCs and used mainly in logistics. The readers are getting smaller, making it feasible to integrate them into smaller mobile terminals such as PDAs or mobile phones, and thus more closely integrating the tags with the other functionalities of these mobile devices.

As can be seen in the next section physical browsing as a concept is not new. Various ways to associate physical and digital entities have been suggested and implemented [1–5]. By combining the aforementioned two technologies, ubiquitous

mobile terminals and cheap tagging systems, it is finally possible to augment our physical environment in a grand scale. As the prices of tags are dropping and without a need for batteries or maintenance we will be able to tag practically anything we want, truly linking the physical and digital worlds. The momentum for physical browsing will thus be created from two directions: a need for easier and more convenient access to digital information, both WWW and local services, and the emergence of enabling technologies—affordable RFID (and visual) tagging technologies and their convergence with mobile phones and PDAs. The first RFID readers have already been released for mobile phones, and cameras able to read visual tags are becoming standard equipment in them.

As a user interaction paradigm, physical browsing is very much analogous to browsing the World Wide Web. A tagged environment includes links to information about the environment and objects in it. By selecting these links with a mobile terminal, the user clicks these physical links, and a terminal displays information related to the objects or activates other services related to them.

We first discuss the previous work done on the topic of combining physical and digital worlds, especially previous physical browsing research. Then we define central terms related to physical browsing and discuss physical selection, the equivalent of clicking a physical link in more detail. We then discuss how physical browsing relates to context awareness and the issues in visualizing the physical hyperlinks. We describe the most prominent technologies for implementing physical browsing, and we look at two demonstration systems we have built.

4.2   Related Work

Physical browsing is akin to Weiser's vision of ubiquitous computing [6] in the sense that in both concepts computational devices are brought to the real, physical world to augment it. Physical browsing can be seen as one user interaction concept in the wider picture of ubiquitous computing—one in which the user controls the interaction between him or her and the world with a mobile terminal. This view is slightly different from the visions of calm or disappearing computing that are connected to ubiquitous computing and included in Weiser's vision. Weiser strove for implicit interaction, whereas in physical browsing the interaction is very explicit. In practice, the most dramatic difference is in the implementation. The traditional view of ubiquitous computing emphasizes intelligence in the environment to make interaction implicit, which puts tremendous requirements on the infrastructure and the power of the appliances. Since physical browsing is based on more explicit interaction the environment can be augmented in extremely economic and simple ways—by augmenting objects with simple and cheap links to information. We see room and need for both views. After all, the important thing is to combine the strengths of the physical and digital worlds.

Coupling between digital and physical objects is thus a central concept of physical browsing. Augmented reality research has explored this area by using virtual reality techniques and devices to display digital information in the physical world, often in visual or aural form. Another approach, which is sometimes used in tangible user interfaces, is to augment the physical objects and make them act as containers

and operators of information [7]. In physical browsing, we use the latter approach. We take first a look at some research on the tangible user interfaces and how they relate to physical browsing, and then on some research projects that are close to our view of physical browsing.

4.2.1   Tangible User Interfaces

Wellner has developed DigitalDesk [8], which is an augmented desk supporting various interactions between electronic and paper documents and is among the first systems creating associations between digital information and physical objects. Ishii and Ullmer have proposed the concept of tangible bits [9], which similarly bridge the gaps between the digital world of information and the physical world of our everyday environments. One of the key concepts of tangible user interfaces is thus association of physical objects to their digital counterparts. A good example of this is mediaBlocks [10], which are small, tagged wooden blocks that can be used as containers, transports, and controls for digital media. The approach of Ishii/Ullmer could be seen as making digital information tangible whereas our approach in physical browsing is slightly different, to create a simple, intuitive mapping between the existing physical objects and the (often existing) information about them.

In their later work, Ullmer and Ishii have redefined tangible user interface to mean *no distinction between input and output* [11], which creates a clear distinction between tangible user interfaces and physical browsing. Physical browsing is more about telling the mobile terminal which physical object we are interested in, rather than directly manipulating the object in question, as is the case in truly tangible or graspable user interfaces. It is not necessarily interacting with smart objects other than the terminal (although the terminal may interact with a smart object); instead, the objects in the environment are simply augmented to support the association between physical and digital.

Fishkin has proposed taxonomy for tangible user interfaces [12] and relaxed Ullmer and Ishii's later definition. The taxonomy involves two axes, embodiment and metaphor. The first axis answers: "To what extent does the user think of the state of computation as being embodied within a particular physical housing?" The alternatives are full (output device is input device), nearby (output takes place near input object), environmental (output is around the user) and distant (output is "over there"). The second axis of the taxonomy is metaphor, which Fishkin quantifies as none, noun (analogy between physical shape and information contained), verb (analogy in the act being performed) or noun and verb. In physical browsing the embodiment can be seen in one sense as "full"; the link and the information it points to seem to be inside the physical object. The metaphor axis is noun: The object physically resembles the information it contains.

4.2.2   Physical Browsing Research

Want and colleagues [13] state that people are at their most effective when they are using familiar everyday objects and that the desktop metaphor fails to capture the ease of use and flexibility of those objects. They propose bridging the gap between digital and physical by activating the everyday objects instead of using the meta-

phor, connecting physical objects with their virtual counterparts via various types of tags. Want and coworkers call these kinds of user interfaces physically-based user interfaces and state that the purpose of these interfaces is to seamlessly blend the affordances and strengths of physically manipulatable objects with virtual environments and artifacts, thereby leveraging the particular strengths of each. They have combined everyday physical objects, RFID tags, and portable computing to create several example applications. Their user interaction method is simple; the user waves the tagged object near his orher tablet PC, which has a tag reader, and then some service is launched in the PC. Some sample applications they have built include sending email messages via augmented business cards, linking a dictionary to a translator program, and opening a digital document via a physical document.

Kindberg and colleagues [2, 3] have created Cooltown, in which people, places, and physical objects are connected to corresponding Web sites. Their vision is that both worlds would be richer if they were connected to each other and that standard Web technologies are the best way to connect them. The user interaction theme of Cooltown is based on physical hyperlinks, which the users can collect to easily access services related to people, places, and things. Cooltown utilizes infrared (IR) and RFID technologies to transmit these links to the users' mobile terminals. In Cooltown museum the visitor can gather links related to the display pieces and view WWW pages related to them. In Cooltown conference room the users can collect a link to the room at its door and access, for example, the printers and projectors of the room via the link. Kindbergand coworkers have also created eSquirt, a drag-and-drop equivalent for physical environments. With eSquirt the users can collect links and "squirt" them to other devices—for example, squirting a previously collected link to a projector to display the corresponding Web page.

Ljungstrand and colleagues have built WebStickers [7, 14], a sample system to demonstrate their token-based interaction [1]. WebStickers is not a mobile physical browsing system, but a desktop-based system to help users better manage their desktop computer bookmarks. However, WebStickers illustrates well the association between digital and physical worlds by using tags. The idea of WebStickers is to make physical objects (tokens) act as bookmarks by coupling digital information (URLs) to them. The physical objects can be organized, stored, and handled outside the desktop in many ways that are not possible within the desktop environment. WebStickers can also provide the user with cognitive cues of the content of the page in more natural ways than textual lists of bookmarks.

WebStickers uses barcode stickers to store links to WWW pages in the form of an ID number, which, after reading, is coupled to a URL in a database. The users can associate the barcode stickers with Web addresses by themselves, print their own barcodes and even associate existing barcodes on products to URLs. When the barcode is read with a reader connected to the desktop computer, the browser can open the page the URL points to or display a list of addresses if several addresses are associated with one WebSticker.

Rekimoto and Ayatsuka have introduced CyberCode [4], a two-dimensional visual tagging system as part of their augmented reality research. They see visual codes as a potential tagging technology because of the increasing availability of cheap digital cameras in mobile devices and mobile phones. They have implemented several interesting applications and interaction techniques using CyberCode, some

of which are possible only with a visual tagging technology. They have for example studied the possibility of transmitting links in a regular TV program, so that the user can point the camera of the mobile terminal towards the TV screen and access the visual tag that way. In addition to using CyberCodes as simple links to digital information, they have also implemented drag-and-drop operations similar to Cooltown's eSquirt and used visual codes for identifying the location and orientation of objects in their augmented reality environments.

Toye and colleagues [15] have developed an interaction technique for controlling and accessing site-specific services with a camera-equipped mobile phone, visual tags, and public information displays. The users can interact with local mobile services by aiming and clicking on tags using their camera phones. Toye and coworkers [15] have developed a mobile phone application, which, when activated, displays the view from the camera on the phone screen, highlighting tags it recognizes. The user can then click the link by clicking a button on the mobile phone when a tag is highlighted. This system can also utilize nearby public displays as a kind of "touch screen" in addition to the mobile phone screen—something that is only possible using visual tags.

GesturePen by Swindells and colleagues [16] is a pointing device for connecting devices. GesturePen works as a stylus for a PDA device, but it also includes an IR pointer, which can read IR tags attached to electronic devices and transmit the tag information to the PDA. The tag contains the identification of the device it is attached to so that the user can select the device for communication by pointing at it, instead of the traditional list, that mobile phones display when they scan the environment for Bluetooth devices. The motivation of Swindellsand coworkers is making the selection process easier, since it is more natural for people to think the device they want to communicate with as "that one there" instead of a menu item in a list.

In addition to the aforementioned research projects, there are also some commercial ventures utilizing physical browsing. Hypertag[1] is a commercial physical browsing system based on mobile phones and IR tags, which can send, for example, WWW addresses to the phones. Integrated RFID readers are appearing in mobile phones—for example, Nokia has released an RFID kit[2] and a Near Field Communication (NFC) shell[3] for their mobile phones. NFC [17, 18] is a technology for short-range (centimeters) wireless communication between devices. The devices are typically mobile terminals and consumer electronics devices, but also smart cards and low-frequency RF tags can be read with NFC readers. There are four applications, as defined by Philips: (1) "Touch and Go" allows simple data gathering from the environment or using the mobile terminal as a ticket or an access code; (2) "Touch and Confirm" lets the user confirm an interaction, (3) "Touch and Connect" helps in linking two NFC-enabled devices, making the device discovery and exchanging communication parameters easier, and finally (4) "Touch and Explore" allows the user or device to find out what services or functions are available in the target device. Interaction in Near Field Communication is based on virtually making the objects touch each other.

1.   http://www.hypertag.com.
2.   http://www.nokia.com/nokia/0,,55738,00.html.
3.   http://www.nokia.com/nokia/0,,66260,00.html.

All these projects or products implement some specific parts and aspects of physical browsing. In the following sections we introduce a generic definition for physical browsing, what steps physical browsing consists of, methods for selecting the link, and technologies that can be used to implement physical selection.

4.3   Physical Browsing Terms and Definitions

4.3.1   Physical Browsing

The term physical browsing has first been introduced by Kindberg [3]. He describes it as: Users obtain information pages about items that they find and scan. Holmquist and colleagues have introduced a taxonomy for physical objects that can be linked with digital information and call it token-based access to digital information [1]. The precise definition of token-based access to digital information is a system where a physical object (token) is used to access some digital information that is stored outside the object, and where the physical representation in some way reflects the nature of the digital information it is associated with. This is a basis for what we call physical browsing—accessing digital information about physical objects and from the objects themselves. We can thus define physical browsing as a system where a mobile terminal and a physical object are used to access some digital information that is stored outside the object, and where the physical representation typically[4] in some way reflects the nature of the digital information it is associated with.

4.3.2   Object

Holmquist and colleagues [1] classify generic objects that can be associated with any type of digital information as containers. A container does not have to have any physical properties representing the information it contains. If there is a physical resemblance between the object and the information it contains, it is called a token. Token is thus more closely tied to the information it represents. In Cooltown [2], the object can also be a place or a person. We take that broader view of objects, including, for example, environments and people in addition to artifacts.

4.3.3   Information Tag

We have defined information tag [19] as a small and inexpensive unique identifier, which is attached to a physical object but has limited or no interaction with the object itself; contains some information, which is typically related to the object; can be read from near vicinity.

Holmquist's and coworkers' [2] token or container is thus the physical object itself, while information tag is the technical device that augments the token or container.

---

4.   The physical object does not necessarily have to reflect the information, as Holmquist et al., [1] note in their definition of a container.

4.3.4   Link

The information tag provides the user a *link* to digital information about the object. This link may be an ID, which is mapped into some other link type (URL in the WebStickers example) a direct Web address, or a phone number as a link to a person, just to mention a few possibilities.

4.3.5   Physical Selection

The basic sequence of physical browsing can be divided into the following phases.

1. The user discovers a link in his or her environment and wants to access the information it leads to;
2. The user selects the link with his or her mobile terminal;
   The link activates an action in the mobile terminal.

Physical selection covers the second phase in physical browsing. In physical selection, the user tells the mobile terminal which link he or she wants to access—that is which tag he or she wants the terminal to read.

Our user interaction paradigm for physical browsing is based on three simple actions for selecting objects: pointing, touching, and scanning, which we also call PointMe, TouchMe, and ScanMe and which we describe in more detail in the following section. All these are initiated by the user, which corresponds to our view of ambient intelligence in which the user controls the interaction through a mobile terminal. In addition to these user-initiated selection methods, there is also NotifyMe, a selection method in which the selection is initiated by the tag or the mobile terminal.

4.3.6   Action

After the mobile terminal has read the link, some kind of action happens. Some example actions are opening a Web page, placing a phone call, or turning on the lights of the room. An important distinction is that this action is separate from the selection method. Actions can also be grouped into categories, as in the NFC user interaction paradigm [18]: for example Touch & Explore, and Touch & Connect.

We seek to create an association between physical browsing and more traditional Web browsing, so we will be using vocabulary that is more closely related to the existing Web browsing vocabulary, hence the term physical browsing itself instead of token-based access to digital information.

4.4   Physical Selection Methods

In this section, we describe the user-initiated physical selection methods, pointing, touching, and scanning. In addition to them, we discuss NotifyMe, a tag or terminal initiated selection method.

4.4.1 PointMe

Pointing is a natural way to access visible links that are farther away than in touching range. In PointMe, the user selects the tag by aligning his or her mobile device toward the intended tag. For pointing, a direct line of sight to the tag is needed, meaning that it does not work with tags embedded under the surface of the object. The tag can be accessed if it is within the range of the reader, typically at most a couple of meters for passive RFID tags, but enough to access tags within a room, for example. If the implementation technology is not directional as itself (for example, visual tags or infrared), the tag must have a sensor to detect when it is pointed. One possibility to implement this is adding a pointing beam to the mobile terminal as seen in Figure 4.1. When the sensor of the tag detects the light (visible or invisible), it knows it is being pointed and should respond.

In environments where tags are close to each other, pointing may be an ambiguous selection method. It is possible that the user hits the wrong tag or more than one tag with the pointing beam. The optimal width for the pointing beam is thus a compromise between easier aiming (wider beam is easier to point with) and probability for multiple hits (wider beam hits several tags more probably). Multiple hits may be presented in the same way as in the ScanMe method (see Section 4.4.3).

4.4.2 TouchMe

Touching is another natural way to access a visible link. In TouchMe, the tag is selected by bringing the mobile terminal close to it, virtually touching the tag. Touching is the most unambiguous physical selection method, and while it lacks in range, it is a powerful method when there are many tags close to each other making accurate pointing difficult. While pointing can be seen as selecting "that one there," touching is about "this one here."
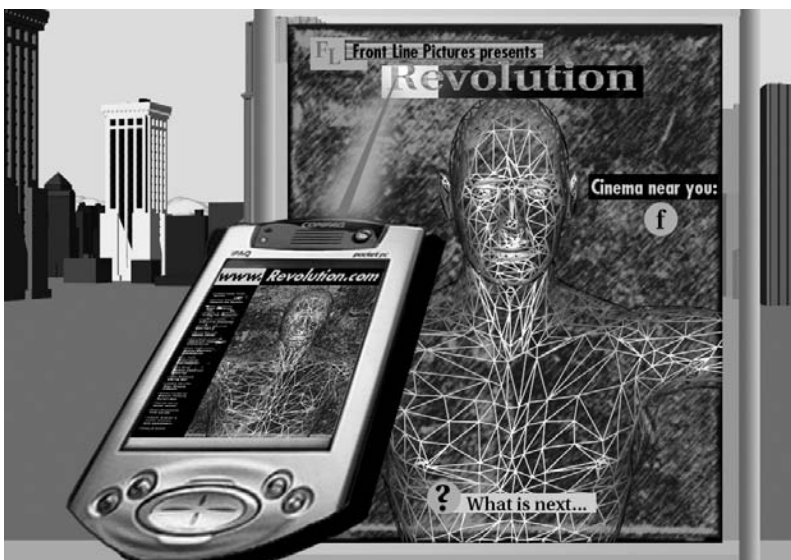

Figure 4.1   PointMe. The user points a mobile terminal at a link in a movie poster. The mobile terminal reads the tag in the poster and displays the Web page of the movie.

Most previous physical browsing systems have used touching as their selection method. This is partly due to the short range of current RFID tags and readers and visual tagging systems. It is also easiest with regards to the user interface

4.4.3   ScanMe

In ScanMe selection method the user uses a mobile terminal to read all the tags in the environment. ScanMe is at its most useful when the user either knows with which object he or she wants to interact but does not know the exact location of the link or does not know which links are available in the environment. After scanning, a result of the scan is presented in the mobile terminal as seen in Figure 4.2.

ScanMe is similar to establishing a Bluetooth connection in current mobile phones. The phone is set to search for Bluetooth devices in the environment and it displays the result of the search as a list, which allows the user to select the target device from the GUI. In an ambient intelligence setting in which dozens or hundreds of objects are augmented the list quickly becomes too long to effectively navigate, which demonstrates the need for single-object selection methods such as TouchMe and PointMe.

ScanMe presents challenges for the design of the tags and the overall infrastructure. One question is how to map the information from tags to the visual presentation in the GUI, for example, the link texts in the Figure 4.2. Should the tag dedicate a part of its memory to the description information or should the description reside in a remote server?

Another challenge in physical selection is what to do after the tag is read. Should it be displayed in the mobile terminal for confirmation before activating the link? Different actions call for different confirmation policies; for example, simple physical Web browsing becomes quickly very tedious if every selection has to be followed by a confirmation to really open the page. On the other hand, a phone call to an expensive service number by accident would not be wanted. The selection methods

Figure 4.2   ScanMe. The user scans the whole room and all the links are displayed in the GUI of the mobile terminal.

also have an effect on the need for confirmation. If the user is pointing or touching a link, the user knows he or she wants to interact with that specific link, so confirmation may not be needed. However, when he is scanning an environment, he or she may not know which services are available and only wants to list them. In that case, opening Web pages or activating other actions is not what the user had in mind.

4.4.4   NotifyMe

In addition to the user-initiated selection methods, active tags can also start the communication. We call this tag-initiated selection NotifyMe. In NotifyMe, the tag sends its contents to the terminal, possibly based in the context it actively monitors, and the terminal displays the information to the user without the user having to select the tag. Another similar case is when the reader constantly, or based on context, reads tags from its environment. This is similar to the ScanMe selection method but the selection is not initiated by the user but by the terminal device. The terminal can then, based on the preferences of the user, display services and information it thinks are interesting to the user.

It is clear that the user should have the ultimate control of which services get pushed by NotifyMe to his or her terminal. A simple scenario involves the user walking through a shopping mall, past several small stores. If every advertisement these stores send to his or her mobile phone gets through, the simple stroll through the mall becomes a rather daunting task.

Siegemund and Flörkemeier [20] divide interaction between users and smart objects to interaction initiated by user (explicit, user in control) and interaction initiated by smart objects (implicit, user not so much in control). The latter is based on context information, and they call it invisible association. They state that explicit and implicit association methods may not be adequate in their pure form in a pervasive computing setting with a massive amount of smart objects, so a hybrid form is needed. They call this hybrid form invisible preselection, in which only probable candidates for user interaction are presented to the user for explicit selection. Physical selection can be seen as a complement to their pre-selection method, as a visible pre-selection. In the next section, we investigate further the relationship of physical browsing and context from the point of how physical browsing can help in setting the context information.

4.5   Physical Browsing and Context-Awareness

Dey [21] defines context as information that can be used to characterize the situation of an entity. The entity in his definition can be the user or another person, a place, a physical object or even an application. The important thing is that the entity is considered relevant to the interaction between the user and the context aware application. Dey considers a system context-aware, if it can use the context to provide the user information or services relevant to the task the user is involved in.

Pradhan [22] has explored one aspect of context, the location of the user, as a customization parameter for Web services in the CoolTown project []. Location can be represented as a point in a coordinate system (for example, latitude and longi-

tude), or as a hierarchical location (in Pradhan's example Palo Alto is part of California, which in turn is part of the United States). Another representation for a location could be proximity to an object or place, which is often a more useful concept in determining context than the absolute location. Pradhan has defined semantic location as an orthogonal form of location, which can be represented as a universal resource identifier (URI). In CoolTown, different technologies for determining the semantic location have been explored. Of those technologies, encoding URLs corresponding to places in barcodes or CoolTown beacons and reading them can be seen as physical browsing. Pradhan illustrates this usage with an example in which the user is traveling in a bus and receives the URL of the bus to a PDA. With the URL and the coordinate location of the bus itself the user has access to customized Web services relevant to his situation in the bus.

Schmidt and Van Laerhoven [23] state that the traditional concept of explicit human-computer interaction is in contrast to the vision of invisible or disappearing computing that emphasizes implicit interaction. They propose context as a means of moving the interaction toward implicit interaction, so that the user does not need to explicitly input context information. However, perfect and always reliable context-awareness in dynamic real-world settings seems to be a very difficult, if not impossible, goal to obtain. Instead, there may exist a balance between the totally implicit interaction and traditional human-computer interaction. We propose that a simple, natural interaction such as physically pointing or touching an object—physical selection—could be an acceptable compromise between the reliability of explicit interaction and the naturalness of implicit interaction. The entities of Dey's context definition [21] can be detected by physical browsing, for example, by reading a tag attached to an artefact, the user tells the mobile terminal or application that he or she is interested of interacting with the object. The terminal or application can then be reasonably sure that the user is located near the object in question and that the object has some relevance to the user at that situation.

Physical browsing can thus be harnessed to help determine at least a part of context for context-aware applications. In the traditional context-aware computing approach, a large amount of sensor data is collected and then the context reasoned from the measurement data. In the MIMOSA project [24] we have defined several scenarios that utilize physical browsing in explicitly setting the context for a mobile terminal. For example, in one scenario, the user sets his or her mobile phone in silent mode by simply touching with the phone a sign that tells people entering the lecture room to silence their phones. In other MIMOSA scenarios, the users use many different mobile applications to interact with services in their environments. In our scenarios the user reads a "context tag" that explicitly tells the mobile terminal to start a service that is relevant to the situation, in the golf scenario example the application is a golf swing analyser application. This explicit interaction model also keeps the user in control all the time. In MIMOSA we are also using methods similar to physical browsing to detect when the context changes and either to start an application or change the state of a running application. For example when a golf player picks up a new golf club, the tag reader detects the change and starts measuring the club movements in order to analyse the swing.

It should be noted, however, that physical browsing is only one aspect to determining context information. While it has many benefits such as explicitness, reli-

ability and ease of implementation, there are also disadvantages. The users must act to give the context information to the system, which may become too laborious if it has to be done too frequently. In addition, physical browsing can by no means give all the information needed for context awareness in all situations, but it could be a useful method to aid in determining context in some situations.

4.6   Visualizing Physical Hyperlinks

One challenge in physical browsing, especially in the physical selection phase is the visualization of the hyperlinks in a physical environment. Some tagging systems are inherently visual, for example, barcodes and matrix codes must be visible for the reader device to be of any use, and their visual appearance and data content are inseparable. RFID tags on the other hand are not so visible and they can even be inserted inside objects. This is sometimes an advantage since the tag will not inter-fere with the aesthetics of the object it augments, but from the user's point of view, it can present a problem.

The main challenges of visualizing physical hyperlinks are as follows.

1. How can a user know there is a link in a physical object or in an environ-ment?
2. How can a user know in which part of the object the link is if he or she knows of its existence?
3. How can a user know which action or functionality will follow from activat-ing the link?
4. How can a user know how a link can be selected if the link does not support all selection methods?

Currently, links are widely available in WWW pages. Optimally, the links are clearly marked as links and the link anchor is a word, a short phrase or image describing the link destination. WWW links are thus usually visible and include information about the action. In desktop WWW, browsing users have also learned to guess from the context what happens when they select a link; that is, from the text or images surrounding the link, or the broader context of the link, such as the whole page.

Physical hyperlinks should follow the same kinds of conventions. The link should be visible, its visual appearance should preferably follow some similar con-vention as the widely used underlining and link color serve in desktop WWW. The context of the link—the physical object it augments and the environment of the object—is a powerful cue for the user, but does not necessarily tell anything about the action that follows from the selection.

The link should thus optimally include information about the following.

1. The presence of a link;
2. The selection methods supported;
3. The action that follows selection.

This is a lot of information to be included in a small visual presentation, and the orthogonality of the second and third requirements makes the number of combinations large.

Another issue to consider is the aesthetics of the link. Not many of us would be happy to have barcodes all over our homes and even more beautifully designed tags may not gain any more popularity in certain environments. One strength of RFID tags is that they may be embedded under the surface of the object so they do not disturb the visual appearance of the object, making it impossible to detect the tag by human senses.

4.7   Implementing Physical Browsing

The three main alternatives for implementing physical browsing are visual codes, infrared communication and electro-magnetic methods. Wired communication methods are left out, since they require clearly more actions from the user than the physical selection paradigm implies.

4.7.1   Visual codes

The common barcode is by far the most widely used and best-known machine-readable visual code. It is a one-dimensional code consisting of vertical stripes and gaps, which can be read by optical laser scanners or digital cameras. Another type of visual code is a two-dimensional matrix code, typically square shaped and containing a matrix of pixels [25]. Optical character recognition (OCR) code consists of characters, which can be read by humans and machines. Special visual code, called SpotCode, is a means to implement novel applications for mobile phones [15]. For example, information or entertainment content can be downloaded by pointing the camera phone at the special circular SpotCode, which initiates a downloading application in the phone.

Visual tags are naturally suitable for unidirectional communication only, since they are usually printed on a paper or other surface and the data in them cannot be changed afterward [7]. The tags, usually printed on paper or plastic, are very thin and can thus be attached almost anywhere. The most significant differences between barcode, matrix code, and OCR are in the information density of the tag and the processing power needed to perform the image recognition. Barcodes have typically less than 20 digits or characters, while matrix tags can contain a few hundred characters. The data content of an OCR is limited by the resolution of the reading device (camera) and the available processing power needed for analyzing the code. Visual codes do not have any processing capability, and they do not contain active components, thus their lifetime is very long and they are inexpensive. The reading distance ranges from contact to around 20 centimeters with hand held readers and it could be up to several meters in the case of a digital camera, depending on the size of the code and resolution of the camera. By nature, visual codes are closer to the pointing class than touching type of selection.

Barcodes are widely used for labeling physical objects, especially in retail commerce and logistics. There are already a myriad of barcode readers, even toys, on the

market. Software for interpreting barcodes in camera pictures is also available. The presence of barcodes in virtually all commercial products makes it possible to use them as links to various information sources ranging from manufacturer's product data to Greenpeace's boycott list.

4.7.2   Electromagnetic Methods

Barcodes notwithstanding, radio frequency identifiers are the most widely employed machine-readable tags used for identifying real-world objects. RFID-based tag technology has been so far used mainly for logistics and industrial purposes. The solutions have been typically vendor specific. Recently, there has been renewed interest in the use of RFID tags, partly due to strong academic and industrial alliances, most notably Electronic Product Code[5] and NFC Forum.

   RFID systems incorporate small electronic tags that communicate with a compatible module called reader [26]. The communication may be based on a magnetic field generated by the reader (inductive coupling), or capacitive coupling which operates for very short distances. Longer operating ranges, even several meters, can be achieved by long-range RFID tags based on UHF (ultra high frequency) technologies [27]. The tags are typically passive, which means that they receive the energy needed for the operation from the electromagnetic field generated by the reader module, eliminating the need for a separate power supply. Additionally, there are active RFID tags which incorporate a separate power supply for increasing the operating range or data processing capability. RFID technology can be applied for physical selection by integrating a tag in the ambient device and a reader in the mobile device or vice versa.

   An RFID tag typically contains an antenna and a chip for storing and possibly processing data. Tags can be unidirectional or bidirectional. Unidirectional tags are typically used in public environments, where the contents of the tags can only be read. Bidirectional tags are used when the user can freely change the contents of the tag. Most widely used communication frequencies are 125 kHz and 13.56 MHz, which is favored by the NFC Forum. Reading distances range from few millimeters to several meters. Applications include logistics (palette identification), antitheft devices, access cards and tokens, smart cards, and vehicle and railway car identification.

   Certain technologies based on magnetic induction [28] and radio frequency wireless communication [29] particularly aimed for short ranges have been presented. These technologies can be applied for identification tags, especially when long reading ranges in order of several meters are needed.

4.7.3   Infrared Technologies

Infrared is widely used in local data transfer applications such as remote control of home appliances and communication between more sophisticated devices, such as laptops and mobile phones. In the latter case, the IrDA standard is widely accepted, and it has a high penetration in PC, mobile phone, and PDA environ-

---

5.   http://www.epcglobalinc.org.

ments. Due to the spatial resolution inherent to the IR technology, IR is a potential technology for implementing physical selection applications based on the pointing concept.

An IR tag capable of communicating with a compatible reader module in the mobile device would consist of a power source, an IR transceiver and a microcontroller. The size of the tag depends on the implementation and intended use, but the smallest tags could easily be attached practically anywhere. The data transfer can be unidirectional or bidirectional. The operation range can be several meters, but a free line of sight (LOS) is required between the mobile device and the ambient device. In the IrDA standard, the specified maximum data rate is 16 Mbit/s and the guaranteed operating range varies from 0.2 to 5 meters, depending on the version used. One possible problem of IrDA, concerning especially the ambient device, is its high power consumption. For reducing the mean power consumption and thus extending the lifetime of the battery, if used, the IR tags can be woken up by the signal from the reader module [30, 31]. It is also possible that the tag wakes up periodically for sending its identification signal to the mobile device in its operating range.

In general, IR technologies are very commonplace. Many home appliances can be controlled by their IR remote controller. Several mobile phones and laptops incorporate an IrDA port, and with suitable software, they can act as tag readers. Components and modules are also available from several manufacturers.

4.7.4   Comparison of the Technologies

The most potential commercial technologies for implementing physical selection are compared in Table 4.1. Bluetooth is included for reference since it is the best-known local wireless communication technology. Obviously, exact and unambiguous values are impossible to give for many characteristics and this is why qualitative descriptions are used instead of numbers. When a cell in the table has two entries, the more typical, standard or existing one is without parenthesis, and the less typical, non-standard or emerging one is in parenthesis.

When considering the existing mobile terminals, it can be concluded that Bluetooth and IrDA are widely supported both in smart phones and PDAs; camera phones support visual codes to certain extent, and RFID readers are just emerging to mobile phones. Both barcode and RFID readers are available as Bluetooth connected accessories. Barcodes are cheap and extensively used; other visual tags are also cheap but not standardized; IrDA tags are virtually nonexistent and RFID tags are still rare and somewhat costly but gaining popularity due to EPC and NFC launches.

4.8   Demonstration Applications

We have created demonstrations as proofs of concept for physical browsing, and to study the user interaction in physical browsing. In this section, we describe our RFID emulation demonstration, which implements all three selection methods, and a genuine RFID and mobile phone-based system.

Table 4.1   Comparison of Potential Commercial Technologies for Physical Selection (Bluetooth included as a reference)

	Visual code	IrDA	RFID, inductive	RFID, UHF	Bluetooth
Selection concept	PointMe (TouchMe)	PointMe	TouchMe	ScanMe (TouchMe) (PointMe)	ScanMe
Data transfer type	Unidirectional	Bidirectional	Unidirectional (bidirectional)	Unidirectional	Bidirectional
Data rate	Medium	High	Medium	Llow-medium	High
Latency	Very short	Medium	Short	Sshort	Long
Operating range	Short-long	Medium (long)	Short (medium)	Medium-long	Long
Data storage type	Fixed	Dynamic	Fixed (dynamic)	Fixed (dynamic)	Dynamic
Data storage capacity	Limited	Not limited	Limited (not limited)	Llimited (not limited)	Not limited
Data processing	None	Yes	Limited	Limited	Yes
Unit costs	Very low	Medium	Low	Low	Medium-high
Power consumption	No	Medium	No (low)	No (low)	Medium-high
Interference hazard	No	Medium	Low-medium	Medium-high	Medium-high
Support in PDAs or mobile phones	Some (camera phones)	Yes	Emerging	No (emerging)	Some

4.8.1   PointMe, TouchMe, ScanMe Demonstration

To evaluate the physical selection methods and other aspects of physical browsing, we have built a prototype system emulating UHF RFID tags that can be read from a distance. The system supports the three selection methods, PointMe, TouchMe, and ScanMe.

The physical browsing system consists of a reader and tags. Because pointable UHF RFID tags with photo sensors are still under development, we have used Soap-Box units [] to emulate them. The remote SoapBoxes acting as tags communicate with the tag reader via RF communication. They are equipped with proximity sensors based on IR reflection to detect touching, but the IR receivers within the sensors detect pointing by IR as well. The SoapBoxes also contain illumination sensors, which may optionally be used for pointing by visible light, for example, laser.

As the mobile terminal, we use an iPAQ PDA with a central SoapBox as tag reader (Figure 4.3.). The black box on the right side is the central SoapBox, which communicates with the tags and it includes an IR LED for pointing. Between the PDA and the SoapBox is the battery case with a laser pointer that aids in pointing. When the user presses the red button of the SoapBox half way in, the laser pointer is activated to show where the invisible IR beam will hit when the button is fully pressed. The width of the IR beam is in the range of about 15–30 centimeters in a typical pointing distance of one to two meters, so the user has to get the laser indicator only near the tag he or she wants to point to.

Figure 4.3   The mobile terminal for physical browsing.

Pointing and touching display the page or other service immediately after the tag contents are read, assuming only one tag was hit. Scanning displays all found tags as hyperlinks in the graphical user interface (GUI) of the PDA. The user can then continue by selecting the desired tag from the GUI with the PDA stylus.

We have augmented physical objects with tags to demonstrate and explore the possibilities of physical browsing. Some actions we have associated with the tags are browsing Web pages, downloading a movie via a tag in a poster, sending email messages and adding an entry about an event into the calendar application of the PDA. The inner workings of our system are not exactly equivalent to passive UHF RFID tags, but the user experience is the same and this system can be used to explore the user interface and applications enabled by real RFID tags.

4.8.2   TouchMe Demonstration Using a Mobile Phone

For testing the TouchMe paradigm using a mobile phone, we have built another prototype in collaboration with Nokia. The prototype is comprised of an RFID reader device attached to a Nokia 6600 mobile phone and a set of tags. The reader module replaces the back cover of the phone as seen in Figure 4.4. Tags for this prototype can accommodate 63 characters, which is enough for most of the URLs and other identifiers. Nokia 6600 uses a Symbian operating system with open programming interfaces, making it possible to explore various scenarios using the RFID reader.

Although the design of the reader is proprietary, it shares many common characteristics with NFC devices. Reading range of the reader is 1 to 2 cm, which makes it suitable for TouchMe scenarios. Latency in communication and data transfer with the tag are negligible, making the touch action easy and convenient to use. In the prototype, the reader is connected to the mobile phone via Bluetooth connection, which is activated by pressing the button, located in the RFID cover. In this phone model, Bluetooth is the only feasible way for third-party hardware to transfer data from the reader to the phone. This causes high power consumption while the

Figure 4.4   RFID reader attached to the back of a Nokia 6600 phone.

reader is on, and the button is used to turn the reader off to minimize consumption while the reader is not used. When the Bluetooth is connected, tags can be read without any button presses or actions, except the touching gesture made with the phone.

A tag is read simply by bringing the antenna of the reader close to the tag. In the prototype, the antenna is located around the lens of the phone's camera. In this case, when the tag is almost touched with the camera of the phone, the reader gives a short sound to notify that the tag is read, and the action indicated by the tag is immediately launched. In the commercial version of the reader, user's might want to be sure of the action the tag initiates and the phone should prompt the user, especially if tags offer commercial services user have to pay for, for example buying a soda from a vending machine.

Several scenarios were explored with the RFID reader device. Web page loading is perhaps the most used scenario used in the literature. We used a tag with a URL to launch the Web browser of the mobile phone. In addition, mobile phone-centric services were implemented: a tag attached to a business card or a picture launched phone call to the person in question. Text messaging service was used to get local weather information. In this case, the tag contained the message and phone number needed to get a weather forecast in a return message. A tag was also used to open other communication methods such as Bluetooth. This scenario was used to send text messages from the phone to the laptop for backup (Figure 4.5.).

Some scenarios with commercial elements were also demonstrated. These scenarios enhanced already available services by augmenting them with a tag. Background image of a mobile phone was changed by touching picture in a magazine. This scenario used already available text messaging service, where the action initiated by the tag replaced the inconvenience of writing and sending the message. In another commercial scenario, a soft drink was bought by touching the tag attached to the vending machine. The machine already had a possibility to buy a drink by calling a certain number, and the tag replaced the need for making the call manually. Both of these scenarios were easy to use and especially in the background image loading scenario, the need for a user prompt before making the purchase was evident because the service could activate accidentally if the phone were placed on top of the magazine.

Figure 4.5 Text messages are sent to a laptop for backup.

4.9 Conclusion

In this chapter, we have discussed the concept of physical browsing, a user interaction paradigm for associating physical objects and digital information related to them. As economic and practical tagging systems are emerging, and being integrated with ubiquitous mobile terminals—mobile phones and PDAs—it will be possible to create ambient intelligence settings in grander scale than ever before and truly connect the worlds that have for a long time been largely separate, the physical world of our everyday environment and the digital world of the World Wide Web and other information systems.

We have discussed the terminology for physical browsing: the user uses physical selection to select a link in a physical object, and the mobile terminal reads an information tag attached to the object and launches some action. Physical selection can be initiated by the user, which is the case in PointMe, TouchMe and ScanMe selection methods, or it can be initiated by the tag, as is the case in the NotifyMe method. We have also described implementation technologies for physical browsing: RFID, visual tagging and infrared technologies. Our two physical browsing demonstrations are described, one that implements all three user-initiated selection methods by emulating near-future RFID tags, and another genuine RFID demonstration, which utilizes the TouchMe selection method.

Physical browsing opens many research challenges, ranging from the user interaction in physical selection to how to visualize the links in the environment. In this chapter our point of view has been user interaction, but several interesting questions

can be found from other viewpoints. The contents of the tags have great impact on how and where they can be interpreted, for example if the tag contains only an identifier number it has to be resolved somewhere into a form the mobile terminal can put to use, for example a URL. If, on the other hand, the tag content is usable as is, using the tag requires only local connectivity unless it contains a link to the outside world. Optimally, standardization will provide at least some solutions to these infrastructure questions. The content of the tag also affects the way the tags can be displayed in the terminal of the user. Other important issues that will have impact on the design of physical browsing systems and infrastructures are privacy and security, which will affect the acceptability of applications utilizing tags and physical browsing.

4.10   Acknowledgments

References

[1]   Holmquist, L. E., Redström, J. and Ljungstrand, P., "Token-Based Access to Digital Information," *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing*, Karlsruhe, Germany, 1999, pp. 234–245.

[2]   Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B. and Spasojevic, M., "People, Places, Things: Web Presence for the Real World" Mobile *Networks and Applications*, Volume 7, Issue 5, October 2002, pp. 365–376.

[3]   Kindberg, T. Implementing Physical Hyperlinks Using Ubiquitous Identifier Resolution. Proceedings of the Eleventh International Conference on World Wide Web, Honolulu, Hawaii, USA. ACM Press, New York NY, USA, 2002. 191–199.

[4]   Rekimoto, J. and Ayatsuka, Y. CyberCode: Designing Augmented Reality Environments with Visual Tags. *Proceedings of DARE 2000 on Designing Augmented Reality Environments*. Elsinore, Denmark, 2000, pp. 1–10.

[5]   Want, R., Fishkin, K. P., Gujar, A. and Harrison, B. L. Bridging Physical and Virtual Worlds with Electronic Tags. *Proceedings of CHI 99*, Pittsburgh, PA, USA. 1999, pp. 370–377.

[6]   Weiser, M. The Computer for the 21st Century. 1991. *Scientific American* Vol. 265, No. 3. 1991, pp. 94–104.

[7]   Ljungstrand, P. and Holmquist, L.E. WebStickers: Using Physical Tokens to Access, Manage and Share Bookmarks to the Web. *Proceedings of DARE 2000 on Designing Augmented Reality Environments*, Elsinore, Denmark. ACM Press, New York, NY, USA, 2000. 23–31.

[8]   Wellner, P. Interacting with Paper on the DigitalDesk. *Communications of the ACM*, July 1993. ACM Press, New York NY, USA.

[9]   Ishii, H. and Ullmer, B. Tangible Bits: Toward Seamless Interfaces between People, Bits and Atoms. *Proceedings of CHI '97*, ACM, Atlanta, GA, 1997.

[10] Ullmer, B., Ishii, H. and Glas, D. mediaBlocks: Physical Containers, Transports, and Controls for Online Media. *Proceedings of SIGGRAPH' 98*. ACM Press, New York NY, USA, 1998.

[11] Ullmer, B. and Ishii, H. Emerging Frameworks for Tangible User Interfaces. *In Human-Computer Interaction in the New Millennium* (ed. J. M. Carrroll). Addison-Wesley, 2001. 579–601.

[12] Fishkin, K. A Taxonomy for and Analysis of Tangible Interfaces. *Personal and Ubiquitous Computing.* Springer-Verlag London, 2004. 347–358.

[13] Want, R., Weiser, M. and Mynatt, E. Activating Everyday Objects. *Proceedings of the 1998 DARPA/NIST Smart Spaces Workshop.* 1998. pp. 7–140 to 7–143.

[14] Ljungstrand, P. and Holmquist, L.E. WebStickers: Using Physical Objects as WWW Bookmarks. *Extended Abstracts of ACM Computer-Human Interaction* (CHI) '99, ACM Press, New York NY, USA, 1999.

[15] Toye, E., Madhavapedy, A., Sharp, R., Scott, D., Blackwell, A. and Upton, E. *Using Camera-Phones to Interact with Context-Aware Mobile Services.* Technical report, UCAM-CL-TR-609, University of Cambridge, 2004.

[16] Swindells, C., Inkpen, K. M., Dill, J. C. and Tory, Me. That one there! Pointing to Establish Device Identity. *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology.* 2002. 151–160.

[17] ECMA. Near Field Communication White Paper, ECMA/TC32-TG19/2004/1, 2004.

[18] Philips Semiconductors. Philips Near Field Communication WWW page, , February 21st, 2005.

[19] Välkkynen, P., Korhonen, I., Plomp, J., Tuomisto, T., Cluitmans, L., Ailisto, H. and Seppä, H. A user interaction paradigm for physical browsing and near-object control based on tags. *Proceedings of Physical Interaction Workshop on Real World User Interfaces*, Udine, Italy. University of Udine, HCI Lab., Department of Mathematics and Computer Science, 2003. 31–34.

[20] Siegemund, F. and Flörkemeier, C. Interaction in Pervasive Computing Settings using Bluetooth-Enabled Active Tags and Passive RFID Technology together with Mobile Phones. *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications* (PerCom'03). 2003.

[21] Dey, A. K. Understanding and Using Context. *Personal and Ubiquitous Computing* 5:4–7, 2001.

[22] Pradhan, S. Semantic Location. Springer-Verlag, *Personal Technologies* (2000) 4:213–216.

[23] Schmidt, A. and Van Laerhoven, K. How to Build Smart Appliances. *IEEE Personal Communications*, August 2001.

[24] Kaasinen, E., Rentto, K., Ikonen, V. and Välkkynen, P. MIMOSA Initial Usage Scenarios. 2004. Available at http://www.mimosa-fp6.com.

[25] Plain-Jones, C. Data Matrix Identification, *Sensor Review* 15, 1 (1995), 12–15.

[26] Finkenzeller, K. *RFID Handbook, Radio-Frequency Identification Fundamentals and Applications.* John Wiley & Son Ltd, England, 1999.

[27] Extending RFID's Reach in Europe. *RFID Journal* March 10, 2002. Available at

[28] *Near-Field Magnetic Communication Properties*, White Paper Aura Communications Inc. 2003.

[29] *nanoNET, Chirp-based Wireless Networks*, White Paper Version/Release number: 1.02 2004.

[30] Ma, H. and Paradiso, J. A. The FindIT Flashlight: Responsive Tagging Based on Optically Triggered Microprocessor Wakeup, in *UbiComp* 2002, LNCS 2498, 160–167.

[31] Strömmer, E., and Suojanen, M. Micropower IR Tag—A New Technology for Ad-Hoc Interconnections between Hand-Held Terminals and *Smart Objects, in Smart Objects Conference sOc'2003* (Grenoble, France, May 2003).

[32] Tuulari, E. and Ylisaukko-oja, A. SoapBox: A Platform for Ubiquitous Computing Research and Applications. Lecture Notes in Computer Science 2414: Pervasive Computing. Zürich, Switzerland, August 26-28, 2002.

RFID tag reader system emulator to support touching, pointing and scanning

RFID TAG READER SYSTEM EMULATOR TO SUPPORT TOUCHING, POINTING AND SCANNING

Timo Tuomisto [1]), Pasi Välkkynen [1]), Arto Ylisaukko-oja [2])

Abstract
Integrated RFID readers are appearing in mobile phones, enabling the selection of tags and communication with physical objects in a user-friendly way by touching. We have built a system that emulates UHF tags and supports other physical selection paradigms in addition to touching: pointing by IR or visual light, and scanning of nearby tags. The emulator is based on RF communication and sensing units, SoapBoxes. The emulator is used to study the feasibility and usability of different selection paradigms in the context of physical browsing, in particular when the tag contains directly a URL to a web resource.

1. Introduction

Integrated 13,56 MHz RFID readers are appearing in mobile phones. They enable user friendly interaction with the information and services associated with our physical environment. If RFID tags contain URLs, the tags with their physical environment form an analogue to a web page and its hyperlinks. Analogously to using a mouse to click hyperlinks, we may select and read the tags with our mobile phones by (virtually) touching the tags. We call this activity physical browsing.

In physical browsing, touching is not the only imaginable physical selection paradigm. By using UHF RFID tags, the reading range of passive tags can be extended to several metres, which enables scanning the environment for tags. Development is in progress to integrate sensors with UHF RFID tags (e.g. EU FP6 MIMOSA project [1]). Pointing to tags can be based on, for example, photosensitive sensors in tags. Thus it is foreseeable that tags may be selected using touching, pointing and scanning [2].

Other systems implementing similar concepts have been built, for example, by Want et al. [3] and Kindberg et al. [4]. These implementations have been very application-centric whereas we have focused on building a generic user interface for different physical selection methods.

In this paper, we introduce a tag reader system emulator, which expands the selection concept from touching to pointing and scanning of nearby tags. The prototype will be used to evaluate the different physical selection paradigms in physical browsing. Its functionality is demonstrated with a poster equipped with emulated tags.

---

[1] VTT Information Technology, P.O. Box 1206, 33101 Tampere, Finland, {Timo.Tuomisto, Pasi.Valkkynen}@vtt.fi
[2] VTT Electronics, P.O. Box 1100, 90571 Oulu, Finland, Arto.Ylisaukko-oja@vtt.fi

2. The Physical Browsing System

The physical browsing system consists of a tag reader (physical browser) and tag emulators. Both use RF communication and sensing units, SoapBoxes [5]. The functionality of remote sensor equipped RFID tags are emulated by remote SoapBoxes and the functionality of a reader by a central SoapBox and a PDA.

2.1. SoapBoxes

The SoapBox is a programmable device with an RF receiver/transmitter and wired communications. It is also equipped with a measurement board and a selection of sensors.

A typical SoapBox network consists of a central SoapBox, which is connected by a serial cable to a terminal device, such as a PC or a PDA, and one or more remote SoapBoxes that can wirelessly communicate with the central SoapBox. The communication range is about 10 metres. The maximum bit rate is 10 kbps.
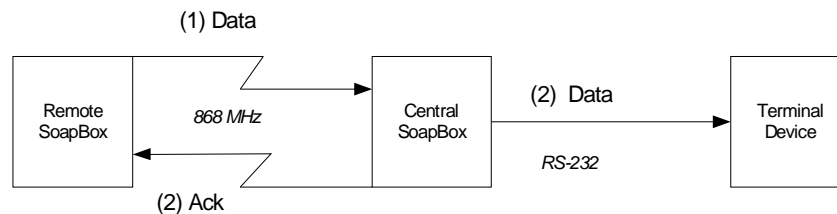
Fig. 1. Remote SoapBoxes always start the RF communication. The central SoapBox sends an acknowledgement message after reading the data.

The order of communication is shown in Fig. 1. Remote SoapBoxes can be programmed to send their data at regular intervals, or when a specific event occurs within the remote SoapBox.

The remote SoapBox sensor board is equipped with a light sensor to detect visible light, and a proximity sensor, which consists of an IR transmitter and receiver. The IR receiver in the proximity sensor is also used to detect the pointing signal emitted by the IR LED of the central SoapBox. The remote SoapBox is programmed to regularly wake up from a sleep mode, measure proximity, and detect any IR pointing signal.

2.2. The Physical Browser

The physical browser, i.e. the tag reader system, consists of a PDA (iPAQ 5400/5500 series) with a WLAN card and a central SoapBox with a laser beam unit. The central SoapBox is connected to the PDA by an RS-232 serial cable (Fig. 2).

The user interface is managed with a web browser (Internet Explorer), which is able to launch applications associated with different resource types, depending on the MIME type of the message, for example, displaying video in a multimedia player.

Fig. 2. The physical browser components: iPAQ (left), central SoapBox (right) and a laser beam unit (middle). The SoapBox has a pointing button (shown in the middle of the SoapBox) to activate the laser unit and the IR transmitter (visible in front of the SoapBox)

The PDA has a lightweight personal HTTP server [6], which supports servlets. The HomePageServlet is used to construct and provide a dynamic home page of available links when scanning, or when multiple tags are selected. The TagReader engine communicates via the SoapBox driver unit with the Java serial port driver for the PocketPC. If a unique tag (containing the URL) is selected, the TagReader engine invokes the default browser of iPAQ with the specific web resource (the URL) as an argument. . When presenting multiple links, the engine launches the web browser with the URL of the local HomePageServlet as an argument. The user then selects the proper link from the PDA GUI, or tries to select a tag again by touching or pointing.

3. The implementation of physical selection paradigm

The physical selection paradigms – pointing, touching and scanning – are implemented with remote and central SoapBoxes as follows.

For touching, the proximity sensor signal level is used to detect the proximity of objects. Whenever the measured reflected beam exceeds a certain threshold, the remote SoapBox data message is transmitted with low RF power to the central SoapBox. The range of low power transmission is about 10 cm. A flag indicating proximity is encoded in the data message. By using low transmission power, artefact detection caused by proximity of objects other than the reader can be eliminated.

Pointing with IR is initiated by pressing the pointing button (see Fig 2): initially the laser pointer is activated, and serves as a visual aid only. When the pointing button is released, the IR LED pointing of the central SoapBox is activated. The IR signal is detected by the remote SoapBox, which then transmits the data message with normal RF power. A flag indicating pointing is contained in the data message. Different IR LEDs with different beam half angles ranging from +/-4 degrees to +/-12 degrees can be used. Extra nozzles attached to the IR LED of the central SoapBox may further be

used to reduce the pointing angle. Optionally, also pointing by laser and detection by illumination sensor can be used.

To mimic scanning, remote SoapBoxes work as beacons. They send regularly (programmable, typically every few seconds) their data messages to the central SoapBox. The TagReader engine keeps track of available tags within the last 10 seconds. Scanning – presenting the dynamic home page with available tags – can be launched with one of the hardware buttons of the iPAQ.

4. Discussion

The prototype presented is not truly a passive RFID tag system in the sense that reading is not initiated by the reader device RF signal. Also a proximity sensor signal is used in addition to low power transmission to enable the detection of touching. In real RFID tags the proximity signal may be omitted altogether. However, for the purpose of studying the user interface issues of physical browsing, the user experience should remain intact. In usability studies we will test for example preferences between the different physical selection paradigms, or whether a user wants to use a button for touching, and resolution preferences when pointing with different IR beam spatial angles and detection ranges.

5. References

[1]   MIMOSA, http://www.mimosa-fp6.com/

[2]   VÄLKKYNEN, P., KORHONEN, I., PLOMP, J., TUOMISTO, T., CLUITMANS, L., AILISTO, H. and SEPPÄ, H. A user interaction paradigm for physical browsing and near-object control based on tags. Proceedings of Physical Interaction Workshop on Real-world User Interfaces. 2003.

[3]   WANT, R., FISHKIN, K. P., GUJAR, A. and HARRISON, B. L. Bridging Physical and Virtual Worlds with Electronic Tags. Proceedings of CHI 99, Pittsburgh, PA, USA. 370-377.

[4]   KINDBERG, T., BARTON, J., MORGAN, J., BECKER, G., CASWELL, D., DEBATY, P., GOPAL, G., FRID, M., KRISHNAN, V., MORRIS, H., SCHETTINO, J., SERRA, B. and SPASOJEVIVIC, M. People, Places, Things: Web Presence for the Real World. Mobile Networks and Applications, Volume 7, Issue 5 (October 2002). 365-376. ISSN: 1383-469X.

[5]   TUULARI, E. and YLISAUKKO-OJA, A. SoapBox: A Platform for Ubiquitous Computing Research and Applications. Lecture Notes in Computer Science 2414: Pervasive Computing. Zürich, CH, August 26-28, 2002. Mattern, F. Naghshineh,M. (eds.). Springer (2002). 125 - 138.

[6]   Acme Java HTTP server, http://www.acme.com/java/software/Acme.Serve.Serve.html

PAPER V

Evaluating touching and pointing with a mobile terminal for physical browsing

In: Proceedings of NordiCHI 2006. 4th Nordic
Conference on Human-computer Interaction.
Changing Roles. Oslo, Norway, October 14–18, 2006.
ACM Press 2006. Pp. 28–37.
Copyright © 2006 by the Association for Computing
Machinery, Inc.
Reprinted with permission from the publisher.

Evaluating Touching and Pointing with a Mobile Terminal for Physical Browsing

Pasi Välkkynen
VTT Technical Research Centre
of Finland
P.O.Box 1300, 33101 Tampere,
Finland
pasi.valkkynen@vtt.fi

Marketta Niemelä
VTT Technical Research Centre
of Finland
P.O.Box 1300, 33101 Tampere,
Finland
marketta.niemela@vtt.fi

Timo Tuomisto
VTT Technical Research Centre
of Finland
P.O.Box 1300, 33101 Tampere,
Finland
timo.tuomisto@vtt.fi

ABSTRACT
Physical browsing is a user interaction paradigm in which the user interacts with physical objects by using a mobile terminal to select the object for some action. The objects contain links to digital services and information related to the objects. The links are implemented with tags that are readable by the mobile terminal. We have built a system that supports selecting objects for interaction by touching and pointing at them. Our physical browsing system emulates passive sensor-equipped long-range RFID tags and a mobile terminal equipped with an RFID reader. We have compared different system configurations for touching and pointing. Additionally, we have evaluated other parameters of physical selection, such as conditions for choice of selection method. In our evaluation of the system, we found touching and pointing to be useful and complementary methods for selecting an object for interaction.

Author Keywords
Physical selection, mobile phone, PDA, laser pointer, RFID, evaluation

ACM Classification Keywords
H5.2. User Interfaces: Input devices and strategies.

INTRODUCTION
Ambient intelligence (AmI) refers to invisible computing everywhere in the environment, providing versatile information and services to users whenever needed. ISTAG [3] states that AmI integrates three key technologies: ubiquitous computing, ubiquitous communication, and intelligent user-friendly interfaces. In AmI visions and scenarios, the interaction between the user and the AmI system is described as natural and intuitive, including gestures, voice, expressions, and large visual screens. One approach to more natural interaction is associating digital AmI services and information directly to physical objects and surroundings. We call this interaction paradigm, in which the user can access digital information and services by physically selecting common objects or surroundings, *physical browsing*. It should be noted that when we speak of physical browsing, we refer to a mobile terminal based approach. The information and services are presented in the mobile terminal, and the terminal is used as the mediator between the user and the environment.

An environment augmented for physical browsing can be seen as analogous to a WWW page, which the user is able to browse by physical means. The environment contains links to different services and the user can 'click' these links with a mobile terminal in the same way desktop WWW links can be selected with a mouse. The links are implemented with tags that can be read using the mobile terminal. This first step in physical browsing is *physical selection*, a method for the user to tell the mobile terminal which tag it should read – that is, which physical object the user wants to access.

A promising enabling technology for physical browsing are radio frequency identification (RFID) tags. In addition to the identifier number, a tag can contain up to a few kilobytes of information, such as a web address. A passive RFID tag is read with a reader device, which also powers the tag so that the tag does not need a power supply of its own. These passive tags are small, simple to produce and cheap – they can be spread everywhere and thus make an attractive solution for enabling physical browsing. The reader can be integrated into a mobile terminal, for instance a mobile phone, which has several facilities to become a common means of interaction with ambient intelligence. In addition, the wireless local networking capabilities of the mobile phone allow for interaction with nearby devices. Due to these reasons, the mobile phone may well be the first realistic platform for everyday ambient intelligence applications.

We have earlier differentiated between three physical selection methods for a mobile terminal: pointing, touching and scanning [24]. *Touching* means selecting the link by

Figure 1. Pointing at a tag at a movie poster with a laser pointer equipped PDA.

bringing the terminal very close to the link. *Pointing* is a directional long-range selection method, analogous to a TV remote control. In *scanning,* all tags within the reading range are read and presented to the user as a list from which the user can continue the selection using the graphical user interface of the terminal. In this study, our focus is in the touching and pointing selection methods.

Wellner [26], and Ishii and Ullmer [6] have proposed coupling physical objects with digital information that pertains to them. Their intention has been to bridge the gap between the two worlds that had remained separate. Want *et al*. [25] call these kinds of interfaces physically-based interfaces. Mobile terminals can be used as a mediator between the user and the pervasive services in the environment. Ballaghas *et al*. [1] have analysed the design space of mobile phones as input devices for ubiquitous computing. Physical selection corresponds in their analysis to the "selection" interaction task. Ballaghas *et al*. have defined "direct pick of tagged objects" interaction technique and physical selection by touching and pointing correspond to it. Selection by scanning is not directly found in their analysis. Token-based interaction by Holmquist *et al*. [5] is "a process of accessing virtual data through a physical object". In a mobile terminal based physical selection, the tagged physical objects correspond to *tokens*, objects that physically resemble the information they represent in some way.

We have carried out a small-scale user study to evaluate different physical selection methods for physical browsing, in particular, the touching and pointing methods. Scanning has been available in commercial systems for some time and is already a fairly well understood selection method. It is closer to the traditional menu-based methods than touching and pointing that are based on physical gestures. Thus we concentrated to evaluate the two other methods; touching and pointing. We have implemented a prototype to emulate the two physical selection methods to study users' preferences. Before proceeding to the system description and user study, we describe how the selection methods can be implemented with passive RFID technology and review

earlier research work on physical browsing and selection methods.

PHYSICAL SELECTION WITH PASSIVE RFID TAGS

Integrated RFID readers have started appearing for commercial mobile terminals. So far, these readers can read tags from a few centimetres range. Research and development is underway for ultra high frequency (UHF) tags that could be read from several metres range and still have a reader small enough to be able to integrate into a mobile terminal[1]. Another technology that is being developed is a sensor interface for these long-range tags. These two new technologies allow the physical selection methods for passive RFID tags to be extended from current touching to pointing (and scanning) as well. We see this development as an important step in the usability of physical browsing, because with the new support for long-range selection methods, interaction with passive RFID tags can be made similar to other tagging systems. Examples of current technologies supporting pointing include IR nodes, sensor-equipped active RFID tags and Bluetooth nodes. The user should not have to worry about the implementation technology of the tag; instead, he or she should be able to select the desired tag using a simple gesture and let the terminal take care of the choice of technology.

In pointing, either the tag or the reader must know towards which tag the reader is aligned to. If the technology for reading the tag is not directional as itself (such as RFID), additional measures must be used to implement directionality. One way is to add a photosensitive sensor to the tag and to add a light-based pointing system to the terminal. We have chosen this approach in our work and we use both visible laser and invisible infrared light for triggering the sensor of the tag.

Both pointing and touching are thus possible to implement using only passive RFID tags. These passive long-range sensor-equipped tags are not yet generally available outside the domain of research prototypes, but they are an emerging technology. The motivation for our work arises from the need to 1) design interaction patterns that take this technology into account, and 2) to integrate passive RFID tags into a unified user interaction paradigm alongside with the other implementation technologies. We see a need for a user interaction paradigm that is as independent of the implementation technologies as possible and takes into account the low-end economic technologies such as passive RFID.

RELATED WORK

In this section, we describe previous ambient intelligence systems that use physical selection in some way in their user interaction. We give a brief overview of the current commercial technologies and systems for physical selection, and an overview of laser pointers as interaction de-

---

[1] www.mimosa-fp6.com

vices. Last, we describe the results of previous user studies regarding physical selection.

Systems Using Physical Selection
The term "physical browsing" originated from CoolTown [7, 8] for which Kindberg *et al*. devised the concept of *web presence*. The idea of web presence is to connect people, places and things to digital information about them in the World Wide Web. Kindberg *et al*. used active infrared beacons and passive RFID tags to bring information about real world objects and locations into a PDA. They also created *eSquirt* protocol for transmitting web addresses from one device to another via an infrared link. With eSquirt, the user can for example 'squirt' a URL from her PDA to a projector to display the page on the projected screen.

In their work with physically-based user interfaces, Want *et al*. [25] have built several prototypes using RFID tags and readers with portable computers. Their prototypes include for example connecting physical books to electronic information about the books and augmenting business cards with digital information about the owner.

Riekki *et al*. [19] have built a system that allows requesting pervasive services by touching RFID tags. The RFID reader in their system is an external handheld device and connected via Bluetooth to a mobile phone.

GesturePen [20] by Swindells *et al*. is a pointing device for connecting devices. GesturePen works as a stylus for a PDA device but it also includes an IR pointer, which can read IR tags attached to electronic devices and transmit the tag information to the PDA. The motivation of Swindells *et al*. is thus making the selection process easier as it is more natural for people to think the device they want to communicate with as "that one there" instead of a menu item in a list.

Visual tags such as barcodes and matrix codes can be used as camera-readable tags. Holmquist *et* al. have built Web-Stickers, a sample system to demonstrate their token-based interaction [5]. WebStickers is not a mobile physical browsing system, but a desktop-based barcode system to help the users better manage their desktop computer bookmarks. However, it illustrates well the association between digital and physical worlds by using tags. Toye *et al*. [21] have developed an interaction technique for controlling and accessing local services with a camera-equipped mobile phone, visual tags and public information displays. AURA [2] is a PDA and barcode based system for linking digital content to physical objects. Rekimoto and Nagao [17] have implemented a computer-augmented environment in which a camera-equipped PDA reads visual codes from the user's environment. Rekimoto and Ayatsuka have introduced CyberCode [18], a visual tagging system as part of their augmented interaction research.

In addition to the aforementioned research projects, there are also some ventures utilising physical browsing. Hypertag[2] is a commercial physical browsing system based on mobile phones and IR and Bluetooth tags. RFID readers are appearing for mobile phones, for example, Nokia has released a Near Field Communication (NFC) shell[3] for their mobile phones. NFC [4] is a technology for short-range (centimetres) wireless communication between devices.

In addition to physical selection by pointing, laser pointers have been used as input devices for large projected screens. For example, Kirsten & Müller [9] have developed a system in which a common laser pointer can be tracked by a camera and used like a mouse to control the application on the projected screen. The laser pointer movements are mapped to mouse movements and turning the pointer on and off is mapped to mouse button up and down. Olsen & Nielsen [15] describe a full suite of interaction techniques with a similar approach as Kirsten & Müller's, but with more laser pointer events and special graphical user interface widgets intended for laser pointer interaction. Myers *et al*. [12] have a different approach. In what they call "semantic snarfing", a laser pointer is used to grab the pointed part of the screen into a mobile terminal. In addition to interacting with projected screens, "snarfing" can be used to grab the user interface of a remote appliance into the mobile terminal.

The systems in mentioned in this section are typically based on one implementation technology and one selection method (with the exception of CoolTown and the prototypes by Want *et al*.).

Previous User Evaluations of Physical Selection
There also exist some end-user evaluations of physical selection methods. In an evaluation of RFID tag based services by Riekki *et al*. [19], touching was the method for user interaction. Using services by touching was easy to learn, and touching was appreciated by interviewees as "clear manual interaction" and because it "gives a better feeling of control" compared to automated service activation by proximity of the user. It is not however always desirable that mere touching activates the tag. For instance, if the reader is in a bag containing tagged objects, touching should be "off" to prevent continuous activation on the tags around. In our study, we compare touching with and without pressing a button simultaneously to activate the tag to find out the preferences of the users.

Swindells and colleagues [20] evaluated their gesturePen in a situation, in which the user tried to make computing devices to communicate with each other. They discovered that connecting the devices by physically pointing at them with the gesturePen was comfortable for the users and signifi-

---

[2] http://www.hypertag.com

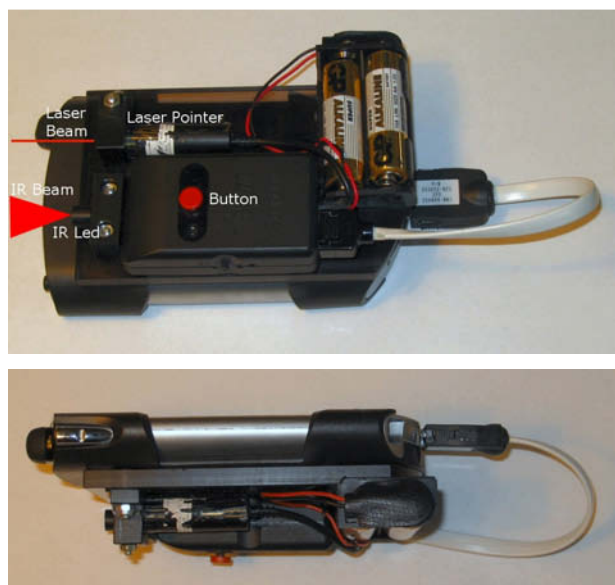[3] http://www.nokia.com/nokia/0,,66260,00.html

Figure 2. The Physical Browser. At the figure on top, the PDA is seen from below, showing the location of the IR and laser pointers and the activation button. The beam locations have been added. In the lower figure, the system is shown from the side.

cantly faster than using the conventional computer user interface.

Pohjanheimo, Ailisto and Plomp [16] have found that it is easier and significantly faster to access web pages by pointing at a physical icon representing a web address with a mobile phone, compared to a "conventional" method of activating GPRS connection and typing the web address to the phone by the user herself. A problem with the pointing method, implemented with an invisible infrared (IrDA) link, was the lack of visual feedback: the users had uncertainty about the accurate direction and time they should point towards an object to activate the service.

Weak-intensity laser beam could provide a suitable visual aid for pointing. Myers *et al*. [13] evaluated experimentally various parameters of laser pointers. They found several human limitations regarding laser pointer interaction: 1) people do not know exactly where the laser beam will point at when they turn it on, and it takes about one second to aim the beam to the target location after it is turned on, 2) the hands of the users are unsteady, which causes the pointer beam to wiggle from and around the target, and 3) when the pointer button is released, the beam often moves away from the target before going off. Although their focus was in interaction with projection screens, these limitations seem to apply to pointing at tags with a laser pointer as well.

We implemented three pointing configurations, with which we could explore the roles of visual feedback and beam width in pointing in a preliminary manner. The configurations were *wide but invisible IR beam*, *narrow but visible laser beam*, and *a combination of both invisible IR and visible laser*. To compensate the shaking mentioned by
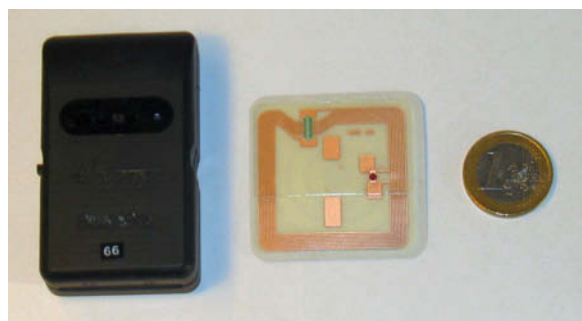

Figure 3. From left to right: SoapBox, RFID tag and one Euro coin for size comparison. The oval-shaped window in the top half of the SoapBox covers the sensors.

Myers *et al*., we decided to use a laser beam that widened slightly in proportion to the pointing distance.

To our knowledge, there are no comparative studies of different touching and pointing configurations. This study attempts to amend this situation, with limitations. The study is not designed to evaluate the user experience of interacting with tags in realistic settings, nor to analyse in which contexts users prefer different methods in an experimental manner. Instead, our intention is to reveal the most pertinent differences between different pointing and touching configurations with a mobile device. Neither do we compare touching and pointing to each other, as they are supposed to be complementary, not competing methods for interacting with the ambient intelligence environment. However, we did explore the threshold distance at which the user shifts the selection method from pointing to touching or vice versa.

SYSTEM

To study the user interaction in physical browsing, we have built *Physical Browser*, a prototype system that implements the three physical selection methods, pointing, touching and scanning [22]. The system includes a mobile terminal with tag reader and tags. The Physical Browser terminal consists of an iPAQ 5450 and an integrated tag reader (see Figure 2). The tags are implemented with remote SoapBoxes [23], and the reader with a central SoapBox. SoapBoxes communicate at 868 MHz band, have a reading range of 10 metres, and a data rate of 10 kbps. SoapBoxes are of the size of a matchbox. Data transmission is always initiated by a Remote SoapBox.

The remote SoapBox (see Figure 3) has an embedded infrared proximity transducer and a separate illumination sensor. The central SoapBox has a laser unit (630 – 680 nm, 5 mW, class IIIA), and an infrared LED, both of which may be used for pointing purposes. The pointing beams are activated with a button in the central SoapBox. The sensor part of the proximity transducer in the remote SoapBox is used to detect the proximity of another object using reflected infrared beam. It also detects infrared pointing by the central SoapBox unit. The illumination sensor is used to detect the laser beam. The remote SoapBox is programmed to
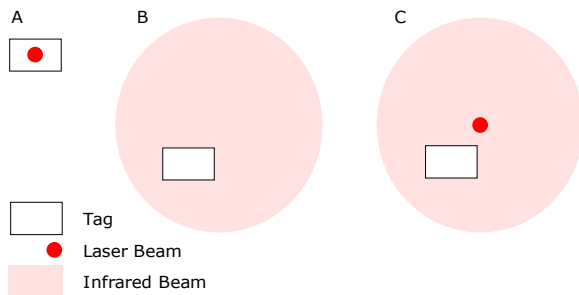
31

Figure 4. Pointing configurations. In A, the tag is triggered by a narrow laser beam only. In B, the tag is triggered by the wider infrared beam. In C, a combination of both is used and the laser works both as a trigger and as an aiming aid. This figure illustrates the relationships between the sizes of the beams and tag at approximately one-metre distance of pointing. The farther the pointer is, the larger the beam diameter is in relation to the tag size, which makes hitting the target tag easier.

regularly detect proximity, IR pointing and laser pointing by the central unit. When any of these events is recognised, the remote SoapBox sends the "tag contents" to the central unit, including also information whether the SoapBox was touched, pointed by IR or pointed by laser. However, when detecting a proximity signal, it only sends a low power signal, which has a range of 10 cm. This eliminates erroneous detection of touching with artefacts other than the terminal.

The Physical Browser has a WWW browser (Internet Explorer), which displays the tag contents. The device has a WLAN connection to the Internet and an internal servlet-based web server to produce dynamic web pages, for example the list of detected hyperlinks. The internal web server may also be used to store other internet resources.

When only one tag is selected, the hyperlink contents are fetched and shown in the PDA display. It is possible that several tags respond to touching or pointing, for example if the wide IR beam illuminates two or more tags simultaneously. In that case, a list of the responding tags is displayed in the graphical user interface of the PDA. The user can either select the desired tag with the stylus, or touch or point again. If no tags are close when the button is pressed, the user gets feedback about no tags being found. The web pages were all generated by the internal web server to eliminate delays from slow web connections.

Touching

Two configurations were implemented to select tag by touching:

- In its simplest form of an activation happens by just bringing the terminal close (~5 cm) to the tag. After the hit, the system waits for 100 ms for other touch messages.

- Activation button must be pressed simultaneously. Otherwise, the conditions are the same as in the first example.

User #	Age group	Sex	Education
1	30-44	F	Vocational
2	30-44	M	Vocational
3	18-29	M	Vocational
4	30-44	M	High School
5	18-29	M	University
6	18-29	F	High School
7	18-29	F	University
8	30-44	F	University
9	30-44	M	Comprehensive
10	30-44	F	University
11	under 18	M	Comprehensive

Table 1. Participants.

Pointing

A tag can be pointed with an IR beam, a laser beam, or both. In our study, laser beam is always activated by pressing the activation button, and turned off when releasing the activation button. IR button is activated only when releasing the activation button. The initial idea of this setup was that the laser beam is only used as visible aid for pointing, and the infrared beam triggers the pointing signal. The three different configurations of pointing are:

- User presses the activation button, and the laser beam is activated. The tag is selected whenever the laser beam hits the target (see Figure 4A). The selection process finishes within 100 ms time window from the first laser-pointing message. All laser hits during this window are accepted as valid.

- User presses the activation button. No laser beam is activated. When the activation button is released, the IR pulse is triggered, and any object hit by the IR pulse is selected (see Figure 4B). Any IR pointing messages within 100 ms from the button release are accepted.

- The combination of these two. The visible narrow laser beam is used as an aiming aid to show the user where the invisible IR beam is directed. Activation may either happen by hitting the tag with laser, or if no laser pointing hits are detected, by pointing with the IR beam (see Figure 4C).

The width of laser beam can be adjusted continuously, whereas the IR beam can be adjusted by using different types of LEDs, and by using nozzles in front of these LEDs to cut off some infrared side bands. In our study, at a distance of 4 metres, the IR beam width was 80 cm and the laser beam dimensions were the same as the sensor window

Figure 5. User pointing at a tag.

(1 cm × 2 cm). At a smaller distance, the beam widths decrease in proportion to the distance.

USER EVALUATION

The aim of this study was to evaluate two methods – touching and pointing – for selecting remote-readable passive tags. The users' experiences and preferences of the different touching and pointing configurations were explored with observing users making selection tasks, interviews, and a questionnaire.

Participants

We had eleven participants; six of them were male and five female (see Table 1) with a diversity of ages and education. All users used computers and mobile phones daily. The users participated individually.

Evaluation method

The users performed several tag selection tasks by both pointing and touching with different configurations:

Touching: A) selection by pressing a button and B) selection without button.

Pointing: A) selection with laser only, B) selection with IR only and C) selection with laser and IR combined.

We had four tags attached to different objects, which were at different distances from the participant. Two tags were attached to medicine packets that were put on a table in front of the user. Their content was product information about the medicines. The third tag was attached to a TV set two metres away from the user and contained a WWW page describing the current programmes showing on TV. The fourth tag was four metres away, attached to a door and simulated an electronic lock.

SoapBoxes have a 1 cm by 2 cm sensor window, under which the sensors are located (see Figure 3). The sensor window had to be left uncovered but otherwise the Soap-Boxes were hidden inside objects (medicine packets) or covered (TV and door). The dark sensor window thus indicated for the user the location of the embedded hyperlink.

The users asked to select the four tags in the order medicine 1, medicine 2, TV and door. The users used one selection method (for example pointing) and a specific configuration

(for example with laser) to select all the tags once and then moved on to the next configuration. To reduce the effect of learning, the configurations were tested in different orders by different users. The two touching configurations were tested consecutively varying the order in which they were tested. The three pointing configurations were also tested consecutively, varying their internal order similarly. Six users started with touching and five with pointing.

The behaviour of the users was observed and recorded on video. The users were encouraged to comment on the system during the tests and their comments were written down.

After the selection tasks, the users were asked to score the different touching and pointing configurations on six usability aspects on a questionnaire. Five of the six usability qualities have been presented by Nielsen [14].

Easy to learn refers to that the user can rapidly start using the system (selection) for her purposes. *Easy to use* is our correspondent to Nielsen's *memorability*, which refers to easiness of starting to make selections again after some period of non-use. We could not evaluate the longer-term (days, weeks) use of the system, but it is important to evaluate easiness of use after initial learning is over.

Efficient refers to the user's ability to reach a high level of productivity with the system, which in our case would mean that the user could select required tags quickly. *Reliable* refers to low *error* rate during use of the system. The user's experience of selection should match to the reality – a reliable system prevents both misses and unintended selections. *Pleasant* refers that the user is satisfied with the selection system and likes it. In addition to these five qualities, selection should be *effortless*. This is important as one of the promises of physical browsing is to make access to digital services effortless compared to current means

The six usability qualities were scored on a five-step scale from one (worst) to five (best). We emphasise that users were not asked to evaluate the usability of the system as such, as it was just a proof-of-concept having a clumsy form factor. Rather the purpose of the questionnaire was to help the participants analyse their use experience with the different of touching and pointing configurations.

After the questionnaire, the users were interviewed about the methods as well as the issues of multi-response and confirmation. We discussed about if several tags response to single selection, how should they be displayed and the interaction proceed. We also discussed the need for confirmation after selection: what kinds of actions could be automatically triggered and what actions should request for confirmation from the user.

Finally, we evaluated the threshold distance between touching and pointing, that is, at what distance do the users switch from touching to pointing. The evaluation was made when the user was sitting, standing, or walking.
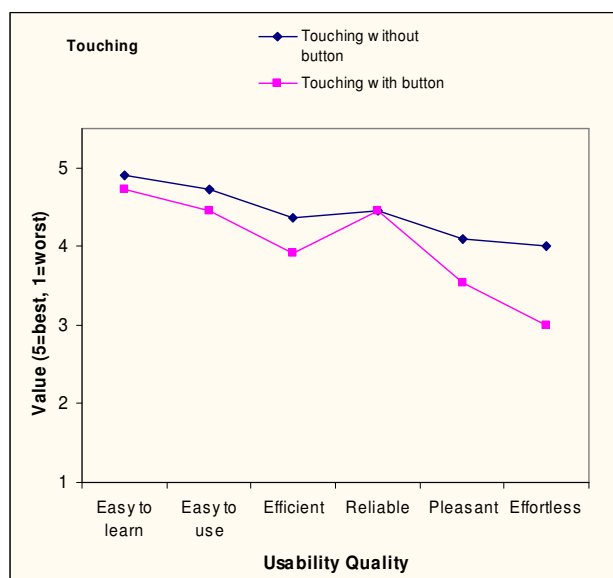
33

Figure 6. Usability qualities of the two touching configurations.

Figure 7. Usability qualities of the three pointing configurations.

For sitting and standing, the distance between the user and the tag was varied by moving the tag closer to the user or farther away from him or her, and the user was asked which selection method he or she would use to select the tag. The distance was varied until a threshold in which touching changed to pointing and vice versa was found for the user.

For evaluating the distance while walking, the user was asked to walk past a tag attached to a door in a wide corridor at different distances and speeds, and either touch or point at the tag. The distances were 1, 2, and 3 metres. The participants were allowed to walk to the tag to touch it, or curve their path to get closer the tag if they wanted. No exact measurements were made, but we observed whether pointing or touching was used, and what was the path selected by the user. We observed, for example, whether the user was willing to bend the path to use touching or to move closer to the tag for easier pointing.

Results

Touching
Figure 6 shows that the users preferred touching without button to touching with button, although the difference is small. In the interviews, nine users of eleven especially appreciated the effortlessness of touching without any extra action.

In spite of the questionnaire results showing that both touching configurations were perceived equally reliable, the users' comments during the evaluations and interviews emphasised the reliability of touching with button. Especially the lock/unlock function of the door was said to feel more reliable and safe when another gesture in addition to the terminal proximity was needed. Evaluating reliability of the touching configuration was problematic also because of the
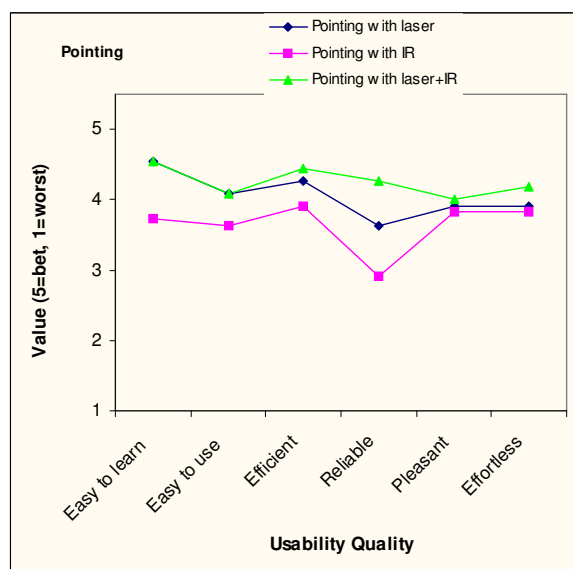
simple setting we had: the users were not confronted with situations in which automatic activation by proximity would have been a problem. For instance, in an environment with high tag density, unintended selection might easily occur if the touching was not confirmed with extra action.

Both touching with and without button were perceived equally easy to learn and use. Touching without button was perceived to be slightly more efficient, effortless and pleasant than touching with button.

Pointing
We had a reliability problem with the IR beam. Approximately 10% of intended selections were missed by the system. The reason for this was possibly the shape of the IR beam; instead of evenly distributed illumination, the projected illumination pattern consisted of circles of alternating intensity. The unreliability of the system affected the users' preferences, as the IR beam was least liked, and scored particularly low in reliability (Figure 7). However, another reason for the low preference of the IR may be the invisibility of the beam. In the selection tasks, we observed that most of the users were initially pointing too high with the IR-only configuration. One explanation for this is that the users could not see well the PDA screen when aligning the PDA towards the tag and they therefore tended to tilt it upwards to better see what was going on in the screen. The main complaint about the IR beam was indeed the lack of visual feedback.

In spite of the relative unreliability of the IR, the combination of laser and IR was rated the best on all usability qualities, and five users reported a preference for the combination in the interview. This indicates that the users intended to select with laser, and IR had a supportive function.

Four users preferred the laser only configuration, remarking that it enables selection that is more accurate when several tags are close to each other. Two users would have used IR only. They reported that they felt the presence of laser "forcing" them to aim precisely, even though they were aware that it was unnecessary due to the relatively wide IR beam.

Based on these results, it seems that a visual aid for pointing is quite critical. Users also appreciated the reliability of selection with the combination of laser and IR, in which the IR beam ensured that the selection happened with high probability even if the laser selection was missed because of an aiming error.

Distance
We also explored the threshold distance between touching and pointing. When sitting, the users chose pointing instead of touching when the distance was more than 1.1 m (median 1.0 m) on average. When standing still, the threshold distance was 1.2 m (median 1.1 m).

The users reported that while sitting they were willing to use touching only as long as they could reach the tag by just leaning forward and extending their hand. While standing, they were willing to take one step towards the tag to touch it. If the tag could not be reached with one step and by extending the arm, it would be pointed.

While walking past a tag at a parallel route, the users were not willing to bend their route towards the tag. If the tag could not be reached by extending an arm, it was always pointed. However, this distance is a complex measure that involves also the number and density of tags, the accuracy of the selection technology and non-technical issues such as the social context and need for privacy. Touching was seen more reliable than pointing, which may also affect the choice of selection method in some situations.

Overall, these results indicate that pointing is a useful selection method and the users prefer pointing to touching if the distance to the tag is more than 1 – 1.1 m.

Multi-responses and confirmation
In the case of more than one tag responding, the users were satisfied with that the responding tags were shown as a list and the user could either select the desired tag from the list, or try re-selection. If the users suspected that the re-selection would not be successful, for example due to tags being very close to each other, they would use the list. The users saw touching as an unambiguous selection method in the case where there are several tags near to each other.

The preferences for confirmation of an action following a selection were categorised according to whether the action was to access information, affect the state of the mobile device, or affect the environment. All users wanted to access information without confirmations, if the information was free of cost. Actions affecting the state of the mobile phone divided the users. Two users felt that a confirmation should always be shown for those actions, one did not see a need for confirmation in any situation, and the other users fell between these two extremes. Interestingly, two users saw a difference between pointing and touching in this regard because they saw pointing more error-prone and needing confirmation. Other issues affecting the need for confirmations were the place where the tag was read (private vs. public) and the security implications: setting a mobile phone to a silent mode would not pose much threat to the phone's security, but the users were afraid of security threats in more complex actions.

Actions affecting the environment also divided the opinions. The difference between touching and pointing was seen also here. In addition, visible feedback influenced the need for confirmation. For example, (un)locking a door is not normally visible at the door itself whereas turning lights on or off is immediately visible in the physical world, so any accidental selection would be seen and reversed quickly. Another difference is the need for security. Opening a lock is an action that has implications on security, so it was seen as needing confirmation.

DISCUSSION
Touching was perceived as an interaction method easy to learn and use, supporting the results of Riekki *et al*. The most salient difference between the two touching configurations was in effortlessness. Touching with button was perceived to require more effort than without the button. Otherwise, the configurations were valued more or less the same although the button configuration scored little lower throughout the questionnaire.

Based on these results, it seems that there is no clear reason to choose or reject either of the touching configurations so far. We suggest that as the configurations are easy to implement both, the final decision should be made by the user. Touching without button should however be a distinct mode that can be easily set on and off to prevent continuous unintentional selections, for instance, when the reader device was in a bag in contact with a number of tagged objects.

Of the three pointing configurations, pointing with only IR beam was preferred least, which probably largely reflected the real unreliability of the prototype. Only circa 90% of IR beam selections succeeded with no certain technical reason for the unsuccessful ones. The laser-only configuration required accurate aiming and thus it was evaluated more negatively than the combined IR and laser. Pointing with the combination of IR and laser was the most preferred. These results show that pointing with a mobile device should provide visual feedback to the user, but in addition to this pointing should have some insensitivity to aiming errors. These results are in line with Myers *et al*. and Pohjanheimo *et al*.

There are also other options to implement pointing with visual feedback and insensitivity to aiming errors than a system of two overlapping beams. The laser beam could be

widened to reduce the effect of shaking and to allow easier aiming. Widening the beam would however lower the intensity of it. A lower-intensity beam is not as visible to human eye in bright sunlight as a narrower higher-intensity beam, and detecting it requires sensors that are more sensitive. Another option would be to use a quickly moving micro-mirror to create a wider raster scan image in the style of projectors. While human eye is slow and sees the raster scan pattern as continuous, the sweep of the laser beam over the sensor may be too quick to trigger the sensor. Another way to implement the visual pointing aid would be to create a circle around the IR beam by using either laser scanning or optics.

In a real world situation, the intensity of laser may be harmful to eyes, but when selecting an alternative it should be taken into account that the beam should be also visible in bright sunlight. To obtain visibility on sunlight class 3A laser pointers have been used. In the classifications, it has been assumed that the laser devices are used in a well-controlled environment, for example as pointing devices by lecturers. However, the debate over safety hazards for laser pointers has increased [10, 11], as they are freely sold in internet, used in toys, and used by children. In some countries, actions have been taken to limit the availability of laser pointers for public to class 1 and class 2 laser pointers only. If laser pointers are included in mobile terminals, similar limits are likely to appear. Currently the classifications make no distinction between lasers within the 400 – 700 nm visible band. The emerging 530 nm green laser is perceived 30 times as intensive as 680 nm red laser. A low power class 2 green laser pointer might be an adequate solution for future mobile terminals.

We did not directly compare touching and pointing as they are supposed to be complementary methods for interaction. Our results show that pointing is a useful selection method and the users prefer pointing to touching if the distance to the tag is more than 1 – 1.1 m. A pointable tag has to have a range that significantly exceeds 1.1 m to be usable at all, and it is not useful to calibrate the default pointing distance (if the technology needs it) to be less than one metre.

Furthermore, social acceptance issues may be relieved with different interaction methods. Earlier studies have brought out that touching tags with a mobile device in public places could be socially embarrassing for the user [19]. Pointing is more unnoticeable and thus perhaps more acceptable in public places, until the interaction methods become common.

Our Physical Browser helped us to evaluate the concept of physical selection by touching and pointing with a mobile device for physical browsing. We made an attempt to clarify that the user should evaluate the interaction by pointing and touching only, and not the prototype system (the physical device and the graphical user interface). It is possible that irrelevant system properties affect the evaluation, like with the unreliable IR beam. In spite of this, we believe that

the results are indicative at the least. The results support and extend those of earlier studies. Necessarily more user evaluations are still required before mobile devices with pointing capabilities are found on the market.

CONCLUSION

To study physical selection, we have built Physical Browser, a system that allows the users to touch, point and scan for tags in the environment. We have evaluated the usability characteristics of touching and pointing and explored preliminarily the conditions in which the users choose between touching and pointing. We found touching and pointing useful and complementary selection methods, that is, they are used in different situations and optimally both selection methods are available to the user. Visual feedback is important for pointing and the pointing beam should preferably be wider than a laser beam to allow easier aiming. Touching without button presses or other extra actions is an effortless way to select objects, but a confirmation or using an extra action makes touching feel more secure and reliable.

REFERENCES

1. Ballaghas, R., Borchers, J., Rohs, M. and Sheridan, J.: The smart phone: a ubiquitous input device. *Pervasive Computing* 5, 1 (2006), 70-77.

2. Bernheim Brush, A.J., Combs Turner, T., Smith, M.A. and Gupta, N. Scanning objects in the wild: Assessing an object triggered information system. In *Proc. Ubicomp 2005*, Springer-Verlag (2005), 305-322.

3. Ducatel, K., Bogdanowicz, M., Scapolo, F., Leitjen, J. and Burgelman, J-C. (eds.) *Scenarios for Ambient Intelligence in 2010*. ISTAG Report, 2001.

4. ECMA. *Near Field Communication White Paper*, ECMA/TC32-TG19/2004/1, 2004.

5. Holmquist, L.E., Redström, J. and Ljungstrand, P. Token-based access to digital information. In *Proc. 1st International Symposium on handheld and ubiquitous computing*, Springer-Verlag (1999), 234-245.

6. Ishii, H. and Ullmer, B. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proc. CHI '97*, ACM Press (1997), 234-241.

7. Kindberg, T. Implementing physical hyperlinks using ubiquitous identifier resolution. In *Proc. Eleventh International Conference on World Wide Web*, ACM Press (2002), 191-199.

8. Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B. and Spasojevic, M. People, places, things: web presence for the real

36

world. *Mobile Networks and Applications*, 7, 5 (2002), 365-376.

9. Kirstein, C. and Müller, H., Interaction with a projection screen using a camera-tracked laser pointer. In *Proc. Multimedia Modeling*, IEEE Computer Society (1998), 191-192.

10. Luttrull J.K. and Hallisey J. Laser pointer-induced macular injury. *Am J Ophthalmol* 127 (1999), 95-96.

11. Marshall J. The safety of laser pointers: myths and realities. *Br J Ophthalmol* 82 (1998), 1335-1338.

12. Myers, B.A., Peck, C.H., Nichold, J., Kong, D. and Miller, R. Interacting at a distance using semantic snarfing. In *Proc. 3rd International Conference on Ubiquitous Computing*, Springer-Verlag (2001), 305-314.

13. Myers, B.A., Bhatnagar, R., Nichols, J., Peck, C.H., Kong, D., Miller, R. and Long, A.C. Interacting at a Distance: measuring the performance of laser pointers and other devices, In *Proc. CHI 2002*, ACM Press (2002), 33-40.

14. Nielsen, J. *Usability Engineering*. Academic Press, Inc., San Diego, CA, USA, 1993.

15. Olsen, D.R. Jr. and Nielsen, T. Laser pointer interaction. In *Proc. CHI 2001*, ACM Press (2001), 17-22.

16. Pohjanheimo, L., Ailisto, H. and Plomp, J. User experiment with physical pointing for accessing services with a mobile device. In *Proc. EUSAI Workshop on Ambient Intelligent Technologies for Wellbeing at Home* (2004).

17. Rekimoto, J. and Nagao, K. The world through the computer: computer augmented interaction with real world environments. In *Proc. 8th annual ACM symposium on User interface and software technology*, ACM Press, (1995) 29-36.

18. Rekimoto, J. and Ayatsuka, Y. CyberCode: designing augmented reality environments with visual tags. In *Proc. DARE 2000 on Designing Augmented Reality Environments*, ACM Press (2000) 1-10.

19. Riekki, J., Salminen, T. and Alakärppä, I. Requesting pervasive services by touching RFID tags. *Pervasive Computing* 5, 1 (2006), 40-46.

20. Swindells, C., Inkpen, K.M., Dill, J.C. and Tory, M. That one there! Pointing to establish device identity. In *Proc. 15th annual ACM symposium on User interface software and technology*, ACM Press (2002), 151-160.

21. Toye, E., Madhavapedy, A., Sharp, R., Scott, D., Blackwell, A. and Upton, E. Using camera-phones to interact with context-aware mobile services. Technical report, UCAM-CL-TR-609, University of Cambridge, 2004.

22. Tuomisto, T., Välkkynen, P. and Ylisaukko-oja, A. RFID tag reader system emulator to support touching, pointing and scanning, in A. Ferscha, R. Mayrhofer, T. Strang, C. Linnhoff-Popien, A. Dey, A. Butz and A. Schmidt (eds.), *Advances in Pervasive Computing*, Austrian Computer Society (2005), 85-88.

23. Tuulari, E. and Ylisaukko-oja, A. SoapBox: a platform for ubiquitous computing research and applications. In *Proc. Pervasive Computing 2002*, Springer-Verlag (2002), 125-138.

24. Välkkynen, P. Korhonen, I., Plomp, J., Tuomisto, T., Cluitmans, L., Ailisto, H. and Seppä, H. A user interaction paradigm for physical browsing and near-object control based on tags, in *Proc. Physical Interaction Workshop on Real World User Interfaces*, (2003), 31-34

25. Want, R., Fishkin, K. P., Gujar, A. and Harrison, B. L. Bridging physical and virtual worlds with electronic tags. In *Proc. CHI 99*, ACM Press (1999), 370-377.

26. Wellner, P. Interacting with paper on the DigitalDesk. *Communications of the ACM*, 36, 7 (1993), 87-96.

Suggestions for visualising physical hyperlinks

Suggestions for Visualising Physical Hyperlinks

Pasi Välkkynen, Timo Tuomisto and Ilkka Korhonen

VTT Technical Research Centre of Finland
P. O. Box 1300, 33101 Tampere, Finland
{firstname.lastname}@vtt.fi

Abstract. Physical browsing is a tag-based and mobile device-centric interaction paradigm for ambient intelligence settings. It is a means of associating physical objects and their related digital information by creating physical hyperlinks. The tags, or physical hyperlinks from the user's point of view, have to be visualised to tell the user of the existence and location of the links, how they can be selected and what action they contain. Based on link visualisation in desktop WWW, we present the problem domain, some suggestions for visualising links in ambient intelligence settings, and example visualisation designs to illustrate our ideas.

1   Introduction

In physical browsing, the user can access information or services about an object by physically selecting the object itself, for example by touching or pointing to the object with a mobile terminal. The enabling technology for this is tags that contain the information – for example a Universal Resource Locator (URL) – related to the object to which it is attached, and mobile terminals equipped with tag readers. A tagged physical environment can be seen as analogous to a WWW page. It contains links to different services and the user can 'click' these links with a mobile terminal in the same way desktop WWW links can be selected with a mouse.

The first step in physical browsing is *physical selection*, a method for the user to tell the mobile terminal which tag it should read – that is, which physical object the user wants to access. We have defined three physical selection methods: pointing, touching and scanning [1] and implemented them on a PDA [2]. Touching means selecting the link by bringing the terminal very close to the link. Touching can be implemented for example with short-range Radio-Frequency Identifier (RFID) tags. Pointing is a directional long-range selection method, similar to a TV remote control. One way to implement pointable tags is to add a photosensitive sensor to a long-range (typically a few metres) RFID tag with a sensor interface[1] and a pointing beam into the mobile terminal. In scanning all tags within the reading range are read and presented to the user as a list from which the user can continue the selection using the

---

[1] Long-range passive RFID tags with a sensor interface are not yet generally available, but they are being developed for example in European project MIMOSA, see http://www.mimosa-fp6.com.

graphical user interface (GUI) of the terminal. Scanning requires from the tags only that they are readable over a long range, and for example Bluetooth devices typically use scanning as the selection method.

Ballaghas *et al.* [3] have analysed the design space of mobile phone based interactions in ubiquitous computing. In their design space, physical selection is situated in the "selection" interaction task, and touching and pointing are in the "selection of object with direct pick device" interaction technique. Scanning does not directly map into the selection technique categories of Ballaghas *et al.*

After physical selection, an *action* follows. The action can be anything the terminal is capable of, for example loading a WWW page, making a phone call, setting a mobile phone to silent mode, turning on the lights of the room, or reading a temperature sensor. The action can be determined indirectly by mapping a tag identifier into some action, or the action can be stored directly in the tag. For example, Mifare1k[2] RFID tags can contain one kilobyte of information, which is enough to directly store URLs in the tag itself.

One research challenge in physical browsing and especially in the physical selection phase is the visualisation of the links in the physical environment. The main challenges are communicating to the user 1) the existence of a link, 2) the precise location of the link, 3) the supported selection methods, if the link does not support all selection methods, and 4) the action of the link. By visualising the link in the physical environment, we mean marking the technical device implementing the link somehow – for example in the case of a Radio-Frequency Identifier (RFID) tag it would mean adding some symbols to the RFID "sticker" itself.

To better illustrate the design challenge, let us consider a simple scenario: a movie poster augmented with several short-range RFID tags. The poster includes touchable links for 1) the movie's WWW page, 2) downloading free movie trailer and theme song via Hypertext Transfer Protocol (HTTP), 3) ordering a chargeable theme (screensaver, background image and ring tone) via mobile phone messaging system, 4) a WWW link to the local movie theatre's ticket buying service and 5) the phone number of the same service. As the poster is located on a bus stop, there are also other links in the environment. The bus stop sign itself is augmented with a pointable long-range tag that contains real-time information about the buses that are stopping in this location. All these services could of course be behind one WWW page, but physical browsing allows for more natural interaction than using a WWW service with a mobile phone. With all these services having a physical link of their own, the user interface is transferred from the small screen to the large poster and from menu selections to touching or pointing the desired links directly on the poster.

In the scenario, it is useful to indicate the presence of the links in the poster so that the user waiting for the bus will know there are some interactive services available. The technology of the poster tags limits their selection method to touching, so the precise locations need to be marked. The user should also have information about how the tags of the poster and the bus stop sign can be selected so that touch-range tags will not inadvertently be pointed. Finally, there is a difference between the ac-

---

[2] See for example
http://www.semiconductors.philips.com/products/identification/mifare/classic

tions. Some of them contain free-of-charge information, some chargeable data and some (the ticket reservation call tag) communications between people.

Visualisation happens in several levels of physical browsing. In addition to the physical link, the information can be presented in the terminal. A useful aid in determining the action of the link would be 'hovering' in the same way it is possible in current desktop WWW browsers. In hovering, the user could select the link, but instead of immediately activating it, the terminal would query the tag for the action and display it to the user. It is also possible to display the link for confirmation after selection or create a separate page containing visualised virtual links after selecting the physical link in the object. However, in our physical browsing evaluations the users have preferred to open simple information links as effortlessly as possible without confirmations and intra-terminal operations. Our whole visualisation work intends to support this single most important case: the simple free-of-charge information content must open with a minimum of confirmations or menu selections. We see this as an essential requirement for natural interaction in ambient intelligence environments and therefore the links must be presented so that they support this.

There is no single correct solution to the visualisation chain, rather this work should be seen primarily as an analysis of the challenges, and as a suggestion for the physical presentation layer for the links. Our simple visualisation example at the end of the paper should be seen only as an example to illustrate our views and the challenges we have found, rather than as a real proposition for visualisations.

2   Physical Hyperlinks in Related Work

Riekki *et al.* [4] have built a system that allows requesting pervasive services into a mobile phone by touching RFID tags with a separate RFID reader connected to the phone. As a part of their work, they have designed visualisations for the tags. The basis of the symbols is a 'general' tag depicting a generic RFID tag. In addition to the general tag, they have added into the general tag special icons for example for printing, setting phone profile and placing a call. These special tags describe additional information related to the action of the tag. Riekki *et al.* report in their results that the users preferred the special tags because "the special tags present their meaning in a concrete way that they felt was logical, fast, and easy to use". Users also felt the special tags were more secure because they could better predict the action. Because the system of Riekki *et al.* only supports selection by touching, they did not visualise the selection method in any way.

Want *et al.* describe several applications [5] in which they have used RFID tags as physical hyperlinks. One of them is Photo Cube, which has a picture of a person on each side. Embedded into the picture is a tag that contains a link to that person's home page. In other applications, the link is embedded under a descriptive text and the user has to figure out the existence and action of the link by guessing from the appearance of the object, or by experimenting.

Kindberg *et al.* have built CoolTown [6], which is implemented with both active infrared tags and passive RFID tags. In CoolTown, the links are often scanned for automatically and displayed in the mobile terminal instead of the environment. They

have also created eSquirt, a protocol to transfer URIs between devices. The user has to point the mobile terminal towards the target device to send the link, meaning that the user has to know the presence and approximate location of the receiving "link". Another directional IR-based system is GesturePen by Swindells *et al*. [7]. In their article, they have a picture of an object they have augmented with an infrared tag, which is visible only as a technical device, but without any other identification of its presence or function.

In visual tagging systems [8], the tags are inherently visible, and thus clearly communicate their presence to the user. However, the tags are all similar in appearance, and the action does not show in any way.

The current commercial physical links visualise their implementation technologies. These links are for example NFC[3] symbols, or logos of commercial Bluetooth[4] and infrared-based local service providers[5], forcing the users to concentrate on the technology instead of the content.

3   Implications from Scenarios

We have designed several tag-based ambient intelligence scenarios [9] that utilise physical selection as one of their basic interaction patterns. From these scenarios, we have identified the following common classes of actions in addition to viewing simple information content:

– making a phone call or sending a message,
– downloading and installing an application,
– starting an application in the terminal,
– connecting an external wireless device to the terminal,
– user identification, for example mobile terminal as a key,
– setting the terminal state, for example to silent mode,
– controlling an external device, for example lights,
– reading a sensor attached to a tag, and
– reading tag contents for future use, for example for 'drop' operation, and 'dropping' saved tag contents to a target device.

In addition to the actions, some non-functional characteristics might be useful to visualise. Examples of these are the cost of the service, and whether the communication is local or external.

---

[3] http://www.nfc-forum.org
[4] http://www.bluetooth.com
[5] See for example http://www.bluecasting.com, http://www.kameleon-europe.com, http://www.wideray.com

4 Lessons from Hyperlinks in Desktop WWW

4.1 Visualising Links

Hypertext Markup Language (HTML) [10] defines two components for each hyperlink: the Universal Resource Identifier (URI) [11], or address of the resource, and the visual presentation of the link. The presentation can be for example a fragment of text or an image. The usual way for presenting a textual link has been to use a different colour for the link and to underline it [12, 13]. Textual links are generally used in two ways: as navigational link lists, or as links inside regular text. For example, an almost standard navigational aid in current websites includes a navigation bar – a list of textual links – on the left side of the page. Nielsen and Tahir state [13] that in the navigation bars it is not strictly necessary to visualise the link by colours and underlining. The navigation bars have become so common that users are normally able to recognise them as link lists without any specific visualisation techniques. We expect the physical links to gain a similar status with time; that is, the users will learn the common places and uses for these links.

An image is another common presentation for a link. Image icons are used in the same way as in non-browser graphical user interfaces and in addition to them, any image can act as a link.

By default, most WWW browsers display the address of the link in the status field or as a tool tip while the cursor is hovering above the link. For the user, this is a useful indicator of the existence of the link and gives a hint about the action by displaying the URI or link title. We suggest hovering as a similar interaction technique for mobile terminals and tags.

4.2 Visualising Actions

With *action*, we mean the result of selecting a link, for example displaying a new WWW page or opening an email client to send mail to an address specified in the link. The resource identified by the URI is typically another WWW page. It may also be of different content type, for example an image or a PDF document. Displaying these resources is left to the WWW browser and it may display them itself or launch another application to display the resource. An example of a content type that typically is handled outside the web browser is text/x-vCalendar[6], which is used to transmit calendar entries and is usually handled in a specific application. The main classes of actions available in desktop WWW are thus:

– Opening a new WWW page or another resource that the browser itself can handle, for example an image and
– Opening a new resource the WWW browser cannot handle on its own, such as a calendar entry, or a mailto link.

---

[6] http://www.imc.org/pdi

Nielsen and Tahir have suggested in their guidelines that any link that opens a resource other than another web page should be designed so that it clearly tells the user what will happen [13]. This visualisation is very important when the link opens another program and takes the user away from the browser application. A classic example of bad link design is to link a person's name to his or her email address. When clicking the link, the user expects intuitively to see the person's WWW page, but is instead thrown into the email application. Instead, the link visualisation should indicate the action, for example by displaying the email address as link text instead, or using a commonly recognised email icon as the presentation. We see the visualisation of physical hyperlinks as a similar problem and that lessons learned from desktop web can and should be applied in physical environments as well. The action of the link should thus be included in the link visualisation.

4.3   Context of the Link

In addition to the immediate link presentation, users interpret the existence and the properties of the link from other sources. WWW pages often follow regular patterns, like having in the upper left corner an image link that leads to the front page of the site [13]. Users have learned these through visiting several sites that behave in similar ways and many designers have learned to adhere to these unwritten standards.

This is possibly the most important lesson from desktop browsing. The physical objects typically already have some function and the users will expect the link to be related to that function. The object itself can thus be seen as one part of the link presentation, just as the whole page or context of the WWW link is part of its presentation. However, there are still questions regarding the physical hyperlink, for example whether a link in a door gives information about the status of the door or toggles the lock.

5   Visualising Physical Hyperlinks

As stated earlier, the first two challenges for visualising physical hyperlinks consider the existence and location of the link. Any visual presentation in the same spot as the tag will tell the user that there is a link, and it will tell the user the exact location of the link. The next two challenges consider the selection method and the action. There are only a limited number of methods for selecting a link in a physical environment, even if new selection methods in addition to pointing, touching and scanning are needed and defined.

The number of actions is, however, large and not limited in any way. In desktop user interfaces, there are some conventions for commonly used icons, such as a floppy disk often representing the save function. However, for more uncommon actions the designers often design their own icons that may be difficult to understand for the users. Another option is to reduce the amount of visualisations by classifying the different actions into wider groups based on what the user expects from selecting the link. The most precise way to visualise the action is a specific icon for each possible action. The other extreme is a single icon that only communicates the presence of a

link. The aforementioned classification categories may be a reasonable compromise. On the other hand, for example setting the terminal state is such an abstract concept that it is hard to design a good icon for it. The more concrete icon for setting the terminal into silent mode is far easier and understandable to the user.

We have chosen not to visualise the implementation technology at all other than implicitly in the selection method. Optimally, the user should not be required to know whether a link can be read using RFID, Bluetooth or some other technology. For example, when the user touches a link, the terminal should use all the short-range technologies it can to read the contents of the tag. The current physical links work exactly opposite to this and concentrate on visualising the implementation technology.

5.1   Constructing Visualisations

Some simple ways to design a visualisation for a link include using a *geometrical shape* for the whole visualisation, a *colour*, and a *shape for the actual content icon.* The actions are more variable than the selection methods and the icon content shape allows a large number of variations. Therefore, we suggest the content shape to be reserved for the action, as in desktop GUI icons.

This will leave for the selection methods the other visualisation techniques, for example the overall shape of the visualisation and the colour. The colour does not associate with any selection method but if for example presenting pointable tags as blue becomes a standard, the users will learn it. It may be possible to design overall shapes for the links so that they naturally map to the different selection methods. In addition to that, we presume that shape will also become a learned notion, if for example simple geometrical shapes like circles, triangles and squares are used around the action icon. Ideally, the link can be selected by touching, pointing and scanning, but the technical limitations of tags may exclude some methods. In the next subsection, we suggest an additive method as an example that can visualise different selection methods by combining selection symbols around the action icon.

Textual links are also a possibility and they are unambiguous if the link text is chosen well. While it could be possible to use link text in physical hyperlinks too, it would cause some problems in practice. In desktop environments, it is easy to control the size of the link text, but links in small objects or links farther away from the user might become difficult to recognise.

5.2   Example Visualisations

In Figure 1, there are icons for visualising actions: connection/control, information retrieval, download/install, messaging and an icon for setting the terminal to silent mode. It should be noted that these are only examples to illustrate the concept of visualisation. For example, the 'i' icon is comprehensible only for people who live in cultures that associate 'i' with information.
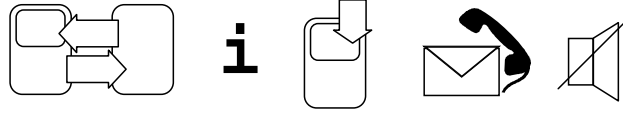
Fig. 1. Actions. From left to right: connecting, information, download, messaging and silent mode.
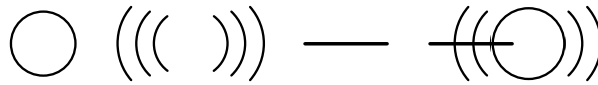

Fig. 2. Selection methods. From left to right: touch, scan, point and a tag supporting all three selection methods.

In Figure 2, there are additional additive visualisations for presenting the selection methods. The first one is for touchable (short-range) tags and the second with 'radio waves' is for tags that can be read by scanning (long-range). The third symbol is for pointing and in the last symbol, the three selection symbols are combined into one visualisation for a link that can be selected by any selection method.
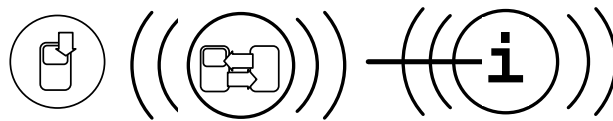
Fig. 3. Actions and selection methods combined.

We show three combinations of actions and selection methods in Figure 3. The first visualisation represents a link that can be read by touching and that downloads or installs software to the terminal. The second one shows a link that can be selected by touching or scanning and that connects the terminal to another device. The third visualisation describes a link that contains a simple information page and can be selected using any selection method.

6   Conclusions

We have analysed how links can be selected in current desktop World Wide Web, and what actions are available both in web and in physical browsing systems. Based on these analyses we have presented suggestions for visualising links in ambient intelligence settings, and a few example visualisations. In addition to the link visualisation itself, the context and perceived affordances of the object will also give suggestions

for the action, for example, a link in a door is likely to open or lock it and not turn on the lights of the room.

We have implemented a preliminary version of the hovering concept as a part of a physical browsing application. Our future work on physical hyperlink visualisation will include further studying hovering and how to divide the visualisations between the physical tag and the graphical user interface of the mobile terminal.

Acknowledgements

References

1. Välkkynen, P., Korhonen, I., Plomp, J., Tuomisto, T., Cluitmans, L., Ailisto, H., Seppä, H.: A user interaction paradigm for physical browsing and near-object control based on tags. In: Proceedings of Physical Interaction Workshop on Real-world User Interfaces (2003) 31-34
2. Tuomisto, T., Välkkynen, P., Ylisaukko-oja, A.: RFID Tag reader system emulator to support touching, pointing and scanning. In: A. Ferscha, R. Mayrhofer, T. Strang, C. Linnhoff-Popien, A. Dey, A. Butz and A. Schmidt (eds.): Advances in Pervasive Computing. Austrian Computer Society, Vienna, Austria (2005) 85-88
3. Ballaghas, R., Borchers, J., Rohs, M., Sheridan, J.: The Smart Phone: A Ubiquitous Input Device. Pervasive Computing 5(1) (2006) 70-77
4. Riekki, J., Salminen, T., Alakärppä, I.: Requesting Pervasive Services by Touching RFID Tags. Pervasive Computing, 5(1) (2006) 40-46
5. Want, R., Fishkin, K. P., Gujar, A., Harrison, B. L.: Bridging Physical and Virtual Worlds with Electronic Tags. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, New York, USA (1999) 370-377
6. Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B., Spasojevic, M.: People, Places, Things: Web Presence for the Real World. Mobile Networks and Applications, 7(5) (2002) 365-376
7. Swindells, C., Inkpen, K. M., Dill, J. C., Tory, M.: That one there! Pointing to establish device identity. In: Proceedings of the 15th annual ACM symposium on User interface software and technology, ACM Press, New York, USA (2002) 151-160
8. Ljungstrand, P., Holmquist, L.E.: WebStickers: Using Physical Objects as WWW Bookmarks. In: CHI'99 extended abstracts on Human factors in computing systems. ACM Press, New York, USA (1999) 332-333
9. Kaasinen, E., Rentto, K., Ikonen, V., Välkkynen, P.: MIMOSA Initial Usage Scenarios. (2004) Available at http://www.mimosa-fp6.com
10. Raggett, D., Le Hors, A., Jacobs, I.: HTML 4.01 Specification. W3C Recommendation 24 December 1999. Available at http://www.w3.org/TR/html4/
11. Berners-Lee, T.: Universal Resource Identifiers. Available at http://www.w3.org/Addressing/URL/URI_Overview.html

12. Lie, H., Bos, B.: Cascading Style Sheets, level 1. W3C Recommendation 17 Dec. 1996, Revised 11 Jan. 1999. Available at http://www.w3.org/TR/CSS1
13. Nielsen, J., Tahir, J.: Homepage Usability: 50 Websites Deconstructed. New Riders Publishing, Indianapolis, USA (2002)

Hovering: visualising RFID hyperlinks in a mobile phone

Hovering: Visualising RFID Hyperlinks in a Mobile Phone

Pasi Välkkynen
VTT Technical Research Centre of
Finland
P.O.Box 1300
33101 Tampere, Finland
+358-20-722 3353

pasi.valkkynen@vtt.fi

ABSTRACT

Physical browsing is a mobile terminal and tag based interaction paradigm for pervasive computing environments. The tags offer to the users physical hyperlinks that can be read with a mobile terminal and that lead to some pervasive services or information. Hovering is an interaction technique, which allows the user to quickly check the contents of a tag by 'hovering' the mobile terminal over the tag. In this paper, I describe a prototype system that implements the hovering concept with a mobile phone and RFID tags. The purpose of the system is to study physical hyperlink visualisations, both in the physical environment, and in the graphical user interface of the mobile terminal.

Categories and Subject Descriptors

H5.2 [**Information interfaces and presentation**]: User interfaces – *input devices and strategies.*

General Terms

Human Factors.

Keywords

RFID, physical browsing, mobile phone, hyperlink, visualisation

1. INTRODUCTION

In physical browsing, the user can access information or services about an object by physically selecting the object itself, for example by touching or pointing to the object. The enabling technology for this is tags that contain the information – for example a Universal Resource Locator (URL) – related to the object to which it is attached. A tagged physical environment can be seen as analogous to a WWW page. It contains *physical hyperlinks* to different services and the user can 'click' these links with a mobile terminal in the same way desktop WWW links can be selected with a mouse.

As links in desktop web, the physical hyperlinks should be visualised to let the user know that 1) there is a link, 2) where it is located, 3) how it can be selected and 4) what will happen after the link is selected. The visualisation can happen in many levels: in the physical object itself the tag may have some icons representing its action and selection method, or the link can be visualised in various ways in the graphical user interface of the mobile terminal.

In this paper, I describe a physical browsing system built on a Nokia 3220 mobile phone. This application enables users to use physical shortcuts to activate digital services in their mobile phone. The purpose of the system is to be a tool for studying user interaction with physical hyperlinks. This system allows similar interaction with links as Nokia's built-in "Service Discovery" application but with extended link visualisation capabilities.

Physical hyperlink visualisation is still a relatively unstudied issue. We have presented some challenges [4] related to the visualisation of the tags. Riekki *et al*. [3] have also studied visualisations of RFID (Radio-frequency Identifier) tags. Generally, physical browsing systems in literature (for example [1], [2] and [5]) do not report in detail their pre-selection visualisations, if they exist. Weinreich & Lamersdorf [6] have implemented a link visualisation system for desktop WWW. Their system takes into account several attributes of a link, for example title, author, language and server response and display them as tooltips when the pointer is hovering over the link.

2. USER INTERACTION

The basic sequence of touch-based mobile interaction with physical hyperlinks is that the user brings the mobile terminal close to the link, after which the terminal reads the contents of the link and displays it to the user. In hovering, the user can 'hover' the mobile terminal over a link similarly to how hovering works in desktop web. In desktop web browser, when the pointer is hovering over a link, additional information about the link is typically displayed. The browser usually displays the address the link leads to in the status bar and if the link has a title, it is displayed as a tooltip next to the link. In this mobile phone based hovering, the link information is displayed in the mobile phone screen before the link is actually selected and activated. This way the user can quickly check the contents of several links before actually selecting any of them (Figure 1).

Figure 1. The user is checking what links the business card contains.

Hovering differs from confirmation dialogs ("Do you want to go to http://www.foo.com?") by not being a question to be answered. It does not present a modal dialog that has to be answered, instead it quickly displays some information about the link and the user can hover over several links to check each of their contents.

There are two main display modes in the hovering application: single and list (see Figure 2). In the single mode, only one link at a time is displayed but more information is available. In the list mode several links are displayed as a list. In either mode, pressing the Select button will activate the link, for example show the information or make the phone call to the number read from the tag. In the list mode, more information about the link, similarly to the single mode, is displayed when the user chooses to view the link list item in its entirety.

Figure 2. On the left is shown the single link display mode. Only one link is displayed but with more information than in the list mode on the right.

Each link has a title, contents and an icon. The title is a human-readable short description of the link. The contents contain the actual content of the link, for example a web address. The icon gives a graphical cue about the type of the content so that the user does not necessarily have to try to figure it out from the content resource.

As seen in Figure 3, each content type has its own visual icon. The purpose of the icon is to give a quick way to see what the type of the link content is. Additionally, it is intended to help differentiate in the list mode between links that have the same title but different content type. For example, a phone number and a web address might both have as the title the name of the person whose phone number and web address they are, but with different icons they can be quickly told from each other.
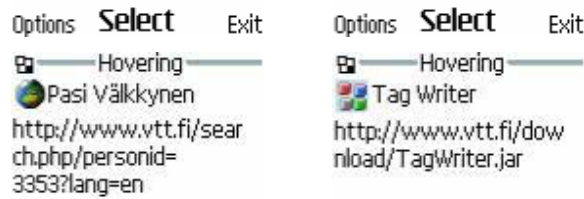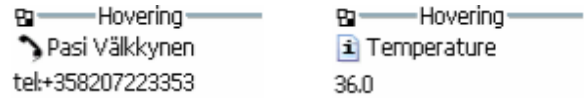

Figure 3. Different content types. On the upper row there are visualisations for phone call and local information, and on the lower row remote information and installable application.

All information links were initially visualised as 'i' but with the current mobile network speeds, there is a huge difference with interacting only with local connectivity or with remote services (for example a WWW page). If a link leads to an external communication, it is visualised with a globe symbol in the hovering application and only local services are visualised as 'i'. One reason for the globe symbol is that Nokia 3220 uses the globe icon for web access. This should make it easy to recognise to a user who is used to the phone.

3. SYSTEM

The hovering system is built on Nokia 3220 mobile phone[1] with Nokia's Xpress-on NFC shell. The software platform of the phone is S40[2] that can run Java MIDlets.

NFC[3] records are used to store the data in the tags in different fields. Each tag has a Title field, and a URI field. The Title field is used to display the title of the link in human-readable form and the URI is used to store the content. The content can be a link to a web resource, a telephone number, a link to a JAR file for downloading and installing applications, or a sensor reading.

The sensor reading does not come from a real sensor, instead it is a random number from a suitable range. The purpose of the sensor "mock-up" is to demonstrate to users how mobile phone based interaction with RFID sensors might look and feel.

---

[1] www.europe.nokia.com/nokia/0,8764,58033,00.html

[2] www.forum.nokia.com/main/0,6566,010_200,00.html

[3] www.nfc-forum.org

The icon is determined from the content of the URI field in the tag. The content type could be checked by querying the web server (in case of WWW resources) but that would take a considerable amount of time with current cellular connection. And after all, the purpose of the system is to be a tool for visualisation studies instead of a physical browsing system that implements all possible security features. I have chosen the same approach as in desktop WWW browsers: the user is given the link title and if he or she can understand how URLs work, the address can also be investigated.

4. CONCLUSION

Optimally the links are visualised also in the physical objects, so that the user can know how to select the link and what action it contains. Hovering can help 1) visualise the action if only selection technology is visualised in the tag (for example NFC symbol), and 2) give additional information about the link such as the actual URL.

The future work on this concept will include building some tagged environments and evaluating the concept with users. Some questions in the evaluation will be the general usefulness of hovering, which display mode (single or list) is more useful and what information the user needs to see about the link. The intention is to study how hovering works with physical visualisations on the tags and how best combine these two visualisation techniques. The current prototype will also be extended to allow interaction with more types of contents, for example SMS messages and tags that can set some context information for the phone.

5. REFERENCES

[1] Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B., Spasojevic, M.: People, Places, Things: Web Presence for the Real World. *Mobile Networks and Applications*, 7(5), 2002. 365-376

[2] Leichtenstern, K., Rukzio, E., Chin, J., Callaghan, V., and Schmidt, A. Mobile interaction in smart environments. *Advances in Pervasive Computing 2006*, 2006. 43-47.

[3] Riekki, J., Salminen, T., Alakärppä, I.: Requesting Pervasive Services by Touching RFID Tags. *Pervasive Computing*, 5(1), 2006. 40-46

[4] Välkkynen, P., Tuomisto, T., and Korhonen, I. Suggestions for visualising physical hyperlinks. *Pervasive 2006 Workshop Proceedings* (Dublin, Ireland May 2006). 245-254.

[5] Want, R., Fishkin, K. P., Gujar, A., Harrison, B. L.: Bridging Physical and Virtual Worlds with Electronic Tags. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, New York, USA, 1999. 370-377

[6] Weinreich, H. and Lamersdorf, W. Concepts for improved visualization of web link attributes. *Computer Networks* 33(1-6), 2000. 403-416.

PAPER VIII

Ambient functionality – use cases

Ambient Functionality – Use Cases

Eija Kaasinen, Timo Tuomisto & Pasi Välkkynen

VTT Information Technology, Sinitaival 6 (P.O.Box 1206), FI-33101 Tampere, Finland
`{firstname.lastname}@vtt.fi`

Abstract

In this paper we describe use cases and user requirements for ambient intelligence applications on personal mobile devices. Wireless connections to tags and sensors provide mobile applications with different identification, measurement and context data. Mobile applications that utilise local connectivity share many common patterns. We have identified these common patterns and describe them as use cases related to physical selection, activating applications, sensing and context-awareness. Based on user and expert evaluations of usage scenarios we also present user requirements for the use cases.

1. Introduction

Mobile devices are increasingly evolving into tools to orientate in and interact with the environment, thus introducing a mobile-device-based approach to ambient intelligence. As mobile devices are equipped with tag and sensor readers, they provide platforms for different applications that make use of local connectivity. These applications share many common features related to collecting measurement data, communicating with objects in the environment, identifying contexts, activating applications and so on. These features include common usage patterns that can be illustrated as use cases.

In this paper we describe common use cases that we have identified for mobile applications that utilise local connectivity. We also describe initial user requirements regarding these use cases, based on scenario evaluations with users. We start with an overview of related research in section 2 and then describe our mobile-device-based ambient intelligence architecture in section 3. In section 4 we introduce our design approach and present excerpts of usage scenarios to illustrate what kinds of applications our design targets at. In section 5 we describe the identified use cases and the user requirements related to them.

2. Related research

A lot of research on ambient intelligence is going on but the results still cover only dedicated services and specific applications [1]. Different basic component technologies such as mobile devices, sensors, ad-hoc networks and computing technologies are already available, but advancements are needed in their integration, scalability and heterogeneity [2].

In the following we will give an overview of research related to local connectivity, especially research related to using tag data for identification, collecting sensor data wirelessly, and using these two to identify the context of use.

2.1. Physical selection

Want, Weiser and Mynatt have identified the coordination of real and virtual objects as a key research problem in ambient intelligence [3]. Kindberg [4] has introduced the term "physical browsing" as "the users obtain information pages about physical items that they find and scan". Physical selection can be seen as the phase in physical browsing by which the user selects with a mobile terminal a physical or a virtual object for interaction. As a result of physical selection, the content of the tag is interpreted and an action is launched.

Want et al. have carried out interesting work regarding the association of physical objects with virtual ones [5]. They have built several prototypes, some of which were implemented with RFID (radio frequency identification) tags read with an RFID reader connected to a PC. Generally their selection method was touching, that is, reading from a short range.

Kindberg et al. [6] have created Cooltown, in which people, places and physical objects are connected to corresponding web sites. The user interaction theme of Cooltown is based on physical hyperlinks, which the users can collect using a mobile terminal to easily access services related to people, places and things. Cooltown utilises IR (infrared) and RFID technologies to transmit these links to the users' mobile terminals.

Holmquist, Redström and Ljungstrand [7] have built WebStickers, a desktop-based system to help the users better manage their desktop computer bookmarks. Their system illustrates well the association between digital and physical worlds by using tags. The idea of WebStickers is to make physical objects act as bookmarks by coupling digital information (URLs, uniform resource locators) to them.

The above-mentioned research projects have used mainly static information or dynamic information composed of different sources in their applications. Measurement data from local sensors has typically not been part of the applications.

2.2. Wireless connections to sensors

In mobile applications that utilise sensor data, sensors can be embedded within a mobile terminal or connected wirelessly to the terminal as separate sensor units via local connectivity.

Sensors can measure environmental information such as location, altitude, illumination, temperature, pressure or surrounding sounds. Personal sensor data on physiological activity and mental state can be measured using electrocardiography [8], accelerometers, galvanic skin resistance [9] etc. Current application areas for personal data monitoring include fitness, sports, wellness and ambulatory monitoring in healthcare.

Bluetooth is currently the most commonly available wireless personal area network (WPAN) for local connectivity. Other low power technologies used for local connectivity are ZigBee and ANT[10] used e.g. in Suunto T6 wristop [11] computer for communicating with a heart rate belt and with an accelerometer installed in the shoe.

Commercial Bluetooth equipped sensor units are still rare [12]. Commercial wearable sensor units typically act as data loggers whose content needs to be uploaded daily or weekly into a user terminal with USB (Universal Serial Bus), infra red, or a proprietary radio link. Common wellness monitors – weighing scale and blood pressure meter – still lack wireless connections or even wired connections to digital infrastructures.

A major challenge with wireless sensor units is how sensors should be introduced and connected to the mobile device. In Bluetooth pairing any two Bluetooth devices form a trusted pair: whenever the other is detected, the paired devices automatically accept communication, bypassing authentication. The Near Field Communication (NFC) Protocol [13] proposes using NFC devices to establish a Bluetooth connection by touch and connect. Hall et al. [14] point out the advantages of linking RFID tags to Bluetooth enabled devices. As the tag contains connection parameters, the system may bypass the discovery process - which otherwise may take even 10 seconds. If the tag performs wakeup for the sensor unit, also power consumption is reduced. The idea of bypassing the Bluetooth device discovery by physically selecting and reading the Bluetooth address has also been utilised in reading visual tags with camera phone [15] or pointing by IrDA (Infrared Data Association) [16].

2.3. Context-awareness

An efficient way to improve the usability of mobile services and applications is adapting the content and presentation of the service to each individual user and his/her current context of use. In this way, the amount of user interaction will be minimised: the user has quick access to the information or services that (s)he needs in his/her current context of use. The services can even be invoked and the information provided to the user automatically.

A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task [17]. The main challenge with context adaptation is the reliable measurement or identification of the context that may include physical, technical, social and emotional elements.

In previous research, context-awareness has mainly been studied in restricted application areas, such as tourist guidance [18], museum guide [19], e-mail [20], mobile network administration, medical care and office visitor information [21]. In these studies, the location of the user is the main attribute used in the context-adaptation. Designing applications for wider contexts of use will require more measurement data.

3. MIMOSA application platform

According to the vision of the MIMOSA project (Microsystems Platform for Mobile Services and Applications) [22], a personal mobile phone is the trusted intelligent user interface to the ambient intelligence environment and a gateway between the tags, sensors, networks of sensors, the public network and the Internet. For realising such a vision, MIMOSA develops and implements an open architecture based on novel low-power microsystems devices integrated on a common technology platform. The approach is based on short-range connectivity that can be set up with relatively modest investments in the infrastructure. At the same time, however, a wide range of consumer applications is covered.

Local connectivity to tags provides access to data and functions related to the environment and different objects in it. Local connectivity to sensors can be utilised to monitor the environment or the user. Tag data and different measurements can be analysed and processed to identify the current context of use.
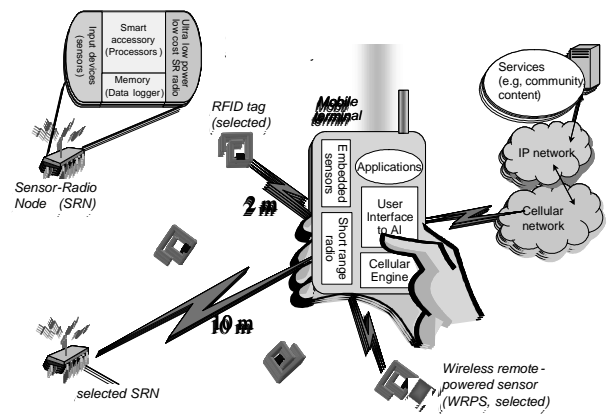

Figure 1. MIMOSA applications can utilise local connectivity to tags and sensors

At the top level the MIMOSA architecture can be divided into four physical entities: mobile terminal, remote application server, sensor radio node (SRN) and wireless remote-powered sensor (WPRS) that have local connectivity capability so that they can communicate with each other (Figure 1). Wireless remote powered sensor includes a sensor and a tag for identification. Sensor radio node is a more versatile unit that may include several sensors, memory and micro-controller that enable the pre-processing of sensor data.

The mobile terminal can collect information from the sensors automatically or by user initiation. The terminal can also read different RFID tags. Optional remote connectivity allows connections to remote application servers on the Internet.

4. The design approach

In our vision, the user feels and really is in control of ambient intelligence applications that are accessible through his/her personal mobile device. These applications help people in their everyday lives: they are useful, usable, reliable, and ethical issues have been taken into account in the design. We are developing demonstrator applications on the platform parallel to designing the MIMOSA platform itself. We are focusing on four representative consumer application fields where user mobility can be combined with measurements of the user and his/her environment: sports, fitness, health care and housing. In addition, we are studying general everyday applications.

Our project team has defined together a common vision of the future and illustrated the vision in the form of usage scenarios. These scenarios (36 short stories) have been used as the basis for user requirements. During the first phases of our project we analysed the scenarios in order to identify common patterns as use cases and related functional requirements. We evaluated the scenarios with end users and application field experts (94 people in total) in focus groups to identify user requirements regarding individual applications, common patterns and the architecture. The following excerpts of our usage scenarios illustrate the kinds of applications that we are targeting at:

1. Lisa notices an interesting poster of a new movie at the bus stop. She points her phone at the poster and presses a button to download a sample video clip for viewing on her bus trip. As the bus has not yet arrived, Lisa points to a tag at the bus stop to check how late the bus will be and to get guidance in redesigning her travel route.

2. John has heard from his doctor that many of his health problems are related to his overweight. Now he has started to use a mobile application that motivates his efforts to eat less and exercise more. As he steps on the scale, his weight is transferred into his mobile phone, where he can see his progress day by day. The scale also measures his fat rate and dehydration level. This morning the welfare application on his phone recommends that he should have two glasses of water to balance his dehydration level. As John starts his daily jogging exercise, his motions are monitored in order to give him feedback of the energy consumption. A lactate sensor indicates if he should slow down in order to keep the right aerobic level for fat-burning.

3. Matthew is quite serious with his golf practise. As he comes to the golf course, his golf application on his wristop is activated. All of Matthew's golf clubs are equipped with tags and sensors. His wristop computer runs a golf application that helps him to keep track of the clubs he has used and the courses he has played at. During the game, the golf clubs send measurement data of each swing to the wristop computer. At home Matthew compares his results to those of Tiger Woods.

Scenario 1 illustrates how physical objects can include links to further information and functions. Scenario 2 illustrates how health measurements can support self-care. Scenario 3 illustrates how an application is activated as the mobile device identifies a certain context as well as more versatile local connectivity to tags and sensors.

5. Use cases and user requirements

MIMOSA use cases are related to communicating with nearby objects, monitoring measurement data as well as identifying contexts and reacting accordingly. In the following we will describe four categories of use cases in detail: the user interface paradigm of physical selection, activating applications, collecting sensor data and context-awareness.

For each category, we will describe a collection of typical use cases with related user requirements. We illustrate the use cases as UML (Unified Modeling Language) style sequence diagrams where the interacting objects are users, mobile terminals, tags, and sensors. User requirements are based on user evaluations of the scenarios as well as analysis of the use cases by usability experts.

5.1. Physical selection

Physical selection is a user interaction paradigm that allows the user to select a link embedded in a physical object. The technical devices providing the links in MIMOSA are RFID tags and sensor radio nodes. The content of the link may be any digital information related to the physical object, for example a link to a web page or a sensor reading. Additionally, physical selection can be used to intuitively connect devices to each other in order to activate communication between them. The use cases for physical selection are the three methods for selecting the tag with the mobile terminal: touching, pointing and scanning.

5.1.1.   Use cases

Touching means selecting the link by bringing the terminal close to the tag. The tags may be read by using continuous polling to detect when the terminal is near a tag, or only when the user also presses a button in the terminal.
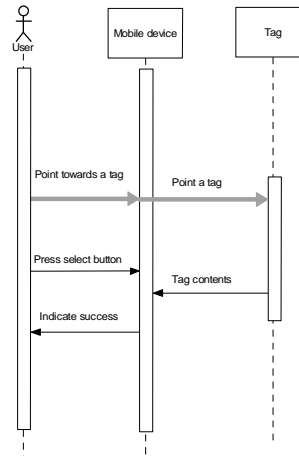

Figure 2. Physical selection by pointing.

In *pointing* the tag is selected by aiming a beam at the tag. The beam may be either visible light, for example laser, or invisible, for example infra red, possibly assisted with a visible beam for easier aiming. To detect the pointing beam the tag has to be equipped with a sensor.

Scanning reads all the tags in the environment and presents them in the mobile terminal's display. The user can then select the tags (s)he needs to interact with.

Figure 2 depicts how pointing works. Grey arrows are used to describe changes in the spatial relations of physical objects, e.g. moving the mobile terminal to point at a tag.

5.1.2.   User requirements

The tags and nodes should support all three physical selection methods: touching, pointing and scanning, and they should behave consistently in the interaction. Consistency supports users in adopting this novel user interface paradigm.

As our user evaluations were based on written and illustrated scenarios, the interviewees based their conception of the look and feel of tags to familiar concepts such as bar codes or infra red controllers. Key user requirements include marking the tags in a consistent way so that the user can identify them on different objects and in the environment. The appearance of the tag could indicate the functions that are included in it. The reading should be easy; the user should not have to wave his/her mobile device back and forth. Fluent reading of tags is a central requirement that becomes especially important in tasks where the reading should take place unnoticed (e.g. while jogging and running pass a check point) or when the user reads a collection of tags. Application-specific requirements for physical selection include for example properties of the pointing beam and required reading ranges for the tags.

The users emphasised the difference between personal tags and public tags. Public tags that can be read by anyone are useful in public places such as the poster in scenario 1. When tags are connected to personal items such as medicine packages, tags that can be read anywhere by anyone were seen as a privacy threat – anyone can unnoticeably read any tags that you are carrying or wearing. Long reading distance is also a threat for privacy. There seems to be trade-off to solve between easy reading and privacy protection.

5.2. Activation of applications

Many applications are bound to a limited space and activity, or are available only locally. Activating contextually relevant applications was a repeated action in our scenarios.

5.2.1. Use cases

A *context tag* is a tag that holds enough information to start an application that is already installed on the mobile device, or to install a new application in the mobile device. For example, in scenario 3, the user comes to a golf course and touches a context tag at the entrance to start the golf software.

After the installation the application can be associated with the ID of the tag so that next time the already installed application can be activated based on reading the tag. Figure 3 illustrates starting an application by reading a context tag.

5.2.2. User requirements

In the user evaluations, the requirements for ease of taking the applications into use were obvious. Although our scenarios described ready made installations, the interviewees often referred to problems that they expected to face in installing and configuring the systems.

Especially in relation to sports and fitness scenarios, the interviewees pointed out that they would not have time for complex set-up operations. The interviewees also saw that if each single sensor has to be activated separately, they easily would forget some of them. In the middle of the exercise it will be too late to start the application or measurements. The system should get started "with a single button". As for the jogging scenario, the interviewees proposed that measuring should start automatically "as the running speed grows".

The users are expected to use their personal mobile device for several different purposes, e.g., when playing golf, for fitness control, home control etc. The coexistence and fluent activation of situationally relevant applications will be a major challenge for taking the services into use. Context-based activation of applications is a promising concept to ease taking the applications into use.
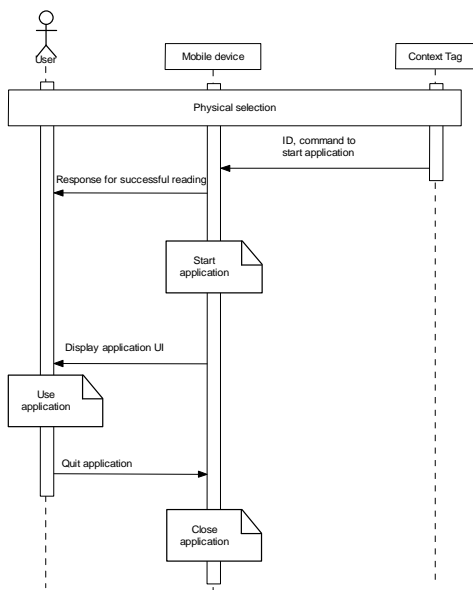

Figure 3. Starting an application by reading a tag.

Additional configuration challenges will be faced when starting applications because each application may require a

different set of sensors and tags. When starting an application for the first time, the physical selection paradigm may be used for introducing and registering sensors and tags to the application. The application may then use these registrations or "pairings" (~analogue to Bluetooth pairing) for automatic connectivity to the sensors. Such a sensor may also be registered as a context tag for the application: by physically selecting the sensor, the application is started automatically.

5.3. Collecting sensor data

In our scenarios, we could identify three kinds of use cases where sensor data was collected: a single reading of a sensor, reading sensor data periodically, and establishing a connection between a smart sensor unit and the mobile device, whereby communication is possible in both directions. These use cases are described in the following subsections. In the use cases we presume that the sensors have already been introduced to the application as described in section 5.2.

5.3.1. Select and read

Single sensor measurements can be made just by reading a wireless remote-powered sensor (WRPS) after physical selection. As an example in scenario 2, the scale is equipped with a tag whose contents include the current measured weight. The welfare application may be continuously active on the mobile device. Being already introduced to the application, the scale can be read by just touching or pointing to it. If the application is not active but the scale tag has already been associated as a context tag with the welfare application, the application can be started by reading the tag.

5.3.2. Select and read periodically

Many MIMOSA scenarios include a background application, running continuously and capable of reading tag contents from a distance. If the application is capable of reserving its own time slot in the tag reader, it should be capable of controlling the reading range and reading interval of the reader as well. The application then also should own all the data associated with the time slot.

As an example, a lactate sensor (scenario 2) may be read continuously by the welfare application. The application periodically reserves the tag reader for its own purposes, increases the reading range from touching to some appropriate level, and reads the contents of the tag.

5.3.3. Select and establish connection

Measurement data can be transferred between a sensor radio node and a terminal by establishing a bidirectional connection between these devices (Figure 4).

The physical selection paradigm and the application activation paradigms described in sections 5.1 and 5.2 make measurements easier by reducing the time for discovery, selection and application activation. This requires that the sensor radio node is equipped with a wakeup tag.

Our scenario 3 utilised bidirectional communication between a wristop device and an integrated motion measurement unit (IMU). Figure 5 illustrates that when a golf swing is recognised by the sensor radio node (SRN), the SRN sends the club ID to the wristop computer, which logs the club used and the current position from GPS. After reading the ID – in addition to reading GPS data – the application might query and read the whole swing data for later analysis.
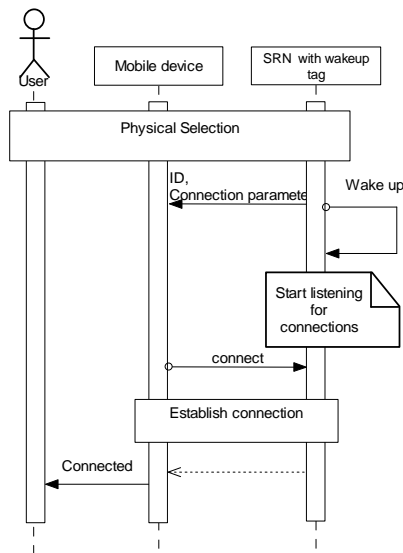
Figure 4. Using sensor radio node with a wakeup tag.

5.3.4.   User requirements

In general the users appreciated the possibility to measure things that you "cannot see or feel". These kinds of measurements could be, for instance, health parameters, proper stretching of body before and after physical exercise, environmental conditions or safety-related measurements at home.

The interviewees innovated several new kinds of sensors that might be useful. Self-care systems are easy to deploy, as they do not require much infrastructure around. There were needs for both short-term and long-term monitoring (months or even years) of physical condition.
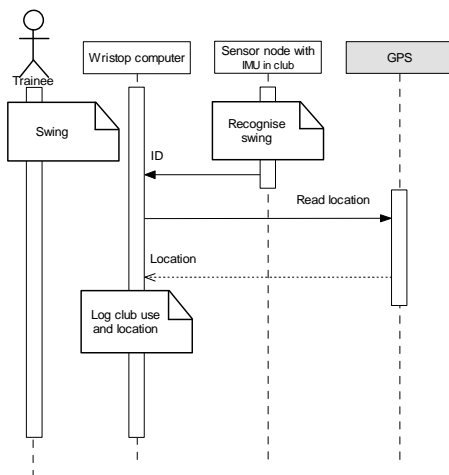

Figure 5. Logging golf club usage.

The running scenario included several measurements. The interviewees pointed out that the user should have the freedom to choose which measurements (s)he wants to include. Situations vary, and the user may want to measure different things on different days.

Motion monitoring that was a central theme in sports-related scenarios, raised some doubts related to the necessity of the information provided. The users also questioned whether measurements could be reliably analysed and synthesised to the kind of feedback that they were expecting.

For instance, the motion of the golf club should be analysed to offer the user not only measurement curves but information on what is wrong with his/her swing.

5.4. Context-awareness

From the system's point of view, context-awareness includes three elements: identifying the context, maintaining user profiles, i.e. information on user needs in different contexts and using information about the context with the user profile to provide situationally relevant services, information and functions to the user.

Contexts can be identified by combining and analysing different measurement data. The measurements take place as described in section 5.3, and the system concludes the context based on the measurements of the user and of the environment. As the context includes several physical, social, technical and emotional elements, the measurements need to be very diverse. Still, a reliable identification of the context is often difficult.

Contexts can also be identified by using tags. Tags may indicate nearby objects and thus define the context. The system can read the context tags automatically or the user can actively select a context tag to inform his/her device about the context. Tag-based context-identification was described in section 5.2 in connection to activating applications. In a similar way, an already active application can utilise context tags. For instance, a golf application can identify the clubs that the user has been using as well as his/her location on the course. This information can be used to identify the current phase of the game. The context can be defined in more detail if the tag-based context information is complemented with measurement data of the user and the environment.

In the user evaluations, the users emphasised the need for ease of taking an application into use and using it. This requirement was crucial for applications targeted at sports and fitness. Tag-based context identification was felt acceptable because it was in the user's control and because it supported well the required ease of use. Depending on the application, the context can be identified totally automatically or by user initiative. The latter alternative can utilise physical selection as described in section 5.1.

Automatic identification of the context presents challenges for tag reading. Each application and each context tag may require different reading distance. The reading distance has to be accurate to identify the context reliably. For instance, in health care applications it would be beneficial to identify that the user is taking medicine. How near does the medicine package then have to be to conclude that the user is about to take the medicine? In the golf application, the club has to be next to the wrist long enough before the system can conclude that the player is planning to swing with that club.

6.   Conclusions and further work

By defining a variety of usage scenarios and analysing them we revealed use cases that repeat in different applications utilising local connectivity. We have learned a great deal about user requirements for those use cases by analysing the scenarios and evaluating them with users and application field experts. Fulfilling the user requirements related to these basic patterns are crucial to the success of ambient functionality. In addition to the MIMOSA specific application fields, the use cases and related user requirements can presumably be applied in other application fields as well.

To study user interaction in more detail, we will continue our work by building proof of concept prototypes that

illustrate key use cases. By evaluating the proof-of-concept prototypes we expect to get more detailed user feedback about viable interaction patterns and user acceptance of local connectivity based applications in general.

Acknowledgements

References

[1] Streitz, N. and Nixon, P. (2005) *The Disappearing Computer*. In *Communications of the ACM*, 48, 3 (March 2005). p. 33-35.

[2] Saha, D. and A. Mukherjee, A. (2003) *Pervasive Computing: A Paradigm for the 21st Century.* in *Computer*, 36, 3 (March 2003). p. 25-31.

[3] Want, R., Weiser, M. and Mynatt, E. (1998) *Activating Everyday Objects.* in *Proceedings of the 1998 DARPA/NIST Smart Spaces Workshop*. 1998. p. 140-143.

[4] Kindberg, T. (2002) *Implementing Physical Hyperlinks Using Ubiquitous Identifier Resolution.* in *Proceedings of the eleventh international conference on World Wide Web*, Honolulu, Hawaii, USA. ACM Press, New York NY, USA. p. 191-199.

[5] Want, R., Fishkin, K. P., Gujar, A. and Harrison, B. L. (1999) *Bridging Physical and Virtual Worlds with Electronic Tags*. in *Proceedings of CHI 99*, Pittsburgh, PA, USA. p. 370-377.

[6] Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B. and Spasojevic, M. (2002) *People, Places, Things: Web Presence for the Real World*. in *Mobile Networks and Applications*, Volume 7, Issue 5, October 2002. p. 365-376.

[7] Holmquist, L. E., Redström, J. and Ljungstrand, P. (1999) *Token-Based Access to Digital Information*. in *Proceedings of the 1st International Symposium on handheld and ubiquitous computing*, Karlsruhe, Germany. p. 234-245.

[8] Polar Electro Heart Rate Monitors, www.polar.fi, accessed 28.4.2005.

[9] HealtWear Armband, http://www.healthwear.com/hw/ healthwear_armband.do, accessed 28.4.2005

[10] ANT personal area network, www.thisisant.com, accessed 28.4.2005

[11] Suunto, Suunto T6 wristop, www.suunto.com, accessed 28.4.2005

[12] Nonin, 4100 Bluetooth enabled Pulse Oximeter, http://www.nonin.com/products/oem/4100.asp, accessed 28.4.2005

[13] NFC, Near Field Communication White Paper, http://www.ecma-international.org/activities/ Communications/2004tg19-001.pdf, accessed 28.4.2005.

[14] Hall, E., Vawdrey, D. and Knutson, C. (2002) *RF Rendez-Blue: Reducing Power And Inquiry Costs In Bluetooth-Enabled Mobile Systems*. in *Proceedings of the Eleventh IEEE International Conference on Computer Communications and Networks (IC3N '02)*, Miami, Florida, October 14-16, 2002.

[15] Scott, D., Sharp, R., Madhavapeddy, A. and Upton, E. (2005) *Using Visual Tags to Bypass Bluetooth Device Discovery*. in *Mobile Computing and Communication Review*, Volume 9, Number 1 (January 2005)

[16] Woodings, R., Joos, D., Clifton, T., and Knutson C. (2002) *Rapid Heterogeneous ad hoc Connection Establishment: Accelerating Bluetooth Inquiry Using IrDA*. in *Proceedings of the Wireless Communications and Networking Conference (WCNC2002)*. p.342-349.

[17] Dey, A. K. (2001) *Understanding and using context*. in *Personal and Ubiquitous Computing*, 5. p. 20–24.

[18] Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, C. (2000) *Developing a context-aware electronic tourist guide: some issues and experiences*. in *CHI 2000 Conference Proceedings*. p. 17–24.

[19] Ciavarella, C. and Paterno, F. (2004) *The design of a handheld, location-aware guide for indoor environments*. in *Personal and Ubiquitous Computing*, 8. p. 82–91.

[20] Ueda, H., Tsukamoto, M. and Nishio, S. (2000) *W-MAIL: An Electronic Mail System for Wearable Computing Environments*. in *Proceedings of the 6th Annual Conference on Mobile Computing and Networking (MOBICOM 2000)*.

[21] Chávez, E., Ide, R. and Kirste, T. (1999) *Interactive applications of personal situation-aware assistants*. in *Computers & Graphics*, 23. p. 903–915.

[22] MIMOSA project home page. www.mimosa-fp6.com

Identifying user requirements for a mobile terminal centric ubiquitous computing architecture

Identifying User Requirements for a Mobile Terminal Centric Ubiquitous Computing Architecture

Eija Kaasinen, Marketta Niemelä, Timo
Tuomisto and Pasi Välkkynen
VTT, Technical Research Centre of Finland
Tampere, Finland
{firstname.lastname}@vtt.fi

Vladimir Ermolov
Nokia Research Center
Helsinki, Finland
Vladimir.Ermolov@nokia.com

Abstract

System level solutions affect many properties of ubiquitous applications and thus also user experience. That is why user point of view should guide the design of mobile architectures although the users will see them indirectly, via the applications. This paper describes our approach in identifying user requirements for a ubiquitous computing architecture that facilitates mobile applications sensing their environment. The sensing is based on wireless connectivity to tags and sensors in the environment. We illustrated a representative set of future applications as scenarios and proof-of-concepts and evaluated them with potential users. Scenarios were analyzed to identify generic use cases and to understand the implications of the user feedback on the architecture. Our experiences show that user requirements for system level solutions can be identified with this approach. We identified several requirements for the architecture dealing with user interaction, wireless measurements, context-awareness, taking applications into use and ethical issues.

1. Introduction

Mobile terminals are increasingly evolving into tools to orient in the environment and to interact with it, thus introducing a mobile terminal centric approach to ubiquitous computing [1]. With ubiquitous computing, many properties of the applications are defined by the underlying system level solutions that are in charge of providing wireless connections to the objects in the environment, providing external connections, identifying contexts and so on. To ensure user acceptance of ubiquitous applications, we should be able to get user feedback already to the design of the system level solutions. Although human-centered design of end-user applications is a well-established and even standardized process [2], approaches to studying systematically user requirements for system level solutions are still rare.

In this paper we describe our experiences in identifying user requirements for a ubiquitous computing architecture. The research work was carried out within an EC-funded research project MIMOSA, Microsystems Platform for Mobile Applications and Services [3]. The project is developing novel microsystems solutions for wireless sensors and tags as well as a mobile platform that facilitates connecting those microsystems wirelessly to mobile phones and other personal mobile devices.

We start this paper with an overview of related research; reported experiences in extending human-centered design approach to system level solutions. After that we introduce our design target: the microsystems platform for mobile services and applications. Then we describe the user requirements definition process that we applied in the MIMOSA project. The rest of the paper describes and analyses key user requirements identified and their implications on the architecture.

2. Related research

Implicit in the idea of ubiquitous computing are the notion of infrastructures – some new, some existing, some virtual, some physical, some technical, some social – all coming together in a seamless way [4]. The potentially vast collection of devices, sensors and personalized applications presents a remarkable challenge to the design of ubiquitous computing infrastructures [5]. Some researchers have touched the issue of focusing more on infrastructures although they do not report actual experiences [6] – [8]. Islam and Fayad [8] emphasize the necessity to develop devices parallel to mobile infrastructures: new types of contents require new types of devices; new network

options require devices that can choose the most suitable network in each context of use and so on.

The well-established human-centered design approach as defined by the ISO 13407 standard [2] is focused on the design of individual applications. The standard guides to iterative prototyping that is not easy in current complex and concurrent application development [9]. For system level solutions that include both hardware and software, iterative prototyping would be almost impossible. We will need new methodological approaches to identify user requirements for system level solutions.

Edwards et al. [10] have presented encouraging experimental results of two case studies where they have identified user requirements for software infrastructures: a file system and a context toolkit. Edwards et al. [10] emphasize the importance of simple proof-of-concepts that illustrate core infrastructure features. User evaluations of these proof-of-concepts gave early feedback to the design of the infrastructure and ensured that the most central system level features were designed to facilitate user-friendly applications.

The case study described in this paper has taken a new challenge as we are focused on a platform architecture that includes both hardware and software. The selection of possible services that our platform will facilitate is much broader than the applications dealt with in the study Edwards et al. [10]. That is why in our case we needed to complement proof-of-concepts with lighter ways to illustrate and evaluate future possibilities – scenarios as described in section 4.

3. The MIMOSA architecture

The planned MIMOSA platform architecture provides mobile users with a smooth transition from current mobile services to ubiquitous computing services. In MIMOSA approach the personal mobile terminal becomes a tool for the user to sense the environment and to interact with it. The approach is based on short-range connectivity that can be set up with relatively modest investments in the infrastructure. A wide range of consumer applications is covered as the architecture facilitates wireless connections to different tags and sensor units [3].

At the top level MIMOSA architecture can be dived into four physical entities and wireless interfaces between the entities. The entities are mobile terminal, remote application server, sensor radio node and wireless remote powered sensor (WRPS). The physical entities are illustrated in Figure 1. As the MIMOSA project is developing novel microsystems solutions

parallel to the development of the architecture, for instance the targeted reading distances are longer and the targeted component sizes smaller than in currently available tag and sensor solutions.
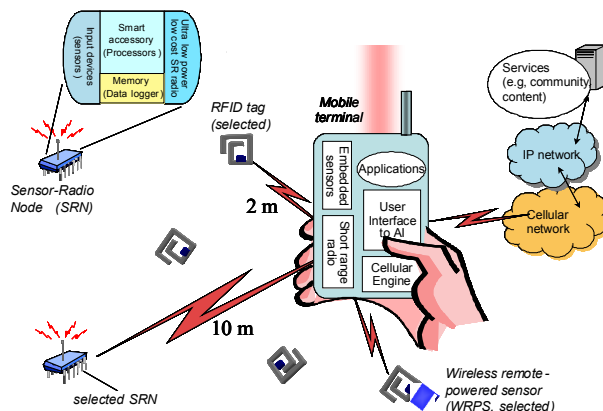

Figure 1. The MIMOSA architecture facilitates local connectivity to tags and sensors

The mobile terminal, the sensor radio node and WRPS all have local connectivity capability so that they can communicate with each other. The terminal can collect information from the sensors automatically or based on user initiation. The terminal is also capable to read different RFID (Radio Frequency Identification) tags. Optional remote connectivity allows for remote application servers in the Internet to apply locally obtained sensor data within the context of specific services and applications.

Terminal device is the trusted device of the user. Most important blocks of the terminal are processor platform, local connectivity module, user interface hardware and embedded sensors. Processor platform is running an operating system. User interface hardware supports easy and flexible interaction between user and terminal. Local connectivity module offers connections to surrounding sensor devices and tags. Individual applications handle information collection, processing and displaying of the results. Context engine is a software module that monitors sensor measurements, identifies context atoms and further contexts and informs applications about contexts according to how the applications have registered for contexts.

Wireless remote powered sensor (WRPS) has architecture similar to RFID tags. Parts of the WRPS are digital control logic, including radio protocols and information controlling, analogue front end, rectifier, power management, and sensor. From the system point of view WRPS can be seen as a special kind of tag from where can be read not only the tag ID but also a sensor reading.

Sensor radio node (SRN) provides a more versatile and longer range sensor connection. SRN consists of a micro-controller-based host (MCU), local connectivity module and sensors. Micro-controller has a processor core that enables application specific software running on the sensor unit. This includes communication protocol stack and sensor information processing. Pre-processing capability in the sensor radio node can provide high-level sensor information for user applications running on the terminal. Sensor radio node can use the radio to advertise services that are available.

As the design and implementation work of the architecture started at the same time as our user requirements definition process, we could not target to influence much on the first architecture prototype but the focus was on the next generation architecture that would be commercialized based on the results of the MIMOSA project.

In the following, we will use the term platform to describe the personal mobile terminal enhanced with local connectivity. Sensor unit is used as a term to describe generally the different sensors and sensor-radio nodes. Architecture describes the whole entity of the mobile terminal, sensors, sensor radio nodes, tags, remotely connected servers as well as related system level software.

4. The process of identifying user requirements

When identifying user requirements for system level solutions, it is essential to see the wide variety of future applications. We decided to focus on four promising consumer application fields: sports, fitness, health care and housing. To enhance coverage we were also studying general everyday applications.

MIMOSA project gathered together a multitechnical project group that included partners developing individual technical solutions and application field experts from different organizations. A scenario workshop was organized in the beginning of the project to define a common vision and to integrate the ideas of the partners. The common vision was illustrated in the form of totally 36 usage scenarios that described comprehensively different application possibilities facilitated by the architecture in the form of excerpts of usage situations.

The scenarios were evaluated with potential users and application field experts (94 people in total) in focus groups (3 to 8 participants in each group interview) and by individual questionnaires where the participants assessed the value of proposed solutions.

In addition we evaluated the scenarios with our project partners with web questionnaires.

Based on the scenario evaluation results, we refined the scenarios ending up with 20 more focused scenarios. Those second generation scenarios presented applications that were accepted by end users and application field experts. Also the credibility of the proposed technical solutions had been checked with project partners. The following excerpt of a golf scenario as well as Figure 2 illustrates the scenarios:

Figure 2. The golf scenario

Matthew is quite serious with his golf practice. As he arrives at the first tee at the golf course, his golf application on his wristop is activated. All of Matthew's golf clubs are equipped with tags and sensors. His wristop computer runs a golf application that helps him to keep track of the clubs he has used and the courses he has played at. During the game, the golf clubs send measurement data of each swing to the wristop computer. At home Matthew compares his results to those of Tiger Woods.

After the scenario refinement, we analyzed the scenarios into use cases and further into sequence diagrams to describe in details what kind of communication they indicated between the user, the application and the different parts of the architecture. The use cases and sequence diagrams were discussed through with the technology developers of the architecture to ensure that the proposed sequences were correctly interpreted and possible to realize [1]. Figure 3 illustrates one of the sequence diagrams related to the golf scenario.

The users could give feedback on many application features based on the illustrated scenarios but features where the look and feel was essential were difficult to imagine and comment. To study and compare alternative look and feel properties, we built proof-of-concept demonstrators. Those simple prototypes were

based on current technology and they illustrated basic application and interaction concepts.
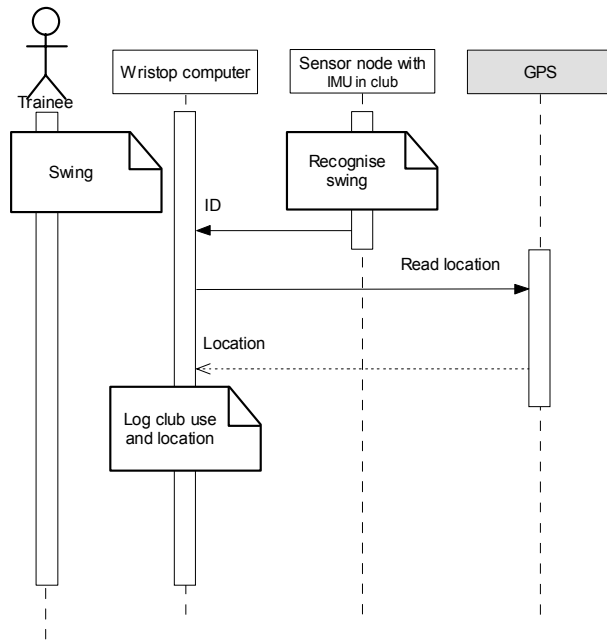

Figure 3. A sequence diagram of the analyzed golf scenario

The proof-of-concepts included for instance a PDA-based prototype that illustrated user interaction with tags in the environment, context-aware everyday services implemented on a tag-reader phone, golf proof-of-concept with motion-monitoring of the club and fitness proof-of-concept integrating different fitness measurements. We evaluated those proof-of-concepts each with 3-8 users in usability tests in laboratory conditions or in the field.

The user evaluations gave us feedback both on proposed functionalities and on the quality attributes for those functionalities. To identify the implications on the architecture, we analyzed the user cases and related user feedback together with architecture designers in several workshops. The emphasis was on generic use cases repeating from one scenario to another as they obviously had strong impact on the architecture [1].

When user requirements are defined for an individual application, we can get quite precise requirements. With system level solutions, the users are commenting individual applications and we need to integrate and interpret their feedback to see the implications on the architectural level. The applicable research methods are mainly qualitative and that is why identified requirements tend to be descriptions of trends rather than statistical evidence. We ended up

with a set of user requirements for the architecture classified as: 1) User interaction by physical selection, 2) Gathering measurement data from wireless sensors, 3) Context-awareness, 4) Taking applications into use and 5) Ethical issues. In the following section we will describe key user requirements identified and their implications on the architecture using this classification.

5. User requirements for the architecture

5.1 User Interaction by Physical Selection

Physical selection allows the user to select an object for further interaction. The selection can be done by touching or pointing with the mobile device but the device can also scan the environment and propose objects for interaction [11]-[14]. In our scenarios physical selection was a recurring usage pattern. The scenarios gave us an overview of the possible ways of user interaction but a user study was required to evaluate and compare the usage patterns more thoroughly and to identify parameters for the physical selection methods. We built a physical selection prototype, which emulated user interaction with MIMOSA style passive RFID tags and supported selection by touching, pointing and scanning. The user study with thirteen test subjects was then conducted using this prototype to find out answers to the specific research questions such as reading distances [15]. In the following, we first describe general user requirements regarding physical selection (gathered by scenario analysis) and then additional requirements for touching and pointing (from the user study).

5.1.1. General User Requirements for Physical Selection. The scenarios indicated need for each of the three selection methods. In an environment with many tags, it will be hard to select the correct one from the list presented after scanning. Therefore, there is a need for touching and pointing when the user knows the approximate location of the tag.

The analysis of the scenarios revealed that there should be a default action depending on the type of the tag read. If the tag contains for example a URL, the terminal should open the web browser and display the page. This means that a) there should be a tag-reading application running in the terminal all the time, and b) it should be able to determine an action for (most) tags the user selects. This also indicates that tags may include meta-information about their contents. Physical selection should make actions happen without the user first starting an application to interpret the tag contents.

5.1.2. User Requirements for Touching. There are three alternative (and possibly complementary) techniques for touching tags: a) touching is always on and available, b) touching is on and available after a "touch mode" is activated, and c) touching requires an additional specific action (for example a button press).

In the physical browsing user study, the users preferred touching being continuously on compared to touching with a button press. However, touching without a button could cause problems, for example if the terminal is put into a purse full of tagged objects as the terminal would then continuously select the tags.

The results indicate that the best technique for touching would probably be having a "touch mode" that can be activated and deactivated either manually or automatically. The alternative use of a button press confirming a touch action could be user configurable.

5.1.3 User Requirements for Pointing. Pointing is a directional selection method. There are three alternative and possibly complementary techniques for pointing at tags: a) pointing with a directed RF field, b) pointing with a visible pointing beam and c) pointing with an invisible pointing beam (for example infrared).

In the physical browsing user study, the pointing technique without a visible aid was more difficult than the techniques in which there was a visual aid.

The results indicate that a visible pointing aid is needed but there may still be a need for invisible pointing technology due to privacy concerns. To ease pointing, a visible pointing aid could be used with a wider invisible beam.

5.2 Gathering measurement data from wireless sensors

Gathering measurement data wirelessly was well illustrated in the scenarios. The scenario evaluations revealed user requirements for special cases such as long term data logging and using multiple devices. The user evaluations of the fitness and golf proof-of-concepts provided additional feedback on activating the measurements.

5.2.1 Physical selection in activating measurements. The initial vision in the scenarios was that every tag and sensor within the reach of the terminal device reading distance is continuously measured. Scenario evaluations with application field experts and user evaluations with the proof-of-concepts revealed that not all sensors can be read continuously as they only intermittently contain valid data. The user evaluations with proof-of-concepts indicated that physical selection can be utilized in activating measurements. E.g., the

user can get a single measurement or initiate a series of measurements by touching a sensor unit.

User feedback has several implications. Physical selection should activate proper application(s) to interpret and analyze the sensor reading. Each sensor reading must be addressed to an application that can validate and process the reading and control the sampling frequency. The validated measurement data can be made available for other applications as well.

5.2.2 Data logging in sensor units. In the scenario evaluations the interviewees told that during many sports activities, for instance when running they would prefer leaving the mobile terminal at home. They asked whether they could just carry the sensor unit and then later at home transfer the data to the mobile terminal.

User feedback implies that data logging in sensor radio node should be supported. This would require time stamping in sensor units. The architecture should facilitate data synchronization between data logging sensor radio nodes and the mobile terminal.

5.2.3 Long term monitoring. Most scenarios described only short-term monitoring. In the scenario evaluations application field experts pointed out the need for long-term monitoring (even years) for instance to get more insight to the symptoms of a patient.

These requirements suggest that the architecture needs to cooperate with external servers that can store the long-term monitored data.

5.2.4. Support for multiple user terminals. In the scenario evaluations the interviewees pointed out that one user may have several personal mobile terminals. For instance (s)he may use a wristop computer when jogging and a mobile phone when walking.

User feedback implies that the user should have seamless access to his/her exercise data with any device. The gathering of measurement data should continue fluently even if the user changes the terminal device temporarily or permanently.

5.3 Context-awareness

User feedback from the scenario evaluations emphasized user control and the role of context-awareness in activating situationally relevant applications. Field evaluations with the proof-of-concept that illustrated the use of context tags confirmed the potential of tags in activating contexts.

5.3.1. Context-based application management. User feedback from scenario evaluations and proof-of-

concept evaluations emphasized effortless use facilitated by context recognition. Initially we were focusing on context-awareness within individual applications but already when defining the scenarios we noticed that context-aware activation of applications was even more important. For instance in the golf scenario, the golf application activates as the user arrives at the golf course. Context-awareness within the application is presented in the same scenario when the golf application starts to record the swing as it recognizes the club near the wristop computer. Both kinds of context-awareness were repeated in our scenarios and widely accepted by the users.

User feedback implies that the architecture should take care of context-based application management, i.e. applications should be activated and deactivated according to the recognized contexts.

5.3.2. Tag-based context recognition. In the scenarios tag-based context recognition often turned out to be adequate because contexts are frequently related to certain places or certain everyday objects. Embedded tags can reveal the place or the objects around the user and thus the context. In the proof-of-concept evaluations the users accepted well context-recognition based on the user actively selecting a tag by physical selection because it saved user efforts.

User feedback implies that the architecture should support context tags: special tags that facilitate easy activation or even download of applications, or context-aware actions within an application. Contexts could be recognized more accurately if the architecture could measure distances and directions to the tags.

5.3.3. User control. In the proof-of-concept evaluations the users wanted to confirm most context-based actions. Only simple actions that are easy to cancel were accepted without user confirmation.

User feedback implies that the architecture needs to include elements that facilitate user control of context-aware features. Context validation may be required both in automatic and in user-initiated context recognition. Validation can take place by asking the user for confirmation or by reading additional sensors. Context-aware actions may require user confirmation as well as undo or cancel functions available.

5.4. Taking applications into use

Although our scenarios described ready-made installations, in the scenario evaluations the interviewees often referred to problems that they expected to face in installing and configuring the systems. Especially in relation to sports and fitness

scenarios, the interviewees would not have time for complex set-up operations. The interviewees also saw that if each single sensor would have to be activated separately, the user would easily forget some of them.

Taking an application into use requires installation of the application, introducing appropriate sensor units to it and configuring the setup. The usage patterns can be defined according to the temporal state of usage: a) the user connects the terminal and a sensor unit for one session, b) the user introduces the terminal and the sensor unit permanently for future sessions, c) the user activates the application, or the application activates itself automatically and d) the application connects itself to previously introduced sensor units.

5.4.1   Activating applications. In the scenario evaluations the users were worried about the installation of the applications that they expected to be complex. For many of the applications the interviewees could see only occasional usage, thus keeping the applications permanently active does not make sense. User feedback in the scenario evaluations confirmed that the user may use his/her personal mobile terminal for several different purposes. Each application may require a different set of sensors and tags, and the same application can be used with varying sets of sensors. Different applications can be active in parallel and by turns. The user may have applications that need to be continuously active, for instance health monitoring applications. The activation of other applications should not disturb these. The user evaluations with proof-of-concepts confirmed the potential of context tags as solutions to ease taking applications into use.

User feedback implies that the architecture should support the co-existence and fluent activation of situationally relevant applications. The architecture should support easy downloading and setting up of new applications as well as easy uninstallation of applications that the user no more needs. Installation and activation of applications can be made effortless to the user with tag-based context-awareness as described in the previous chapter.

5.4.2   Private sensors. In the scenario evaluations the users wondered how complicated it would be to connect the sensors wirelessly to the mobile device. They also were worried about how to make sure that nobody else is reading their personal sensors indicating for instance information on their health parameters.

User feedback implies that the architecture should support both public and private sensors. The architecture should provide the user with tools for the management of private sensor units. The tools should facilitate introducing the sensor units to the personal mobile terminal and the application on it, initiating a

measurement period and measuring. Introduction is needed to associate the sensors to the terminal and to a specific application, and to set up authentication and encryption parameters between the terminal and the sensor unit.

Using physical selection methods to introduce the sensors to the terminal device has good potential as it keeps to common usage patterns. Touching is a preferable method due to small range and minimal chance for eavesdropping. Scanning can be used in connection to setup wizards with user confirmations.

5.5. Ethical issues

The ethical issues raised up in the user evaluations were in line with earlier research [16]-[18]. We however found that many ethical issues can actually be dealt with at the level of the architecture, and the personal mobile terminal centric approach has potential to respond to the ethical challenges. The personal mobile terminal is a trusted device for personal data, providing facilities to ensure keeping the user in control. This constitutes a good basis for ethically acceptable solutions.

We shortly describe key ethical issues considering privacy, security, and trust that were arisen in the user evaluations and the scenario analysis.

1. What kind of information is retrieved of the user? The mobile terminal centric approach is focused on personal information (e.g. health measurements) that is very private to the user. Furthermore, context-related information such as nearby tags and personal context history are private information. Even if the user's identity cannot be directly revealed from the information, delivering it may be ethically arguable.

2. Who can get the information? The user should have the right to decide who can receive personal information. Physical selection by pointing and scanning, and generally reading from a distance, increase possibilities for eavesdropping. Privacy of the data communication should be protected by authorization mechanisms.

3. The user needs feedback when personal information is transferred. Tags are normally read to a certain application on the user device. However, there may be other tag readers around. The user should know if tags and sensors that (s)he wears or carries along are readable also to external readers and should get feedback if this kind of reading takes place.

4. The user needs effortless ways to protect valuable information. Personal data stored on the user's mobile terminal must be protected against theft and losing by providing back-up and authorization mechanisms. Connections to external servers need to be secure.

5. The mobile terminal and applications should be protected against external attacks. Context tags and downloadable applications may provide access to virus programs and other hostile attacks. The user needs means to ensure the reliability of downloadable applications and their providers.

6. Reliability limitations with a mobile terminal based solution should be considered. The terminal may get lost, the batteries may drain out, and server connections may break. Backup mechanisms are needed to handle these situations.

These issues highlight the user requirements of awareness, control, and feedback of data stored in or mediated via the personal mobile device. The requirements are important as personal data may be maliciously used to threaten the user's privacy or security. The user should not however be responsible alone for protecting against these threats, but ethical issues should be taken into consideration at the level of the architecture, in compliance with laws and common moral norms and rules.

6. Conclusions

Our experiences point out that several system level design decisions concerning a mobile terminal centric ubiquitous computing architecture have significant influences on the end user experience. Our experiences also point out that user requirements regarding the architecture can be identified early in the design by illustrating the forthcoming applications as scenarios and proof-of-concepts and by evaluating those illustrations with potential users. Rich scenario material is essential to identify common usage patterns and further use cases. User feedback needs to be analyzed thoroughly in cooperation with technical experts to interpret user requirements to implications on the architecture.

Our vision of mobile terminal centric approach to ubiquitous computing was well accepted among the users. The occasional usage needs identified in connection to many of the proposed solutions backed up our design approach of a low-cost infrastructure built on the user's personal mobile terminal and providing platform for several kinds of applications and services. However, user feedback changed quite a lot our initial vision of the functionality of the architecture.

With some functionalities, especially context-awareness, we were focused on overly complex solutions while user feedback proved that simple solutions were sufficient in most cases. Some functionalities turned out to be more complex than foreseen. Tag and sensor connection distance and

directionality need to be precise because proof-of-concept evaluations revealed that indefiniteness in them caused a confusing user experience. Once the connection has been set up, it should be kept on by tuning the reading distance.

The new user interaction paradigm of physical selection was well accepted as it enables the user easily to get information of the environment and near-by objects. Physical selection can also be used in user-initiated context recognition and in introducing personal sensor units to applications.

Ease of taking applications into use and ethical issues were repeated concerns of our interviewees and test users. Taking these issues into consideration at the architectural level helps to ensure that they are responded in all the applications and services built on the architecture. This will increase users' trust on the applications.

7. References

[1] E. Kaasinen, T. Tuomisto and P. Välkkynen, "Ambient functionality - use cases," in *Smart Objects Conference,* Grenoble, 2005.

[2] ISO 13407:1999, Human-centered design processes for interactive systems. International standard. International Standardization Organization, Geneve, Switzerland, 1999.

[3] MIMOSA project home page. www.mimosa-fp6.com

[4] S. D. Mainwaring, M. F. Chang and K. Anderson, "Infrastructures and Their Discontents: Implications for Ubicomp," in: N. Davies, E. Mynatt and I. Siio (eds.), *UbiComp 2004: Ubiquitous computing. 6th International Conference. Proceedings*, 2004, pp. 418-432.

[5] N. Streitz and P. Nixon, "The Disappearing Computer," in *Communications of the ACM* 48(3), 2005, pp. 33-35.

[6] G. Cockton, "Value-Centred HCI," in *Proceedings of the third Nordic Conference on Human-Computer Interaction (NordiCHI)*, ACM Press, New York, 2004, pp.149-160.

[7] G. Gay and H. Hembrooke, *Activity-Centered Design: An Ecological Approach to Designing Smart Tools and Usable systems.* The MIT Press, Cambridge, MA, 2004.

[8] N. Islam and M. Fayad, "Toward ubiquitous acceptance of ubiquitous computing," in *Communications of the ACM*, 46(2), 2003, pp. 89-92.

[9] P. Ketola, *Integrating usability with concurrent engineering in mobile phone development*, Academic dissertation, University of Tampere, Department of Computer and Information Science, A2002-5, ISBN 951-44-5359-x, 2002.

[10] W. K. Edwards, V. Bellotti, A. K. Dey and M. W. Newman, "Stuck in the Middle: The Challenges of User-Centered Design and Evaluation for Infrastructure," in *Proceedings of CHI2003*. ACM Press, New York, 2003, pp. 297-304.

[11] H. Ailisto, L. Pohjanheimo, P. Välkkynen, E. Strömmer, T. Tuomisto and I. Korhonen, "Bridging the Physical and Virtual Worlds by Local Connectivity-Based Physical Selection," in *Personal and Ubiquitous Computing*, Online First, 2006.

[12] R. Ballaghas, J. Borchers, M. Rohs and J. Sheridan, "The Smart Phone: A Ubiquitous Input Device," in *Pervasive Computing* 5(1), 2006, pp. 70-77.

[13] R. Want, K. P. Fishkin, A. Gujar and B. L. Harrison, "Bridging Physical and Virtual Worlds with Electronic Tags," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, New York, 1999, pp. 370-377.

[14] J. Riekki, T. Salminen and I. Alakärppä, "Requesting Pervasive Services by Touching RFID Tags," in *Pervasive Computing* 5(1), 2006, pp. 40-46.

[15] Välkkynen, P., Niemelä, M. and Tuomisto, T. Evaluating touching and pointing with a mobile terminal for physical browsing. Accepted to Proc. NordiCHI 2006.

[16] M. Friedewald, E. Vildjiounaite and D. Wright, *The brave new world of ambient intelligence.* Deliverable D1. A report of the SWAMI consortium to the European Commission under contract 006507. June 2005. Available: http://swami.jrc.es

[17] S. Lahlou, M. Langheinrich, C. Röcker, "Privacy and trust issues with invisible computers," in *Communications of the ACM* 48(3), 2005, pp. 59-60.

[18] M. Ohkubo, K. Suzuki and S. Kinoshita, "RFID privacy issues and technical challenges," in *Communications of the ACM*, 48(9), 2005, pp. 66-71.

COMPUTER
SOCIETY

Author(s)
Välkkynen, Pasi

Title
Physical Selection in Ubiquitous Computing

Abstract

In ubiquitous computing, the computing devices are embedded into the physical environment so that the users can interact with the devices at the same time as they interact with the physical environment. The various devices are connected to each other, and have various sizes and input and output capabilities depending on their purpose. These features of ubiquitous computing create a need for interaction methods that are radically different from the desktop computer interactions.

Physical selection is an interaction task for ubiquitous computing and it is used to tell the user's mobile terminal which physical object the user wants to interact with. It is based on tags that identify physical objects or store a physical hyperlink to digital information related to the object the tag is attached to. The user selects the physical hyperlink by touching, pointing or scanning the tag with the mobile terminal that is equipped with an appropriate reader. Physical selection has been implemented with various technologies, such as radio-frequency tags and readers, infrared transceivers, and optically readable tags and mobile phone cameras.

In this dissertation, physical selection is analysed as a user interaction task, and from the implementation viewpoint. Different selection methods – touching, pointing and scanning – are presented. Touching and pointing have been studied by implementing a prototype and conducting user experiments with it. The contributions of this dissertation include an analysis of physical selection in the ubiquitous computing context, suggestions for visualising the physical hyperlinks in both the physical environment and in the mobile terminal, and user requirements for physical selection as a part of an ambient intelligence architecture.

VTT PUBLICATIONS

645   Laitila, Arja. Microbes in the tailoring of barley malt properties. 2007. 107 p. + app. 79 p.

646   Mäkinen, Iiro. To patent or not to patent? An innovation-level investigation of the propensity to patent. 2007. 95 p. + app. 13 p.

647   Mutanen, Teemu. Consumer Data and Privacy in Ubiquitous Computing. 2007. 82 p. + app. 3 p.

648   Vesikari, Erkki. Service life management system of concrete structures in nuclear power plants. 2007. 73 p.

649   Niskanen, Ilkka. An interactive ontology visualization approach for the domain of networked home environments. 2007. 112 p. + app. 19 p.

650   Wessberg, Nina. Teollisuuden häiriöpäästöjen hallinnan kehittämishaasteet. 2007. 195 s. + liitt. 4 s.

651   Laitakari, Juhani. Dynamic context monitoring for adaptive and context-aware applications. 2007. 111 p. + app. 8 p.

652   Wilhelmson, Annika. The importance of oxygen availability in two plant-based bioprocesses: hairy root cultivation and malting. 2007. 66 p. + app. 56 p.

653   Ahlqvist, Toni, Carlsen, Henrik, Iversen, Jonas & Kristiansen, Ernst. Nordic ICT Foresight. Futures of the ICT environment and applications on the Nordic level. 2007. 147 p. + app. 24 p.

654   Arvas, Mikko. Comparative and functional genome analysis of fungi for development of the protein production host *Trichoderma reesei*. 100 p. + app. 105 p.

655   Kuisma, Veli Matti. Joustavan konepaja-automaation käyttöönoton onnistumisen edellytykset. 2007. 240 s. + liitt. 68 s.

656   Hybrid Media in Personal Management of Nutrition and Exercise. Report on the HyperFit Project. Ed. by Paula Järvinen. 121 p. + app. 2 p.

657   Szilvay, Géza R. Self-assembly of hydrophobin proteins from the fungus *Trichoderma reesei*. 2007. 64 p. + app. 43 p.

658   Palviainen, Marko. Technique for dynamic composition of content and context-sensitive mobile applications. Adaptive mobile browsers as a case study. 2007. 233 p.

659   Qu, Yang. System-level design and configuration management for run-time reconfigurable devices. 2007. 133 p.

660   Sihvonen, Markus. Adaptive personal service environment. 2007. 114 p. + app. 77 p.

661   Rautio, Jari. Development of rapid gene expression analysis and its application to bioprocess monitoring. 2007. 123 p. + app. 83 p.

662   Karjalainen, Sami. The characteristics of usable room temperature control. 2007. 133 p. + app. 71 p.

663   Välkkynen, Pasi. Physical Selection in Ubiquitous Computing. 2007. 97 p. + app. 96 p.