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
P.O.Box 1206
FIN-33101Tampere, Finland
+358 3 316 3111

[?]) VTT Information Technology
P.O.Box 1200
FIN-02044 VTT, Finland
+358 9 4561

*) VTT Electronics
P.O.Box 1100
FIN-90571 Oulu, Finland
+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

This paper was presented at "Physical Interaction (PI03) - Workshop on Real World User Interfaces", a workshop at the Mobile HCI Conference 2003 in Udine (Italy). September 8, 2003.
Further information and online proceedings are available at http://www.medien.informatik.uni-muenchen.de/en/events/pi03/

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).

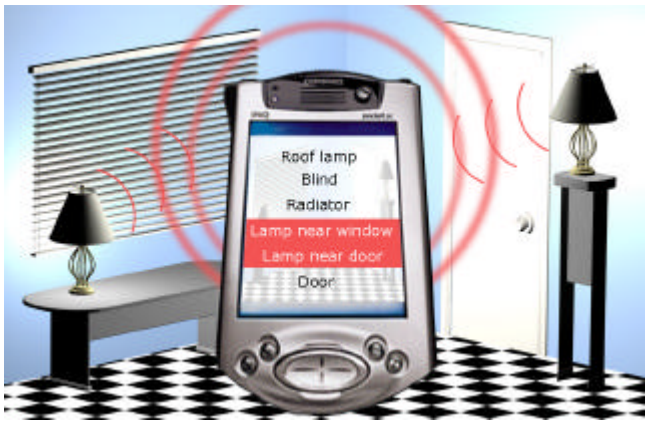

Figure 1: ScanMe

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, 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.

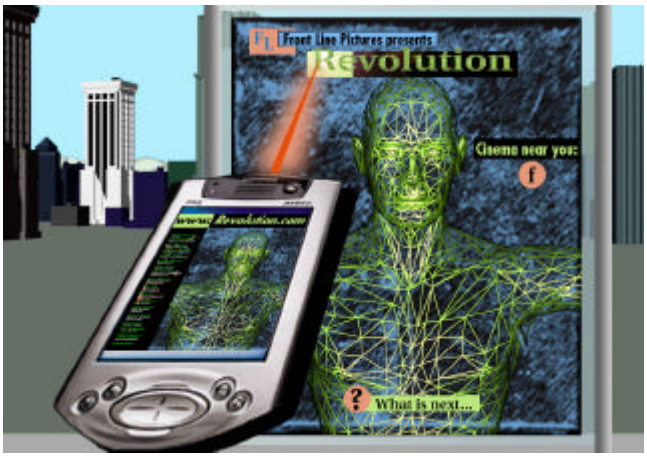

Figure 2: PointMe

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

[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 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, Juanles-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/