Real Time Inspection of Hidden Worlds

Alois Ferscha, Markus Keller
Department of Applied Computer Science, University Linz, Austria
ferscha@soft.uni-linz.ac.at, markus.keller@studnets.jku.at

Abstract

"Smart Things" are commonly understood as wireless ad-hoc networked, mobile, autonomous, special purpose computing appliances, usually interacting with their environment implicitly via a variety of sensors on the input side and actuators on the output side. Such smart appliances have started to populate the "real world" with "hidden" or "invisible" services, thus building up an "invisible world" of services associated with real world objects. With the embedding of invisible technology into everyday things, however, also the intuitive perception of "invisible services" disappears. We believe that it has potential advantages to support the perception of smart appliance services via novel interactive visual experiences. For this purpose we have developed and built DigiScope, a see-through based visual real time perception system for "invisible worlds" to support interactive theater experience in mixed reality spaces. A case study is presented that demonstrates the use of DigiScope to percept the "invisible services" of our smart Internet appliance SmartCase. Opposed to previous work on mixed reality based augmentation of reality, the DigiScope approach allows for a multiuser, collaborative real time perceptual experience.

1. Hidden Worlds

"The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" was Mark Weiser's central statement in his seminal paper [15] in 1991. His conjecture, that "we are 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" has fertilized not only the embedding of ubiquitous computing technology into a natural human environment which responds to people's needs and actions in a contextual manner, but has also caused "hidden" functionality and services volatilize out of sight of humans. "Smart Things" functionality is characterized by the autonomy of their programmed behaviour, the dynamicity and context-awareness of services and applications they offer, the ad-hoc interoperability of services and the different modes of user interaction upon those services [3]. Since many of these objects are able to communicate and interact with global networks and with each other, the vision of "context-aware" [2] smart appliances and smart spaces – where dynamically configured systems of mobile entities by exploiting the available infrastructure and processing power of the environment – has become a reality. Common in the trend of context aware environments is that they interact with the user in a pro-active, autonomous, sovereign, responsible and user-authorized way. Common is also that the provision of their services is based on their ability of being aware of the presence of other objects or users, and being sensitive, adaptive and responsive to their needs, habits and emotions. Their services tend to become ubiquitously accessible via natural interaction. Embodied into real world objects like furniture, clothing, crafts, rooms, etc., those services are usually "invisible", i.e. cannot be perceived visually or orally, thus leaving the user "un-aware" of their presence. It must be assumed, that services carried by these kinds of ubiquitous computing technology in a mostly invisible manner, is considered and perceived by human user to be "not there" just because they cannot be seen, i.e. because they are invisible. In certain cases of ubiquitous computing therefore it will be desirable, to at least visualize invisible services to the user. We believe that the utility of many of those "invisible services" of smart appliances being better exploited when presented to the user in a more intuitive and natural way, thus raising the need for a better perception of smart environments by the user. To support people living in the real world populated with a variety of digital artefacts as created by the digital components in a smart environment, when acting, perceiving and interacting with objects in their environment, we propose a see-through based theatre experience of visual perception, seamlessly merging the artefacts of the real and the digital world. After discussing related work (Section 2) and motivating our approach (Section 3), we present the DigiScope system (Section 4), a 6DOF visual see-through tablet that we have developed to support "invisible world" inspection. The runtime environment is outlined in Section 5. In Section 6 we present use cases of DigiScope for "hidden world" applications that we have developed based on our contextware framework [4] .

2. Related Work

Comparable oncepts for annotating real world objects to be perceived in an augmented or mixed reality setting via the see-through metaphor have appeared in the recent literature, mostly based on Head-Mounted-Display (HMD) equipment.

One of the first approaches avoiding HMDs is MEDARPA [11][13], a medical augmented reality (AR) application, which supports the surgeon during operation. The support consists of the presentation of “missing information” about the patient’s anatomy during surgery. The “missing information” is a virtual model of the patient’s anatomy, which is based on acquired information (CT, MRT). This information is useful when the physician inserts an instrument into the patient’s body without seeing the actual anatomy under the skin. The use of AR augments the physician’s view with the virtual model. The realization of MEDARPA’s AR window is a pivoted transparent display based on a 17” active-matrix LCD (TFT) screen allowing a resolution of 1024 x 768 at 75 Hz. Due to the available display technology the transparency is restricted, but if sufficient light can be supplied to the observed scene (the patient’s body) the display meets the transparency requirements. This approach allows the physician to see hidden information about the patient through a window (or tablet) like device.

A similar approach, but with the option of controlling the virtual viewpoint, is followed in the project Boom Chameleon [14]. The output and also the control device is an LCD screen, which represents the virtual eye perceiving a virtual reality world. Moving the screen in space by hand changes the virtual viewpoint accordingly in real time.

Yet another “window” based approach is Augurscope [12], which uses an LCD tablet screen for presenting the virtual objects annotated to the scene. A GPS receiver and an electronic compass is used for position and orientation tracking of the screen, with an accuracy of about two to four meters. The compass data has a typical accuracy of about 1°. Also a delay between moving the device and receiving a position update of more than a second is mentioned, which makes real time interaction almost impossible. Opposed to Augurscope, which attempts outdoor usage scenarios, DigiScope is addresses indoor use and employs an ultrasonic tracking device to obtain its position and orientation at an update rate of 180Hz.

A hand held video based see-through system is the AR Pad [8], which utilizes a small LCD screen for video output. The AR Pad is able to view AR environments and it is also capable to interact with virtual objects in a collaborative and intuitive way. Opposed to AR Pad, which implements a video based see-through solution, DigiScope utilizes the more intuitive optical see-through system approach.

3. Inspecting the Invisible in Real Time

Generally, more than the previous approaches, the DigiScope aims at supporting “human to ubiquitous computer interaction” processes by bringing back visual clues to the user on how to interact. It is motivated not by the requirements of a particular application, like surgery in the MEDARPA case or medieval castle inspection Augurscope, but most generally by the user needs raising with the growth and popularity of services in the pervasive and ubiquitous landscape: Once computers have disappeared from desks, hiding in the background, their services will most likely still be there. New artefacts and smart appliances [10] are evolving that “carry” invisible services, such that manipulating the appliance controls a service. Even if the service is not integrated into the artefact but merely “linked” to a background system [7], the manipulation of the physical object can manipulate their virtual representative on that background system respectively. To this end it is necessary to link the physical world with the virtual world [4], i.e. the linking of physical objects with their “virtual counterparts” [9]. Tangible interface research [6] has contributed to this issue of physical-virtual linkage by considering physical artefacts as representations and controls for digital information. A physical object thus represents information while at the same time acts a control for directly manipulating that information or underlying associations. Examples for approaches were input and output are fused into physical object manipulation include architecture and landscape design and analysis, object shape modeling interfaces using Lego-type blocks or triangular tiles, or cubes used as a bi-directional user interfaces.

A common problem of tangible interfaces is the lack of visual clues of the kind of services hidden in an artifact and how to access or use them. Since not all of those artefacts are designed (or even able) to provide their own visual interface, an annotation of the related physical objects with digital information appears appropriate. The visualization of those annotations can be done in a situative way by means of virtual or mixed reality technologies. Mixed reality (MR) environments according to Milgram’s classification are those in which real world and virtual world objects are presented together on a single display. It covers the whole reality-virtuality spectrum, involving the physical reality, augmented reality and virtual reality.

The implementation of visual interfaces for MR applications usually involves the use of output devices that allow users to percept the world via coalesced physical and artificial views. The most prominent technologies to implement such interfaces today are video- and optical see-through systems, realized either as head-mounted- (HMD) or non-head-mounted displays. User experience however has revealed that such devices are very unnatural and klutzy in handling. We therefore follow an approach that allows for an obstruction-free perception of ubiquitous computing rich environments. Much like Paradiso's metaphor of a flashlight to "find" (tagged) objects in the "darkness", and Koleva's metaphor of a handheld position and orientation sensitive 3D environment inspection pad [1], we propose a 6DOF optical see-through inspection tablet, DigiScope, to provide a seamless interface between real and digital artifacts.

We exploit the metaphor of digital annotations for real world objects, and display these annotations along the line of sight to real world objects that are seen through a holographic display. The user gets the ability to interact with the virtual object and its digital information by viewing the corresponding real (physical) artefact. With DigiScope, the user is handling a holographic display tablet just like a 6DOF window that opens a view into the virtual world. The tablet is an optical see-though display which allows for a very natural viewing and scene inspection. To implement correct views in real time, the angle and perspective of the DigiScope is being tracked, instead of tracking the position and orientation of the user. Thus the user is freed from any system hardware obstacles like HMDs, stereoscopic glasses, trackers, sensors, markers, tags, pointers and the such.

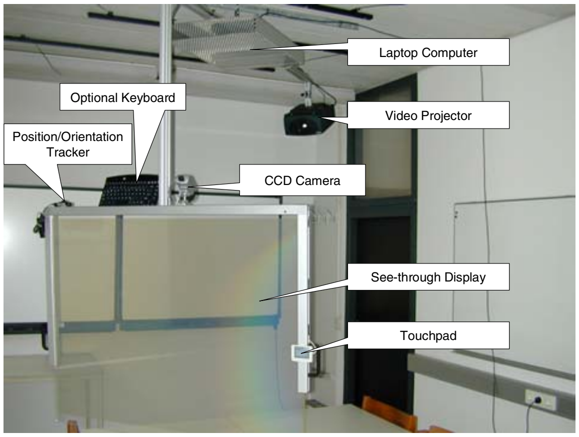

Figure 1. The DigiScope System

To support free navigation in the scene, the DigiScope can be fully tilt and rotated in space by hand, delivering the look and feel of a steerable window to perceive the annotations associated with the real world objects in the focused scene. Several design options – beyond the current DigiScope appearance – have been investigarted, ranging from wearable (microptical) displays to holographic tablets, from stationary (ceiling mounted) to mobile (wheel chair) settings.

4. DigiScope Hardware Components

To support a free choice of views into a scene, the DigiScope can be fully tilt and rotated in space by hand. A projecting beamer is fixed in a proper projecting angle on top of a 6DOF mounting frame (see Figure 1), and is used to project the computer generated image encoding the scene annotation onto a holographic display. For tracking the position and orientation of the DigiScope frame, an Intersense IS-900 tracker (www.isense.com) is used. The mounting frame holds also a video camera (Sony DFW-X700) for optical marker recognition and a touchpad for user interactions. All hardware components are connected to an integrated notebook computer (See Figure 2). As such, DigiScope appears as a stand-alone, autarkic system.

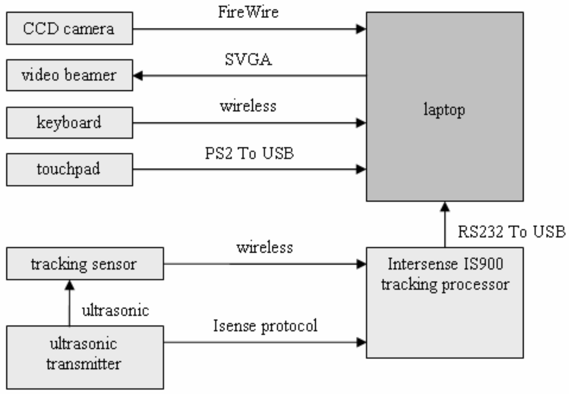

Figure 2. DigiScope System Layout Diagram

5. The DigiScope Runtime System

The runtime system of the DigiScope is able to identify two types of physical objects in the scene in real time (while the spectator is viewing the scene). It recognizes movable objects (not fixed in position), so called *dynamic objects*, and *static objects* which have a fixed position in the scene. Physical objects are split up into these two classes, because they are handled with different approaches and technologies. For dynamic objects, identification is implemented based on visual markers and marker recognition at runtime, for static objects RFID based idetification mechanisms have been implemented. This choice is just for technological convenience of the scenarios DigiScope is handling at the moment, and can be modified or replaced by other identification technologies at any time.

Visual markers attached to dynamic objects are captured with the CCD camera mounted on the DigiScope. Marker recognition and identification is processed by ARToolkit (www.hitl.washington.edu). ARToolkit retrieves the position and orientation of the identified visual (optical) marker relative to the viewpoint position. Static objects are not tagged with optical markers and thus can not be recognized by the ARToolkit. Static objects are fixed in space and are registered with their physical position in the virtual world scene. For this purpose, every static object holds a configuration file with the respective real world coordinates. The postion data file is loaded into a Java3D based virtual scene model at start-up time. Equipped with the positions of all static objects and the position and orientation of the DigiScope, the runtime system calculates the view into the annotated scene and displays it on the Holo screen. In the scene model, the virtual representatives of physical world objects are overlaid with their physical counterparts, their annotations can be displayed in close spatial proximity and deliver the realistic impression of digital annotations to the reality.

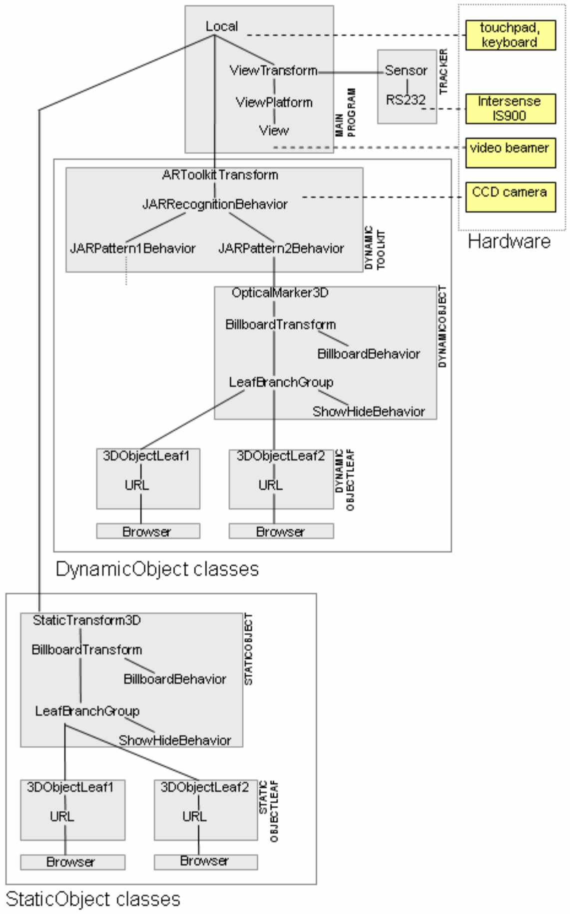

Figure 3. Classes and Hardware Interfacing

Within the DigiScope runtime systems, physical objects are as well classified into by dynamic and static objects and modelled by the classes *DynamicObjects* and *StaticObjects*. (see Figure 3). shows the software structure of the DigiScope including the classification of *DynamicObject* classes and *StaticObject* classes. The runtime system has its entry point in the *MainProgram*, where the virtual model is created. It includes the creation of the *Universe*, the *Local*, the *ViewTransform*, the *View* (virtual viewpoint), *Behaviors* for user interactions and the 3D rendering engine, which provides the 3D scene for the prohjection subsystem. The *ViewTransform* encapsulates the transformations of the static objects path, which is directly manipulated by the *Tracker* class. Altering the *ViewTransform* changes accordingly the attached *View*, which is the virtual viewpoint in the 3D scene. The *Tracker* class has the task to communicate with the Intersense IS900 processor via a serial port. The virtual representatives of the static objects are linked directly to the *Local*. Therefore manipulating the *ViewTransform*, changes the position of the *StaticObjects* relatively to the virtual eye. Moving the tracked DigiScope is changing the *ViewTransform* and thus the virtual representatives of the static objects are moving on the Holo screen accordingly and are always presented along the line of sight of the spectator.

DynamicObjects are not manipulated by the *Tracker* class. They are controlled by the *DynamicToolkit* class which utilizes the ARToolkit and interfaces the CCD camera to capture the live image of the view through the DigiScope screen. Because each dynamic object can be moved individually, each of them has its own *OpticalMarker3D* transformation, which is the 6DOF information of the optical marker relative to the camera position. Therefore each *DynamicObject* gets its own *JARPatternBehavior* within the *DynamicToolkit*. A *JARPatternBehavior* is an individual software thread that can react to a position change of an optical marker instantly and in real time. With this approach, the virtual *DynamicObjects* can move in the 3D scene accordingly to the physical ones without moving the DigiScope and without interfering the virtual presentations of the static objects. *DynamicObjects* can move on the screen, while the *StaticObjects* remain at their location.

Annotations to both static and dynamic objects can rang from simple text, to image, to any kind of a 3D geometry model (e.g. represented in VRML). If an annotation is displayed, and if the user is interested in the information "hyper"-linked to that object, he can select the annotation with the cursor by using the touchpad. The selected annotation pops up a menu with different options (*StaticObjectLeafs*, *DynamicObjectLeafs*). Selecting a menu item displays the information linked to, which can be any WWW reference like the URL to a web page, an online manual, a multimedia document or web service.

Scene annotations can be configured for each real world object separately in individual XML files. The configuration files for *DynamicObjects* and *StaticObjects* are very similar and loaded at start-up time. Because a dynamic object is associated with an optical marker and a static object is fixed in space, their XML files contain different content. For example, an XML configuration file for a *DynamicObject* contains the element tag `<PATTERNNAME>` which holds the path to the optical marker pattern file. This line is linking the physical optical marker to the annotation.

```
<DYNAMICOBJECT>
   <NAME>DisplayedName</NAME>
   <PATTERNNAME>Data\\patt.kanji</PATTERNNAME>
   ...
</DYNAMICOBJECT>
```

An XML file for a *StaticObject* contains element tags to specify the location of the static physical object wrt. to its cartesian coordinates in the scene:

```
<STATICOBJECT>
   <NAME>DisplayedName </NAME>
   <X>1000.0</X>
   <Y>300.0</Y>
   <Z>2500.0</Z>
   <ROTX>0.0</ROTX>
   <ROTY>0.0</ROTY>
   <ROTZ>90.0</ROTZ>
   ...
</STATICOBJECT>
```

6. Inspection Scenarios

To illustrate the performance of DigiScope, a few scenarios will be presented in this section. A first impression of the look and feel of DigiScope, the appearance of annotations and the navigation through the associated hyperlinks is given in Figure 4.

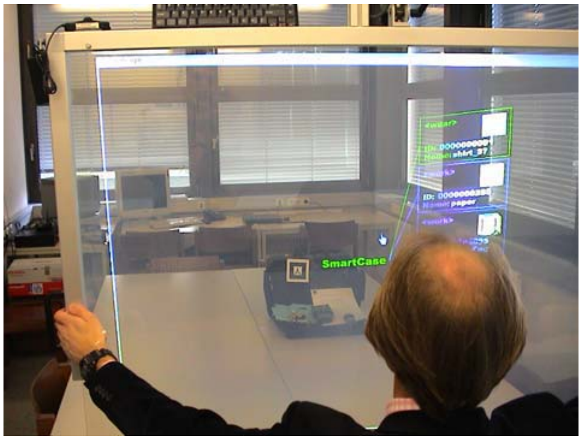

Figure 4. Inspection demonstration

The user is controlling the DigiScope by hand, and viewing the scene from different perspectives through the Holo screen by steeering the DigiScope. The scene observed in Figure 4 includes a suitcase tagged with an optical marker. The camera which captures the visual marker invokes the *DynamicToolkit*, which calculates the position and orientation of the suitcase. The identified pattern on the optical marker is configured to the *DynamicObject* of the suitcase, and associated with the annotation containing the text “SmartCase”. If the suitcase is moved in the scene, the text of the annotation will move accordingly. A list with the suitcase inventory appears after the user has selected and clicked the “SmartCase” hyperlink with the touchpad. Each item of the inventory list – again represented as a hyperlink annotation – can be selected separately to get more detailed information about the desired inventory object.

The scenario in Figure 5 shows how the geometry of cubical annotation is perceived via the DigiScope. A colored cube geometry is transformed (translated, rotated and scaled) according to the rotation of the DigiScope. A wireframe grid, aligned with the walls of the room, helps to perceive that the cube has a fixed position relative to the physical scene. In the sequence of pictures (1-6) the DigiScope performs a 180° rotation.

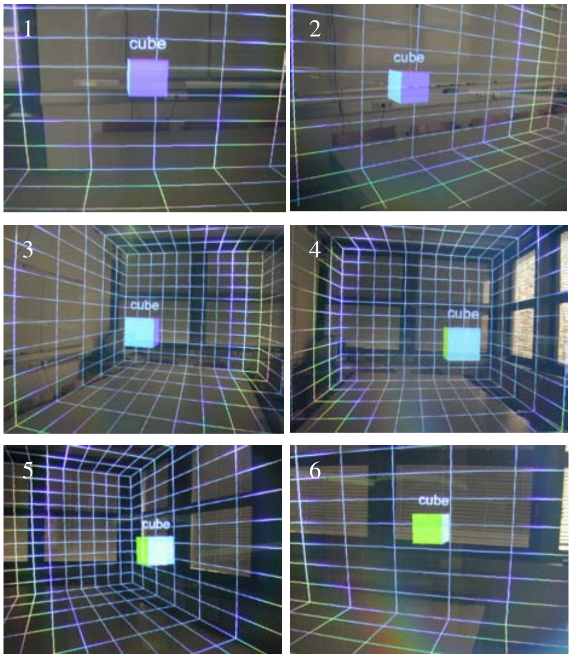

Figure 5. Registration and Perspective

The scenario shown in Figure 6 demonstrates the use of a hyperlink annotation for a static object, in this case the base station of the Intersense IS900 tracking device in our lab. This device is statically located on a desk, the text of its annotation is “IS900”, the hyperlinks associated with this annotation are organized in a two item menu, one is “Manual”, a Web reference to a pdf document residing at the manufacturers homepage, the second is “Tracker values”, a link to the file system of the tracking device.

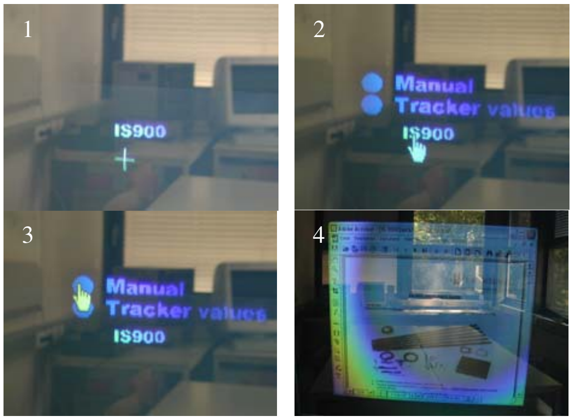

Figure 6. Annotation of a Static Object

In this example, DigiScope offers the ability to interact with the annotation by controlling the mouse cursor via the touchpad. The mouse cursor appears usually as a small cross, and swaps to the hand symbol when the user is going to select the “IS900” annotation. A single in this node brings up the bullet point list of the menu. Clicking on the “Manual” item will spawn an internet browser instance referring to the web site which is linked with the annotation.

Figure 7. Dynamic Objects: Annotating Persons

Figure 7 demonstrates the ability of the framework to link dynamic (physical) objects to a DigiScope scene. Optical markers are printed on shirts, allowing the DigiScope to track position and orientation of moving persons, while displaying appropriate annotations, like e.g. a persons private web page. Here, the annoation of a person is just a hyperlinked text object showing that persons name. The sequence of pictures (1-4) demonstrates how two dynamic objects are simultaneously tracked based on optical marker recognition.

A more complex scenario is developed in Figure 8 , which refers to our previous work, the development of SmartCase [4]. SmartCase is an exaple of a context aware smart appliance, providing a variety of “invisible services” to the user.

The hardware for the SmartCase demonstration prototype uses an embedded single board computer integrated into an off-the-shelf suitcase, which executes a standard TCP/IP stack and HTTP server, accepting requests wirelessly over an integrated IEEE802.11b WLAN adaptor. A miniaturized RFID reader is connected to the serial port of the server machine, an RFID antenna is integrated in the frame of the suitcase so as to enable the server to sense RFID tags contained in the SmartCase. A vast of 125KHz and 13,56 MHz magnetic coupled transponders are used to tag real world objects (like shirts, keys, PDAs or even printed paper) to be potentially carried (and sensed) by the suitcase.

A unique ID associated with every real world object is the ID encoded in its RFID tag. It is sensed by an RFID reader which triggers a script to update the state information on the embedded Web server. Considering now the inventory of the SmartCase as an “invisible” service, then, once an object (e.g. shirt) has been put into the SmartCase, this service can be queried to check whether the shirt is in the case or not. A straightforward way to access this information would be via a classical http interface to the embedded web-server. Observed via the DigiScope however, changes to the SmartCase inventory are displayed as a graphical annotation of the real world. Figure 8 illustrates a sequence of user interactions with objects in the real world and the corresponding state changes in the augmented world representation as seen through the DigiScope. In (1) the SmartCase and the tagged objects are shown. (2) shows the SmartCase annotated with the hyperlink “SmartCase” on the DigiScope display. A shirt is put into the SmartCase. RFID sensors detect the presence of the shirt and immediately after the digital counterpart of the shirt is shown as content of the SmartCase on the DigiScope (3). Similarly, the action of putting a paper document into the SmartCase is detected by the system and consequently a hyperlink (Web link) to the paper appears on the DigiScope display. After adding further items (PDA, a bunch of keys), annotations for all objects are visible and can be manipulated (6).

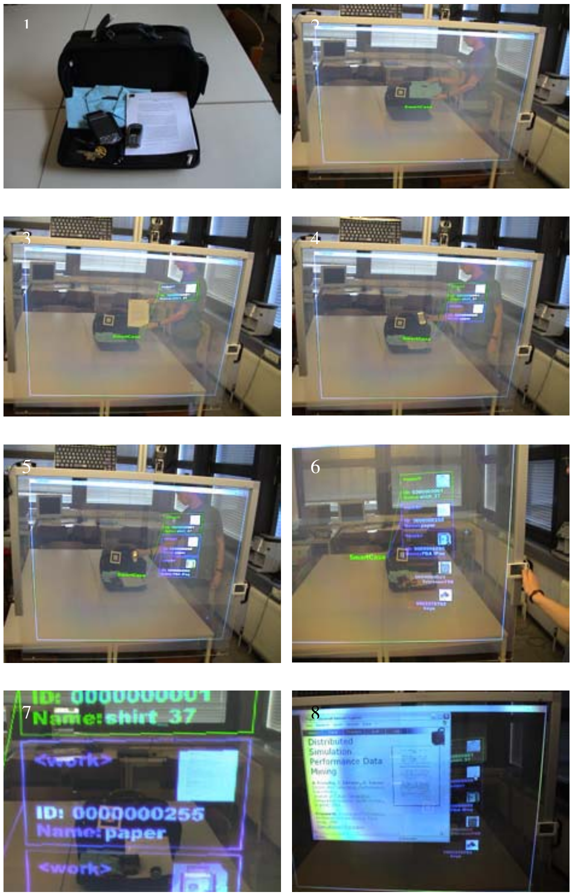

Figure 8. Packing and linteracting with the SmartCase

A major feature of the DigiScope hyperlink annotation concept when accessing "hidden" services is the seamless access to the WWW. Figure 8 illustrates how the inventory of the packed SmartCase is accessed via the URL links associated with each item in the inventory list. The user is interested in the paper document (item 2 in the list), and thus points the cursor to the URL link in the annotation and clicks on it (7). This event is captured by the DigiScope and as a response, a Web browser is spawned displaying html related to the paper ID, which contains a Web link to the pdf document itself (8). Browsing this link spawns a pdf viewer allowing to interact with the digital counterpart of the physical document in the usual manner.

7. Conclutions

The emerging problem of developing intuitive interfaces for the perception and inspection of natural human environments populated with an increasing number of smart appliances in the pervasive and ubiquitous computing landscape together with their "invisible" services has been approached. Our DigiScope envisions a new type of MR interface with two main features: (i) a new exploration experience of the physical world seamlessly merged with its digital annotations via a non-obtrusive MR interface, and (ii) an integration of ubiquitous context-awareness and physical hyperlinking at the user interface level. Related to comparable real world annotation approaches, e.g. based ond CAVE or head mounted display technologeies, DigiScope works in real time and enables even multiuser collaborative scene inspection.

8. References

[1] S. Benford, H. Schnadelbach, B. Koleva, B. Gaver, A. Schmidt, A. Boucher, A. Steed, R. Anastasi, C. Greenhalgh, T. Rodden, and H.W. Gellersen, "Sensible, sensable and desirable: a framework for designing physical interfaces", Technical Report Equator-03-003, Equator, 2003.

[2] A.K. Dey, "Understanding and Using Context" Personal and Ubiquitous Computing, Special Issue on Situated Interaction and Ubiquitous Computing, Vol. 5 No. 1, 2001, pp 4-7.

[3] D. Estrin, D. Culler, K. Pister, and G. Sukhatme, "Connecting the Physical World with Pervasive Networks", Pervasive Computing, Vol. 1 No. 1, 2002, pp. 59-69.

[4] A. Ferscha, "Contextware: Bridging Virtual and Physical Worlds", Reliable Software Technologies, AE 2002, LNCS 2361, Berlin, 2002, pp. 51-64.

[5] H.W. Gellersen, M. Beigl, and A. Schmidt, "Sensor-based Context-Awareness for Situated Computing", Proc. of Workshop on Software Engineering for Wearable and Pervasive Computing, Limerick, Ireland, June, 2000, pp. 77-83.

[6] M.G. Gorbet, M. Orth, and M. Ishii, "Triangles: Tangible Interface for Manipulation and Exploration of Digital Information Topography", Proc. CHI1998, 1998, pp. 49-56.

[7] T. Kindberg, and A. Fox, “System Software for Ubiquitous Computing”, IEEE Pervasive Computing, Vol. 1 No. 1, 2002, pp. 70-81.

[8] D. Mogilev, K. Kiyokawa, M. Billinghurst, and J. Pair, “AR Pad: An Interface for Face-to-face AR Collaboration”, Proceedings of the ACM Conference on Human Factors in Computing Systems 2002 (CHI '02), Minneapolis, 2002, pp. 654-655.

[9] K. Römer, T. Schoch, F. Mattern, and T. Dübendorfer, “Smart Identification Frameworks for Ubiquitous Computing Applications”, Proc. PerCom, 2003, pp. 253-264.

[10] A. Schmidt, and K. Van Laerhoven, “How to Build Smart Appliances?”, IEEE Personal Communications 8(4), 2001, pp. 66-71.

[11] M. Schnaider, H. Seibert, B. Schwald, T. Weller, S. Wesarg, and P. Zogal, „Medarpa - Ein Augmented Reality System für Minimal-Invasive Interventionen”,2. Int. Statustagung "Virtuelle und Erweiterte Realität", Leipzig, 2002.

[12] H. Schnädelbach, B. Koleva, M. Flintham, M. Fraser, P. Chandler, M. Foster, S. Benford, C. Greenhalgh, S. Izadi, and T. Rodden, “The Augurscope: A Mixed Reality Interface for Outdoors”, Proceedings CHI 2002, ACM Press, Minneapolis , USA, 2002, pp. 9-16.

[13] B. Schwald, F. Seibert, and T. Weller, “A Flexible Tracking Concept Applied to Medical Scenarios Using an AR Window”, Poster at ISMAR2002, 2002.

[14] M. Tsang, G.W. Fitzmaurice, G. Kurtenbach, A. Khan, and B. Buxton, “Boom Chameleon: Simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display”, ACM CHI Letters Vol. 4, No. 2, 2002, pp. 111-120.

[15] M. Weiser, “The Computer of the Twenty-First Century”, Scientific American, 1991, pp. 94-100.